XoveTIC: Impulsando el talento científico

Proceedings XoveTIC
2023

October 2023

Editorial Board
Manuel Lagos Rodríguez
Álvaro Leitao Rodríguez
Tirso Varela Rodeiro
Javier Pereira Loureiro
Manuel Francisco González Penedo

Editorial Board

Manuel Lagos Rodríguez
Universidade da Coruña
Spain

Álvaro Leitao Rodríguez
UOC
Spain

Tirso Varela Rodeiro
Universidade da Coruña
Spain

Javier Pereira Loureiro
Universidade da Coruña
Spain

Manuel Francisco González Penedo
Universidade da Coruña
Spain

Editorial Office

Servizo de Publicacións (SPU)
Universidade da Coruña
Maestranza 9, 15001 A Coruña, Spain

ISBN XXX-XXXXX-XXX-X (Hbk)
ISBN XXX-XXXXX-XXX-X (PDF)

Cover image courtesy of CITIC—Research Center of Information and Communication Technologies, University of A Coruna, Spain

Proceedings XoveTIC 2023. Impulsando el talento científico.

© 2023

This volume, including all its contents, is licensed under a Creative Commons Attribution 4.0 International license, and made available in Open Access at https://xovetic.citic.udc.es/. The authors of the individual contributions, who are identified as such, retain the copyright over their original work.

For more information on the CC BY 4.0 license, please refer to:
https://creativecommons.org/licenses/by/4.0/deed.en.

This volume was typeset in LaTeX by the editors using *varianTeX* — a reusable template for journals in the Humanities, developed by Wout Dillen. *varianTeX* is open source, available on GitHub, and deposited in the Zenodo Open Science Repository. DOI: 10.5281/zenodo.3484651.

Contents

1 ***Jacques.B. Ngoua-Ndong Avele, and Marc Bernice Angoue Avele***: Simulation of a Unmanned Aerial Vehicle (UAV) Swarm Control System **3**

2 ***Marc Bernice Angoue Avele, Darian Pérez-Adán, and Dariel Pereira-Ruisánchez***: Eavesdropping and Jamming via Pilot Attacks in 5G Massive MIMO **9**

3 ***Sergio Marcos-Vazquez, Carlos Fernandez-Lozano, Adrian Carballal, and Francisco Cedron***: Music Recommendation System Based on Ratings Obtained from Amazon **17**

4 ***Raquel Simões de Almeida, Paulo Veloso Gomes, António Marques, and Javier Pereira*** : Forest Bathing Environments in Virtual Reality: Enhancing Well-being through Immersive Nature Experiences **25**

5 ***Patrícia Silva, Paulo Veloso Gomes, Jorge Condeço, António Marques, Henrique Curado, Rita Veloso, and Javier Pereira***: m-Health System for the Self-management of Patients with Heart Failure **31**

6 ***Hilda Romero-Velo, Susana Ladra, José R. Paramá, and Fernando Silva-Coira***: Indexing and Retrieval of Scores by Humming based on Extracted Features **35**

7 ***Paulo Veloso Gomes, Henrique Curado, António Marques, Bárbara Gomes, and Javier Pereira***: Distributed Database Model for Mobile Health Telemonitoring Applications **43**

8 ***Daniel Garcia-Gonzalez, Enrique Fernandez-Blanco, Daniel Rivero, and Miguel R. Luaces***: Introducing a Human Activity Recognition Dataset Gathered on Real-Life Conditions **47**

9 ***Julián Montoto-Louzao, Alba Camino-Mera, María J Martín, and Antonio Salas***: GUANIN: GUi-driven Analyzer for NanoString Interactive Normalization **55**

10 ***Renato Magalhães, Rita Veloso, Paulo Veloso Gomes, António Marques, and Javier Pereira***: Extended Reality in the Operating Room: Robot-Assisted Orthopedics Surgery with Live and Interactive Streaming for Medical Students **63**

11 ***Renato Magalhães, Rita Veloso, Adélio Vilaça, Ana Oliveira, and David Terroso***: Applicability of Mixed Reality in Operating Room **69**

12 ***Roberto Prado-Rodríguez, Patricia González, Julio R. Banga, and Ramón Doallo***: Self-adaptive Cooperation Scheme in a Parallel ACO Algorithm for Binary Combinatorial Problems **75**

13 ***José Domingues, Paulo Veloso Gomes, Raquel Simões de Almeida, Constantino Martins, Nuno Fonseca, Joaquim Santos, António Marques, and Javier Pereira***: Self-management Web Application for Persons with Schizophrenia **83**

14 ***Lucía Álvarez-Crespo, and Laura M. Castro***: Unveiling the Dark Side of Social Media: Developing the First Galician Corpus for Misogyny Detection on Twitter and Mastodon **87**

15 ***Alberto Fernandez Sanchez, Juan Ramon Rabuñal Dopico, Daniel Rivero Cebrian, Alejandro Celestino Pazos Sierra, Marcos Gestal Pose, and Luis Cea Gomez***: Predicting Inflow Flow in Hydraulic Dams Using Artificial Neural Networks **91**

16 ***Ángel Niebla-Montero, Paula Fraga-Lamas, José Varela-Barbeito, and Tiago M. Fernández-Caramés***: Ultra Wideband for Shipyard 5.0 Indoor Gantry Crane High-Precision Positioning **97**

17 ***Gabriel Fernández-Blanco, and Tiago M. Fernández-Caramés***: A Validation System for Academic Records based on Blockchain **103**

18 ***Naomi Diz-Rosales, María José Lombardía, and Domingo Morales***: Mapping the Poverty Proportion in Small Areas under Random Regression Coefficient Poisson Models **109**

19 ***Delfina Ramos-Vidal, and Nieves R. Brisaboa***: Efficient Database Evolution in Digital Library Reengineering **117**

20 ***David Soler, Iván Cillero, Francisco J. Nóvoa, Carlos Dafonte, Ana Fernández Vilas, and Manuel Fernández-Veiga***: Analysis of the NS-3 module QKDNetSim for the Simulation of QKD Networks **121**

21 ***Iago Fernández-Garrido and Ángel Gómez***: Virtual Reality for Cognitive Stimulation in the Elderly and Individuals with Disability **129**

22 ***Noé Vila, Paula M. Castro, and Adriana Dapena***: Reinforcement and Homework Control for Children with ASD using a Mobile Application for Apple Watch **135**

23 ***Emilio López-Varela, Noelia Barreira, Nuria Olivier Pascual, and Manuel G. Penedo***: Enhancing Pathological Detection and Monitoring in OCT Volumes with Limited Slices using Convolutional Neural Networks and 3D Visualization Techniques **143**

24 ***Julio Pombo-Romero, Oliver Rúas-Barrosa, and Carlos Vázquez***: Valuation and Risk Modeling of Renewable PPAs **151**

25 ***Alberto Pérez-Aboli, Carlos Dafonte, and Ángel Gómez***: Embedded System for Monitoring and Prevention of Hardware/Software Incidents in Data Center Devices **157**

26 ***Manuel Lagos Rodríguez, Nereida Canosa, and Thais Pousada***: Integration of Captive Movement Systems in Virtual Environments for Rehabilitation in People with Disabilities **163**

27 ***Carlos Andrés Correa-Guillén, Ángel Gómez, and Estefanía López-Salas***: A System for the Management, Visualization and Annotation of Digital Collections of Handwritten Historical Documents **169**

28 ***José Bobes-Bascarán, Ángel Fernández Leal, Eduardo Mosqueira-Rey, David Alonso-Ríos, Elena Hernández-Pereira, Vicente Moret-Bonillo***: Understanding Machine Learning Explainability Models in the context of Pancreatic Cancer Treatment **175**

29 ***Blanca E. Monroy-Castillo, M.A. Jácome, and Ricardo Cao***: Estimation of Distance Correlation: a Simulation-based Comparative Study **183**

30 ***María Isabel Limaylla-Lunarejo, Nelly Condori-Fernandez, and Miguel R. Luaces***: Towards a FAIR Dataset for Spanish Non-Functional Requirements **191**

31 ***Julián González-Muñoz, Mario Casado, Daniel Garabato, Francisco J. Nóvoa, and Carlos Dafonte***: Hardening of a Continuous Behavior-based Authentication Distributed System **197**

32 ***María A. Baamonde-Seoane, María del Carmen Calvo-Garrido and Carlos Vázquez***: Method for Pricing Renewable Energy Certificates **203**

33 ***Alfonso Torralba Mantiñán, Cristina López Bravo, José Luis Rivas López***: Development and forensic study of a ransomware for Android 8.1 devices. **213**

34 ***Mario Noriega, José Ribelles, Antonio Seoane, Margarita Martínez, Javier Taibo, and José A. Iglesias-Guitián***: Highway Remodeling: Harnessing Georeferenced Data for Procedural Modeling **219**

35 ***Elsa Martin-De Arribas, Kelly Conde-Pérez, Pablo Aja-Macaya, Juan A. Vallejo, Margarita Poza, and Susana Ladra***: Automation Proposal for the Intermediate Steps in the 16S FFPE Samples Analysis Pipeline **227**

36 ***Francisco Prado-Valiño, Roi Santos-Ríos, Carlos Gómez-Rodríguez, and Jesús Vilares***: Prototype of an Entity Recognition System for Antimicrobial Resistance Data Management. **235**

37 ***Patricia Concheiro-Moscoso, Betania Groba, María del Carmen Miranda-Duro, and Diego Alvarez-Estevez***: Performance of Xiaomi Mi Band 5 Against Polysomnography in Measuring Sleep **243**

38 ***Davide Trevisani, José Germán López, Chris Kenyon, Carlos Vázquez, and Mourad Berrahoui***: Rationale and Design of a Scope 3 Capital Charge **249**

39 ***João Donga, Paulo Veloso Gomes, António Marques, and Javier Pereira***: Interactive Virtual Reality Public Speaking Training Simulator to Empower Confident Communicators **257**

40 ***Pablo García-Rivada, Paula Fraga-Lamas, and Tiago M. Fernández-Caramés***: An Intelligent IoT Wearable for Monitoring and Preventing Asthma Attacks **263**

41 ***João Donga, João Azevedo, Marta Carrapa, and Paulo Veloso Gomes***: Reimagining Art: Virart, a Platform Using Virtual Reality for Creative Expression and Cultural Engagement **269**

42 ***Paulo Veloso Gomes, Vítor J. Sá, João Donga, António Marques, Bárbara Gomes, Raquel Simões de Almeida, and Javier Pereira***: The Use of Artificial Intelligence in Interactive Virtual Reality Adaptive Environments with Real-Time Biofeedback Applied to Phobias Psychotherapy **275**

43 ***Daniel J. García-Paz, Amparo Alonso-Betanzos, Bertha Guijarro-Berdiñas, and Alejandro Rodríguez-Arias***: Simulation of Virus Propagation and Acceptance of Socio-Sanitary Measures Through an Intelligent Model **281**

44 ***Andres Molares-Ulloa, Maria del Rocio, Julian Dorado, Nieves Pedreira, and Alvaro Rodriguez***: Re-Identification of Rats with Transfer Learning **289**

45 ***Inés Lloves García and Diego Álvarez Estévez***: Use of Deep Learning Techniques for Motor Events Detection in Polysomnographic Records **297**

46 ***Antonio Díaz-Longueira, Míriam Timiraos, Álvaro Michelena, Óscar Ramos-Fontenla, and Jose Luis Calvo-Rolle***: Development of a Virtual Sensor for COD Measurement in a Wastewater Treatment Plant **305**

47 ***Manuel Silva, Omar A. Mures, Antonio Seoane, and Jose A. Iglesias-Guitian***: Understanding the Influence of Rendering Parameters in Synthetic Datasets for Neural Semantic Segmentation Tasks **313**

48 ***Adrián Nieto Antelo, Diego Fernández Iglesias, and Francisco J. Nóvoa***: Use of Machine Learning Algorithms for Network Traffic Classification **321**

49 ***Eva Pardo-Otero, Iago Fernández-Garrido, Santiago Fernández-Noriega-Balseiro, María Martínez-Pérez, and Laura Nieto-Riveiro***: Solution for Capturing Data from Wearable Devices **327**

50 ***Jorge Fernández-García, Martiño Rivera-Dourado, Rubén Pérez-Jove, Cristian R. Munteanu, and Jose Vázquez-Naya***: Improving Authentication in the Amazon Alexa Virtual Assistant by Using a Geofence **335**

51 ***Amara García Adán and María Ángeles Fuentes Loureiro***: Ethical and Legal Risks of Algorithmic and AI Tools Developed to Fight Against Trafficking in Cultural Property in the RITHMS Project **343**

52 ***Manuel Framil, Mario Carpente, David Fraga, Jonatan Enes, Roberto R. Expósito, Guillermo L. Taboada, and Juan Touriño***: BDWatchdogFaaS: A Tool for Monitoring and Analysis of Functions-as-a-Service in Cloud Environment **351**

Simulation of a Unmanned Aerial Vehicle (UAV) Swarm Control System

Jacques.B. Ngoua-Ndong Avele, and Marc Bernice Angoue Avele

Faculty of Quantum and Optical Electronic, Department of Photonics, Saint Petersburg State Electrotechnical University, 197022,Saint Petersburg, Russia.
Aerospace Technology research group, University of Vigo, 36310 Vigo, Spain.
Correspondence: avelejacques@yahoo.fr

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: A promising direction for the use of UAVs is their association into groups or herds, like the insects observe the principle of swarm organization in the nature. In the case of UAVs, after combining into a swarm, each drone is controlled by its own automation, and the behavior of the swarm can be controlled by a program with elements of artificial intelligence or (several) an operator. In this work, we considered the task of improving the safety of the flight of a group of UAVs in the agro-industrial complex. The simulation of an unmanned aerial vehicle (UAV) swarm control system was created by using the 3D simulator Coppeliasim and the software PyCharm, and it was implemented when monitoring plants in the agro-industrial complex.

1 Introduction

Swarm robotics have many potential applications. These include tasks that require miniaturization, such as distributed sensing tasks in micro machines or the human body. Some of the most promising applications of swarm robotics include search and rescue operations Foroutannia et al. (2021). Swarms of robots of different sizes can be sent to places where rescuers cannot reach safely, to explore the unknown environment using on-board sensors. On the other hand, group robotics could be suitable for applications requiring low-cost construction - such as mining or agricultural operations Kumar et al. (2018). The relevance of this paper is to develop a flight control simulation of a UAV swarm for plant monitoring in agro-industrial complex. In agriculture, the practice of using drones to survey the condition of agricultural plants, monitor the performance of work in the fields, assess the quality of crops,is being actively introduced Pham et al. (2020). UAV(s) are now designed to be used in groups, they can be launched at the target en masse, from several carriers at a time; in flight, they will be able to interact with each other and exchange information - this is known as the UAV swarm.In our case, a swarm of unmanned aerial vehicles will be used for agro-industrial surveillance, with the lead UAV controlled by a pilot (or autopilot) and the slave UAVs working in a co-ordinated manner within a group.

2 Analysis of UAV swarming technology

A UAV swarm, also called a drone fleet, is a coordinated set of drones - air, ground, underground or marine - to perform a common task in different types of applications, civilian or military. work together to achieve the tasks that either cannot be performed by a single drone. Each drone may accomplish a similar role or different roles, such as weapons, or communications relaying Jin et al. (2022). Drone swarming requires a variety of advanced capabilities,

such as the ability to maintain a separation to avoid collision and the ability for each tactical drone to predict where its neighbours will be at any one time. These capabilities may be facilitated by real-time sensing as well as artificial intelligence and computer vision algorithms Xiaoning (2020). Swarm communications for UAV may rely on RF (radio frequency), cellular or satellite communications (SATCOM) Tegicho and Bogale (2022).The swarm may use ad-hoc networking technologies, particularly when operating BVLOS (beyond visual line of sight) and over large areas where existing connectivity is not guaranteed. Individual drones may connect to and disconnect from the network all the time, making a decentralized ad-hoc network structure highly Hartley et al. (2022).

2.1 System description

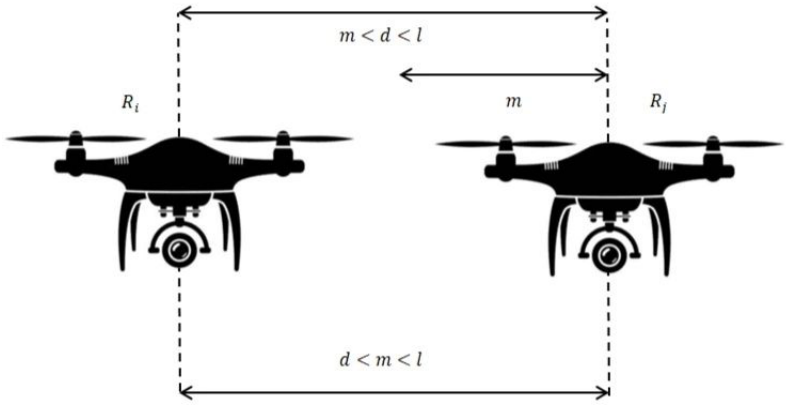


Figure 1: Mathematical model for UAV-to-UAV distance.

Let **m** be the minimum permitted distance between UAVs; **d** - current distance between UAVs R_i and R_j and **L** is the radius of visibility of each UAV R_i and R_j. $\left|\overrightarrow{F_{ij}}(k)\right| \approx \mathrm{d}, \mathbf{m} < \mathbf{d} < \mathbf{L}$, in a straight direction and $\left|\overrightarrow{F_{ij}}(k)\right| \approx \mathrm{d}^{-1}, \mathbf{m} < \mathbf{d} < \mathbf{L}$ in the reverse direction, Where $\left|\overrightarrow{F_{ij}}(k)\right|$ represents the force vector of the distance between UAVs. The UAV(s) in the swarm that has information about the desired direction of motion will be called "leading" UAV(s). And the rest is called "slave" UAV(s). Depending on the nature of information exchange in a UAV swarm, there are two possible scenarios:

- The lead UAV that detects a target tells it neighbour UAV, the coordinates of the target, and so on, and soon all UAVs in the swarm have information about the location of the target and can move in this direction.
- The lead UAV that has detected a target cannot communicate its coordinates to other UAV in the swarm.

2.2 Algorithmic flight program for the UAV swarm

The algorithm shown in Fig. 2 summarises the take-off of the UAV swarm. The first UAV takes off first, followed by the second after 60 seconds if no technical problems are detected. 60 seconds represents the waiting time introduced into the 3D-CoppeliasSim simulation program with python language. **Z** is the height of UAV in reality, and **h** is the height set by the UAV operator.

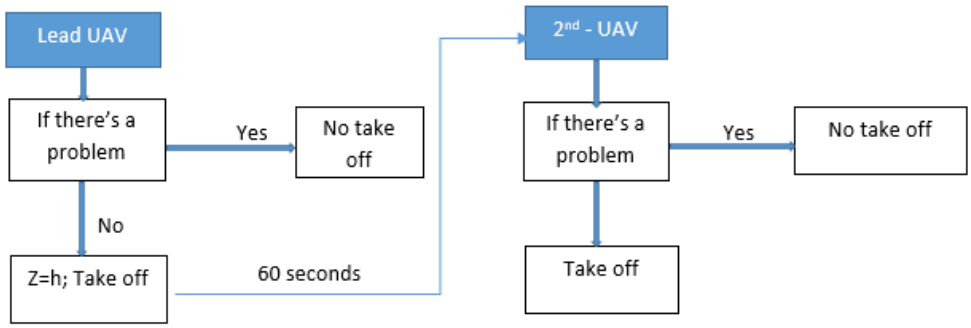


Figure 2: UAV swarm take-off algorithm.

Once the UAV swarm has taken off, its flight will proceed as follows:

- The lead UAV always flies with a time T in front of the second UAV which follows its movement.
- If there is an obstruction, the two UAVs slow down, whether or not they are in the surveillance zone $\{S > L_{min}\}$ (L_{min} is the minimum length covered by the UAVs) and return to the starting point.
- Without obstacles, the two UAVs fly at maximum speed, carrying out their mission.
- During their mission, if there are obstacles, they return to point 2 and if there are no obstacles, they face the condition (L=X, which is the area to be covered).
- The UAV detects the bad plants and complete their mission.

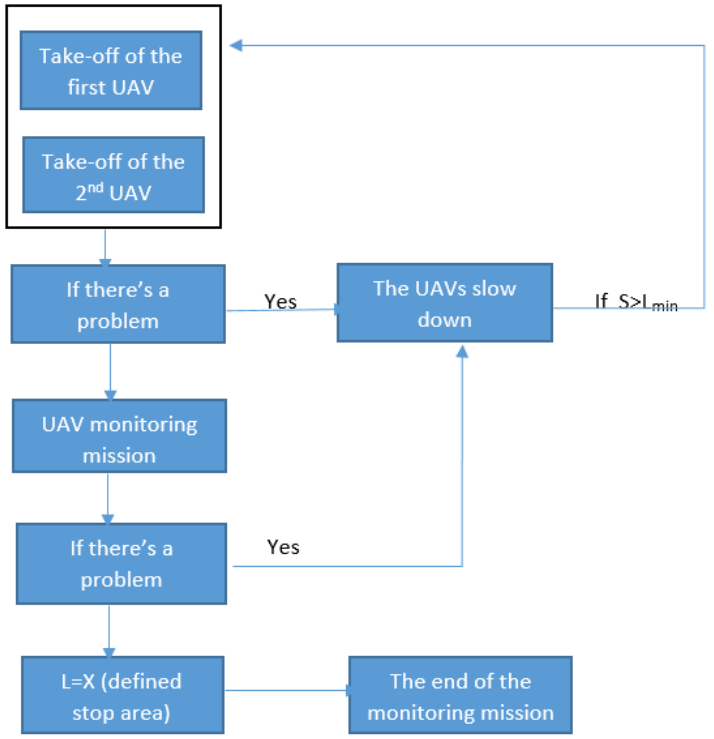


Figure 3: UAV swarm flight algorithm.

After the mission, the UAV swarm turns around to land. The first UAV will then land for 60 seconds, as for take-off, followed by the UAV.

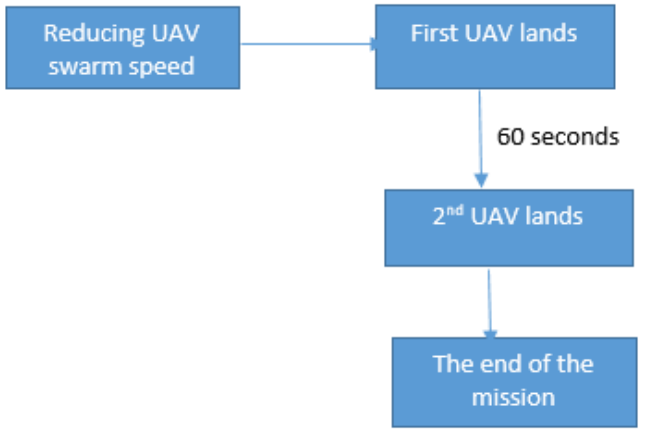


Figure 4: UAV swarm landing algorithm.

3 Results

With the help of the Python programming language application tools and using the CoppeliaSim simulation software environment, the plots of the UAV distribution over the territory were implemented and the distances between the theoretical UAV centers were calculated.

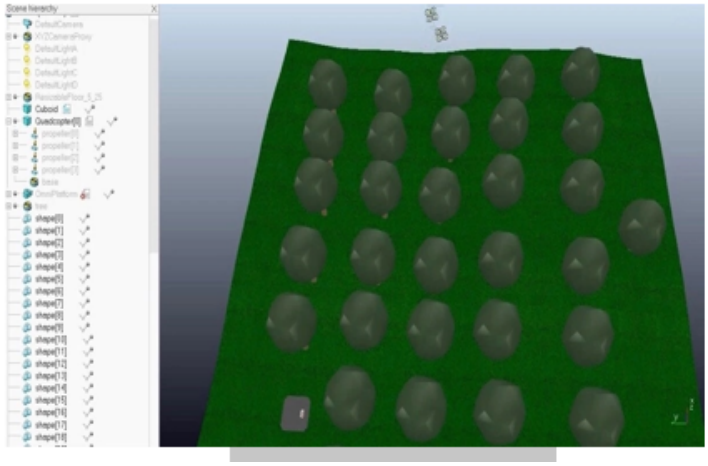


Figure 5: Flight simulation of a swarm of two UAVs in Matlab in 3D Coppeliasim.

the video of this simulation can be seen on this link: *https://drive.google.com/file/d/1Ggsv_jpjV_hM9ej_6L3NXj3b4fNAPMNC/view?usp=drive_link*

4 Conclusions

In this article, the simulation of the drone swarm control system for plant monitoring in the agro-industrial complex has been carried out. For this work, the 3D simulator Coppeliasim and the PyCharm software were used. The following steps were carried out in the process:

- Analysis of drone swarm control technology.
- Development of algorithmic models for UAV swarm takeoff, flight and landing.
- Result of this simulation through a video.

Bibliography

A. Foroutannia, M. Shoryabi, A. A. Anaraki, and A. Rowhanimanesh. Sin: A programmable platform for swarm robotics. In *2021 26th International Computer Conference, Computer Society of Iran (CSICC)*, pages 1–5. IEEE, 2021.

R. J. a. L. Hartley, I. L. Henderson, and C. L. Jackson. Bvlos unmanned aircraft operations in forest environments. *Drones*, 6(7):167, 2022.

H. Jin, Q. Liu, C. Li, Y. T. Hou, W. Lou, and S. Kompella. Hector: A reinforcement learning-based scheduler for minimizing casualties of a military drone swarm. In *MILCOM 2022-2022 IEEE Military Communications Conference (MILCOM)*, pages 887–894. IEEE, 2022.

A. Kumar, S. Santhosh, R. D'souza, et al. Feeder weeder for autonomous farming. In *2018 3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, pages 2574–2578. IEEE, 2018.

T. H. Pham, D. Ichalal, and S. Mammar. Complete coverage path planning for pests-ridden in precision agriculture using uav. In *2020 IEEE International Conference on Networking, Sensing and Control (ICNSC)*, pages 1–6. IEEE, 2020.

B. E. Tegicho and T. E. Bogale. Ris assisted mimo uav swarm communication: Channel capacity margin analysis. In *2022 IEEE 13th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, pages 0347–0351. IEEE, 2022.

Z. Xiaoning. Analysis of military application of uav swarm technology. In *2020 3rd International Conference on Unmanned Systems (ICUS)*, pages 1200–1204. IEEE, 2020.

Eavesdropping and Jamming via Pilot Attacks in 5G Massive MIMO

Marc Bernice Angoue Avele, Darian Pérez-Adán, and Dariel Pereira-Ruisánchez

Aerospace Technology research group, University of Vigo, 36310 Vigo, Spain.
Department of Computer Engineering & CITIC Research Center, University of A Coruña, Spain.
Correspondence: marcbernice.angoue@uvigo.gal

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: In this work, we investigate pilot attacks for 5G single-cell multi-user massive multiple-input multiple-output (MaMIMO) systems with a single-antenna active eavesdropper and a single-antenna jammer operating in time-division duplex (TDD) schemes. Firstly, we describe the attacks when the base station (BS) estimates the channel state information (CSI) based on the uplink pilot transmissions. Finally, we propose a reinforcement learning (RL)-based framework for maximizing the system sum rate that proved robust to the eavesdropping and jamming attacks.

1 Introduction

5G is the name given to the next generation of wireless connectivity, which cellular phone companies are deploying worldwide due to the large demand for high data rates and low latency in mobile service. It is slated to succeed the existing 4G networks, which currently serve as the backbone for most contemporary mobile devices Zhang et al. (2020). All 5G wireless devices are connected to the Internet network by radio waves via an antenna in the cell. Some of the key technologies to be deployed in 5G technology are MaMIMO, device-to-device (D2D) communications, intelligent reflecting surfaces (IRS), and millimeter-wave (mmWave) Perez-Adan et al. (2021). massive multiple-input multiple-output (MaMIMO) is a key 5G technology, which refers to deploying a vast number of antennas at the base stations (BSs) to support multiple users at the same time-frequency resources. MaMIMO has the potential to concentrate the radiation energy in the expected direction by using precoding algorithms, and thus the inter-cell interference can be reduced Wang et al. (2021). Due to the large number of antennas at the BSs and the relatively short channel coherence time, the channel state information (CSI) between the BS and individual users must be frequently estimated by using uplink pilot transmissions Wang et al. (2021). However, the security aspects of MaMIMO systems remain relatively unexplored. Among the critical concerns within MaMIMO systems is the pilot contamination (PC) attack, often referred to simply as a pilot attack Akgun et al. (2018). As a preliminary attempt, the work in Akgun et al. (2018) approaches an important attack scenario, in which a malicious user may send false CSI feedback to a target BS to jam or eavesdrop on messages received by other benign users. Thus leading to corruption in the signal transmitted between the communication ends. In this work, we briefly overview the eavesdropping and jamming attacks in 5G MaMIMO systems and approach reinforcement learning (RL)-based solutions to detect/mitigate communication intrusions.

1.1 Organization

The remainder of this work is structured as follows. We describe the system model and the attacks in Section 2 and Section 3, respectively. In Section 4, we briefly overview RL solutions to detect intrusions and propose an RL-based solution for the precoding design. Simulation results and comparisons are presented in Section 5. Finally, Section 6 is devoted to the conclusions of the paper.

2 System model

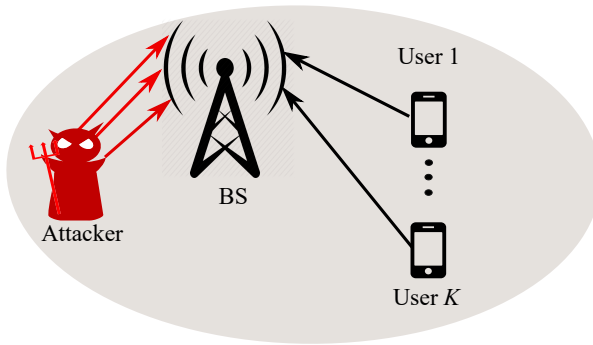


Figure 1: Pilot attack model in a multi-user (K users) uplink MaMIMO system.

Let us consider a MaMIMO system as shown in Figure 1, where a BS is equipped with M antennas to communicate with K single-antenna users, such that $M \gg K$. On the other hand, an attacker (the red devil) is trying to eavesdrop or jam the transmissions between the users and BS. We assume a time division duplex time-division duplex (TDD) system and channel reciprocity holds for coherence time. In MaMIMO systems, the BS needs the channel response of the user terminal to get the estimate of the channel. The TDD protocol establishes that the user sends an uplink pilot sequence which is used by the BS to estimate the CSI for that user in that cell. The BS employs this CSI to estimate the uplink data and for the beamforming design in downlink transmissions.

3 Attack description

The MaMIMO CSI estimation phase is vulnerable to malicious attacks, which can be classified into two forms: Pilot active eavesdropping attack and pilot jamming attack. In the following, we describe both forms of attack and state the corresponding signal model.

3.1 Pilot active eavesdropping attack

Pilot active eavesdropping attacks occur during the uplink. A user transmits a pilot symbol to the BS for channel estimation, and the BS transmits a precoded signal toward the users. At the same time as the uplink pilot transmission, the eavesdropper starts an attack by sending another pilot symbol with the basic assumption that it has the perfect knowledge of the user's pilot symbol and the exact time to transmit. It means that the attacker is synchronized with the legitimate transmission, and this is possible by overhearing the signaling exchange between the BS and the users.

Let $\mathbf{h}_k \in \mathbb{C}^{M\times 1}$ and $\mathbf{h}_{\mathrm{E}} \in \mathbb{C}^{M\times 1}$ be the column vectors representing the uplink channels from the legitimate user (k-th) and eavesdropper to BS, respectively. The elements in $\mathbf{h}_k$ and

$\mathbf{h}_{\mathrm{E}}$ follow a Rayleigh fading model. To estimate the channel response, each user sends a pilot symbol $x_{\mathrm{k},i}$ which is exactly known by the BS. The received pilot at the i-th time instant with $i = 1, \ldots, L$ and without eavesdropping can be expressed as

$$\mathbf{y}_i = \sum_{k=1}^{K} \sqrt{P_k}\mathbf{h}_k x_{k,i} + \mathbf{n}_i, \tag{2.1}$$

where $\mathbf{y}_i$ and $\mathbf{n}_i$ are the $M \times 1$ vectors of received signal and noise, respectively, P_k is the power transmission available in the K users. The vector $\mathbf{n}_i$ contains the additive white Gaussian noise (AWGN) modeled as $\mathbf{n} \sim \mathcal{N}_{\mathbb{C}}(0, \sigma_{\mathrm{n}}^2 \mathbf{I}_M)$. The received signal after L pilot transmissions can be stated as

$$\mathbf{Y} = \sum_{k=1}^{K} \sqrt{P_k}\mathbf{h}_k \mathbf{x}_k + \mathbf{N}, \tag{2.2}$$

where $\mathbf{x}_k \in \mathbb{C}^{1\times L}$ is the block of L symbols corresponding to pilot transmission from the k-th user and $\mathbf{N}$ is the $M \times L$ AWGN matrix. Under a prior knowledge of $\mathbf{x}_k$, the estimated channel (without eavesdropping) would be

$$\hat{\mathbf{h}}_k = \frac{\mathbf{Y}\mathbf{x}_k^*}{\sqrt{P_k}L} = \sum_{i=1}^{K} \frac{\sqrt{P_i}\mathbf{h}_i \mathbf{x}_i \mathbf{x}_k^*}{\sqrt{P_k}L} + \frac{\mathbf{N}\mathbf{x}_k^*}{\sqrt{P_k}L} = \mathbf{h}_k + \tilde{\mathbf{n}}_k, \tag{2.3}$$

with $\tilde{\mathbf{n}}_k \triangleq \frac{\mathbf{N}\mathbf{x}_k^*}{\sqrt{P_k}L} \sim \mathcal{CN}\left(0, \frac{1}{P_k L}\mathbf{I}_M\right)$

If the eavesdropper is active (i.e., it synchronously sends the same pilot sequence as the target users), the pilot-based channel estimation will also contain the component of the eavesdropper to the BS, hence the received signal at the BS is written as

$$\mathbf{Y} = \sum_{k=1}^{K} \sqrt{P_k}\mathbf{h}_k \mathbf{x}_k + \sqrt{P_{\mathrm{E}}}\mathbf{h}_{\mathrm{E}} \mathbf{x}_j + \mathbf{N}, \tag{2.4}$$

where $\mathbf{x}_j$ is the signal sent from the attacker and P_{E} is the attacker power. $\mathbf{Y}$ is the $M \times L$ matrix containing the pilots received by M antennas at BS. We use the model of $\mathbf{x}_j$ by considering that $\mathbf{x}_j = \sqrt{P_{\mathrm{E}}} \sum_{k=1}^{K} \mathbf{x}_k$, as in the reference Akgun et al. (2017). In this case, the estimated channel for the k-th user will be given by Akgun et al. (2017)

$$\hat{\mathbf{h}}_k = \mathbf{h}_k + \tilde{\mathbf{n}}_k + \sqrt{\alpha_k}\mathbf{h}_{\mathrm{E}}, \tag{2.5}$$

where $\alpha_k = P_{\mathrm{E}}/P_{\mathrm{k}}$ is the ratio between the average power at the attacker and the power allocated by the k-th user to the pilot.

The active eavesdroppers aim to disturb the functioning of the network. Specifically, the eavesdropper's aims are the following

- Exploit the weaknesses in the user capacity-optimized pilot sequence design to increase the pilot attacks within the uplink channel estimation.
- Degrade the user signal-to-interference-plus-noise ratio (SINR) to a point where it cannot meet its requirements, even with a large number of antennas at the BS.

3.2 Pilot jamming attack

By considering the uplink of a single-cell MaMIMO system, depicted in Figure 1, let us consider now a single-antenna jammer (the red devil). During the uplink, the users send to the BS a pilot sequence, and an attacker, in this case, sends a jamming signal to interfere with the channel estimation, and it can adopt a strategy according to its knowledge of the system:

- If the jammer does not have prior knowledge of the pilot sequences used by the user, then it will send a random sequence to attack the system Do et al. (2016). It means that during the uplink, the user sends a pilot sequence, while the jammer sends a random jamming sequence. where $\mathbf{n}_j$ is a pseudorandom noise pilot transmitted by the jamming.
- The jammer has prior knowledge of the pilot sequences used by the users, and it can know the transmission protocol and the pilot set. The jammer can obtain this information by listening to the channel for some consecutive blocks.

In any case, this attacker transmits a random jamming sequence. Therefore, the signal received by the BS in (2.4) would be written as

$$\mathbf{Y} = \sum_{k=1}^{K} \sqrt{P_k}\mathbf{h}_k\mathbf{x}_k + \sqrt{P_{\mathrm{E}}}\mathbf{h}_{\mathrm{E}}\mathbf{n}_j + \mathbf{N}, \tag{2.6}$$

and the estimated channel in (2.5) for the k-th user will be a highly distorted version of $\mathbf{h}_k$ due to the noise transmitted by the attacker.

Since one of the main advantages of MaMIMO is achieving high spectral efficiency, the jammer, by launching a pilot attack, will degrade the asymptotic spectral efficiency of the legitimate system Do et al. (2016). In Vinogradova et al. (2016) is shown that if the jammer smartly adjusts its transmission power to match the desired signals, the spectral efficiency is significantly affected.

4 Technical solution

In this section, we discuss a possible technical solution for mitigating pilot attacks in 5G MaMIMO systems. We proposed a method to deal with pilot attacks using RL-based solutions. In particular, we aim to maximize the system sum rate in the downlink system by considering an active eavesdropper or a jamming signal. RL is one of the three primary machine learning (ML) paradigms, alongside supervised and unsupervised learning. In this area of ML, the agent is given a reward for taking actions that lead to desired outcomes. Over time, the agent learns to take actions that maximize the notion of cumulative reward. Some works are approached in the literature to detect communication intrusions. The work in Tu et al. (2021) proposes an RL-based technique to detect impersonation attacks in device-to-device (D2D) communications. They formulate this task as a Markov decision process and learn the optimal policy for detecting impersonation attacks. The authors in Sedjelmaci (2020) propose an RL-based approach to detect attacks in 5G wireless networks. The approach uses a hierarchical RL algorithm to learn the optimal policy to detect attacks.

4.1 RL-based solution

In this subsection, we develop a solution for active eavesdropping and jamming mitigation based on RL. We model a sum-rate maximization problem and design the downlink precoders by considering RL-based optimization.

4.2 Downlink transmission under active eavesdropping

Let s_k be the information signal sent to the k-th user from the BS, which is previously precoded with $\mathbf{v}_k \in \mathbb{C}^{M\times 1}$.

$$y_k = \sum_{i=1}^{K} \sqrt{P_i{}^{(d)}}\mathbf{h}_k^{*}\mathbf{v}_i s_i + n_k^{(d)}, \tag{2.7}$$

where $P_k^{(d)} = \mathbf{v}_k^* \mathbf{v}_k$ and $n_k^{(d)}$ are the allocated power to s_k at BS and the downlink AWGN signal. The achievable rate for the k-th user can be defined as follows by assuming $\sigma_{\mathrm{n}}^2 = 1$

$$R_k = \log\left(1 + \frac{P_k{}^{(d)} \mid \hat{\mathbf{h}}_k^* \mathbf{v}_k \mid^2}{\sum_{l \in \mathcal{K}/k} P_l{}^{(d)} \mid \hat{\mathbf{h}}_l^* \mathbf{v}_l \mid^2 + 1}\right). \tag{2.8}$$

Note that the following sum-rate maximization problem can be addressed to determine the downlink precoders and the power allocation

$$\begin{aligned} \left[\mathbf{v}_1^* \cdots \mathbf{v}_K^*\right] &= \underset{\mathbf{v}_k, \forall k \in \mathcal{K}}{\arg\max} \sum_{k=1}^{K} R_k \\ \text{s.t.} \sum_{k=1}^{K} \mathbf{v}_k^* \mathbf{v}_k &\leqslant P_{\mathrm{A}}, \end{aligned} \tag{2.9}$$

being $P_{\mathrm{A}} \geqslant \sum_{k=1}^{k} P_k^{(d)}$ the total power available at the BS with SNR (dB) $= 10 \log_{10}(P_{\mathrm{A}})$. However, note that the channel estimation is corrupted by the active eavesdropper or the jamming attacker (cf. (2.5)).

Due to the cost function and the design constraints for the precoders, the formulation in (2.9) becomes a non-trivial optimization problem. Besides, handling the eavesdropping/jamming scenario without knowing the statistics of the CSI errors is a cumbersome task. We propose an RL-based solution to solve (2.9) by leveraging the learning capabilities of neural networks.

State, action, and reward function

The state, action, and reward elements that we consider to solve 2.9 are the following. The **state** vector $\mathbf{t}_\ell$ is constructed as $\mathbf{t}_\ell = [\hat{\mathbf{h}}_1^*, ..., \hat{\mathbf{h}}_K^*]$. Notice that the states stack the information related to the channel realizations (BS-users). The action vector is composed of the vectors that we are optimizing (i.e., the precoders), such that $\mathbf{a}_\ell = [\mathbf{v}_1^*, ..., \mathbf{v}_K^*]$. Finally, the reward function will be defined by

$$r_\ell = \sum_{i=1}^{K} R_k \Big| (\mathbf{t}_\ell, \mathbf{a}_\ell). \tag{2.10}$$

Algorithm 1 DCB-DDPG algorithm

Initialize:
set $\pi(\mathbf{s}, \boldsymbol{\vartheta}_\pi)$ with random $\boldsymbol{\vartheta}_\pi$
set $r(\mathbf{s}, \mathbf{a}, \boldsymbol{\vartheta}_r)$, $r(\mathbf{s}, \mathbf{a}, \tilde{\boldsymbol{\vartheta}}_r)$ with $\boldsymbol{\vartheta}_r = \tilde{\boldsymbol{\vartheta}}_r$ and set the buffer $\mathcal{R}$
for $\ell = 0, \ldots, T-1$ **do:**
  set $\mathbf{t}_\ell$ given $\hat{\mathbf{h}}_{\mathrm{k}}, \forall k$
  agent takes $\mathbf{a}_\ell = \pi(\mathbf{s}_\ell, \boldsymbol{\vartheta}_\pi) + \mathbf{n}_{\mathrm{e}}$
  environment returns r_ℓ
  $\mathcal{R}$ stores $\mathcal{E}_\ell = (\mathbf{s}_\ell, \mathbf{a}_\ell, r_\ell)$
  if $|\mathcal{R}| > |\mathcal{B}|$:
    Sample $|\mathcal{B}|$ random experiences $\mathcal{E}_i = (\mathbf{s}_i, \mathbf{a}_i, r_i)$, $i = 0, \ldots, |\mathcal{B}| - 1$
    Compute L_{c} and backpropagate to update $\boldsymbol{\vartheta}_r$
    Compute L_{a} and backpropagate to update $\boldsymbol{\vartheta}_\pi$
    $\tilde{\boldsymbol{\vartheta}}_r \leftarrow \tau \boldsymbol{\vartheta}_r + (1 - \tau) \tilde{\boldsymbol{\vartheta}}_r$
  end if
Obtain $\mathbf{v}_k, \forall k$ by evaluating $\pi(\mathbf{s}, \boldsymbol{\vartheta}_\pi)$

Output: $[\mathbf{v}_1^* \cdots \mathbf{v}_K^*]$

In the algorithmic solution, we use a deep contextual bandit-oriented deep deterministic policy gradient (DCB-DDPG) agent similar to the one approached in Pereira-Ruisánchez et al. (2023). Similar structures are employed for the neural networks in the actor and critic networks but proper adaptations have been performed for the dimensions of the state vectors.

The proposed agent is composed of the actor-network $\pi(\mathbf{t}, \boldsymbol{\vartheta}_\pi)$, the critic network $r(\mathbf{t}, \mathbf{a}, \boldsymbol{\vartheta}_r)$, the target critic network $r(\mathbf{t}, \mathbf{a}, \tilde{\boldsymbol{\vartheta}}_r)$, and the replay buffer $\mathcal{R}$. Note that $\boldsymbol{\vartheta}_\pi$ and $\boldsymbol{\vartheta}_r$ are vectors of weights of the actor and critic network, respectively. The actor and critic networks are trained from stored experiences. We compute the critic and the actor losses as

$$L_{\mathrm{c}} = \frac{1}{|\mathcal{B}|} \sum_i (r_i - r(\mathbf{s}_i, \mathbf{a}_i, \boldsymbol{\vartheta}_r))^2 \tag{2.11}$$

and

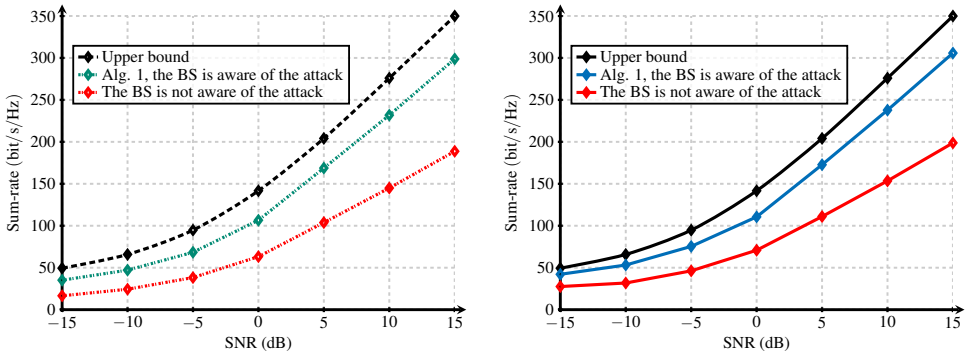
$$L_{\mathrm{a}} = -\frac{1}{|\mathcal{B}|} \sum_i r(\mathbf{s}_i, \pi(\mathbf{s}_i, \boldsymbol{\vartheta}_\pi), \tilde{\boldsymbol{\vartheta}}_r), \tag{2.12}$$

respectively. The exploration noise is defined as $\mathbf{n}_{\mathrm{e}} \sim \mathcal{N}_{\mathbb{C}}(0, \sigma_{\mathrm{n_e}}^2 \mathbf{I}_{D_{\mathrm{action}}})$. Algorithm 1 summarizes the interactions between these elements during the training stage.

5 Simulation results

In this section, we present computer experiments to analyze the capability of the proposed DCB-DDPG to solve the optimization problem in (2.9) while considering active eavesdropping and jamming. We have previously approached the active eavesdropping scheme, however, note that the jamming scenario can be easily modeled by considering (2.6) instead of (2.4). We have considered a multi-user MaMIMO setup with $K = 10$ single-antenna users and $M = 400$ antennas at the BS, where a single-antenna eavesdropper or jammer has attacked during the pilot transmission. The number of pilot symbols in the CSI estimation is set to $L = 20$.

The values of the training steps ($T = 100000$) and the size of the replay buffer, $|\mathcal{R}|_{\max}$, were selected to fit the intended training time. The mini-batch is defined as $|\mathcal{B}| = 16$ whereas the exploration noise variance is set to $\sigma_{\mathrm{n_e}}^2 = 0.05$. The remaining parameters were obtained experimentally to provide the highest system performance.



(a) Active eavesdropping attack.

(b) Jamming attack.

Figure 2: Sum rate by considering $K = 10$ users, $M = 400$ antennas at the BS, $L = 20$ pilot symbols, and different strategies to configure the precoders and power allocation.

Figure 2 shows the achievable sum rates obtained with the proposed algorithm and two baseline strategies: 1-) an upper bound strategy without eavesdropping or jamming; 2-) a lower bound (non-robust) strategy where the BS disregards the effects of active eavesdropping or jamming attack. As shown, the performance of the proposed algorithm comes close to the upper bound scheme while offering large gains over the non-robust strategy for both scenarios active eavesdropping (Figure 2(a)) and jamming (Figure 2(b)) attacks. It is also observed that slightly higher system performance is achieved with our proposed RL-based solution under the jamming attacker scenario over that assessed under the active eavesdropping attack.

6 Conclusions

In this paper, we have briefly discussed the eavesdropping and jamming pilot attacks in 5G massive MIMO. We have also proposed an RL-based system design to deal with active eavesdropping and jamming in the downlink of a multi-user massive MIMO system. The results show large gains (in terms of sum rate) provided by the proposed solution over a baseline strategy that neglects the effects of the attacker in the channel estimation process in the communication system.

Acknowledgments

CITIC is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS).

Bibliography

B. Akgun, M. Krunz, and O. O. Koyluoglu. Pilot contamination attacks in massive MIMO systems. In *2017 IEEE Conference on Communications and Network Security (CNS)*, pages 1–9. IEEE, 2017.

B. Akgun, M. Krunz, and O. O. Koyluoglu. Vulnerabilities of massive MIMO systems to pilot contamination attacks. *IEEE Trans. Inf. Forensics Secur.*, 14(5):1251–1263, 2018.

T. T. Do, H. Q. Ngo, T. Q. Duong, T. J. Oechtering, and M. Skoglund. Massive MIMO pilot retransmission strategies for robustification against jamming. *IEEE Wireless Commun. Lett.*, 6 (1):58–61, 2016.

D. Pereira-Ruisánchez, O. Fresnedo, D. Pérez-Adán, and L. Castedo. Deep contextual bandit and reinforcement learning for IRS-assisted MU-MIMO systems. *IEEE Transactions on Vehicular Technology*, 72(7):9099–9114, 2023. doi: 10.1109/TVT.2023.3249353.

D. Perez-Adan, O. Fresnedo, J. P. Gonzalez-Coma, and L. Castedo. Intelligent reflective surfaces for wireless networks: An overview of applications, approached issues, and open problems. *Electronics*, 10(19):2345, 2021.

H. Sedjelmaci. Attacks detection approach based on a reinforcement learning process to secure 5g wireless network. In *IEEE International Conference on Communications Workshops (ICC Workshops)*, pages 1–6. IEEE, 2020.

S. Tu, M. Waqas, S. U. Rehman, T. Mir, G. Abbas, Z. H. Abbas, Z. Halim, and I. Ahmad. Reinforcement learning assisted impersonation attack detection in device-to-device communications. *IEEE Transactions on Vehicular Technology*, 70(2):1474–1479, 2021.

J. Vinogradova, E. Björnson, and E. G. Larsson. Detection and mitigation of jamming attacks in massive mimo systems using random matrix theory. In *2016 IEEE 17th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 1–5. IEEE, 2016.

Z. Wang, D. Shi, and H. Wu. The role of massive MIMO and intelligent reflecting surface in 5g/6g networks. In *2021 International Conference on Wireless Communications and Smart Grid (ICWCSG)*, pages 309–312. IEEE, 2021.

X. Zhang, X. Liu, G. Wei, Y. Zhao, and Y. Guo. The study of 5G massive MIMO end-to-end MPAC test solution. In *14th European Conference on Antennas and Propagation (EuCAP)*, pages 1–5. IEEE, 2020.

Music Recommendation System Based on Ratings Obtained from Amazon

Sergio Marcos-Vazquez, Carlos Fernandez-Lozano, Adrian Carballal, and Francisco Cedron

RNASA-IMEDIR, Faculty of Computer Science, Universidade da Coruña, 15071 A Coruña, Spain
Software Engineering Laboratory, Faculty of Computer Science, Universidade da Coruña, 15071 A Coruña, Spain
Centro de Investigación CITIC, Universidade da Coruña, 15071 A Coruña, Spain
Correspondence: francisco.cedron@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: In the current context of an era in which a significant portion of people are constantly living online, with various multimedia streaming platforms serving as major sources of entertainment, and with e-commerce playing also a key role, recommender systems are carving out their place as one of the most important and widely used tools for enhancing user experiences on these platforms. This work undertakes a comparative study on some of the techniques used within these systems, mainly focused on those based in collaborative filtering. Multiple recommender systems will be implemented according to each of these methods, taking for this purpose the vinyl records and CDs Amazon's user ratings.

1 Introduction

According to the Recording Industry Association of America, the music industry business has been on an upward trend since 2014 (Recording Industry Association of America, 2023). This may be due to the data provided by the billboard indicating that more than 100,000 songs are uploaded every day(Billboard, 2023).

With such a large number of songs, it is understood that it is impossible for a person to manually select which songs they like the most manually. It is extremely necessary to have a recommendation system for some time that indicates users new songs based on the tastes that it leaves in their musical fingerprint.

One way to obtain a person's personalized tastes is through the reviews they publish in ecommerce. With all the data found in ecommerce, profiles of aesthetic tastes could be established that help propose new content to users based on the evaluations of users who have a similar musical taste.

In this work, reviews of CDs and vinyl obtained from Amazon will be used (Rappaz et al., 2021). The objective is to use the ratings in several recommendation systems based on collaborative filtering, measuring and comparing the results obtained for each of them.

2 Collaborative filtering systems

Collaborative filtering recommendation systems are based on the evaluation of the interest a user may have in a certain item based on the opinions of other users(Schafer et al., 2007). This

is based on the premise that if two users have similar opinions on certain items, it is likely that an item liked by one of the users will also be liked by the other.

In order to capture the preferences and behaviours of users with respect to items, a matrix is created that has as dimensions users and items, and where the value of the cells would correspond to the ratings. This matrix is fundamental in recommendation systems based on collaborative filtering, as it will be used as the basis for the algorithms used in this type of recommendation systems.

These matrices tend to have a peculiarity: it is normal that a user does not rate all the articles, nor is an article rated by all the users. In fact, the number of ratings that a user usually makes is tiny in comparison to the number of articles that he or she can rate, and the same is true for articles, where the number of ratings received is also very small in comparison to the number of existing users. This causes most of the cells of the rating matrix to be empty, generating sparse matrices(Branham, 1990).

Depending on how you work with this matrix, you can find two categories into which collaborative filtering recommendation systems can be divided: collaborative filtering based on models and collaborative filtering based on neighbors(Singh et al., 2020).

2.1 Collaborative filtering based on models

This category attempts to fill in the item-user rating matrix, predicting the ratings that a user would give to an item that they have not rated at the moment. These models are built from a training dataset that contains patterns with item-user relationships.

Alternating Least Squares

The Alternating Least Squares (ALS) method is an algorithm used for matrix factorization. It factors a given matrix R into two smaller matrices U and V such that $R \approx U^T V$(Hastie et al., 2015). The goal is to predict user preferences for items based on past interactions. The algorithm attempts to estimate the ratings matrix R as the product of two lower-rank matrices, X and Y, i.e., $X * Yt = R$(Clarkson and Woodruff, 2017). During each iteration, one of the factor matrices is held constant, while the other is solved for using least squares. The newly-solved factor matrix is then held constant while solving for the other factor matrix.

Singular Value Decomposition

The singular value decomposition (SVD) of an array is a factorization of that array into three arrays $A = UWV^T$(Banerjee and Roy, 2014). U is mxn array of the orthonormal eigenvectors of AA^T. V^T is the transpose of a nxn array containing the orthonormal eigenvectors of $A^T A$. Finally W is an nxn diagonal array of the singular values that are the square roots of the eigenvalues of $A^T A$.

Non-negative Matrix Factorization

Non-negative matrix factorization (NMF) is a group of algorithms in multivariate analysis and linear algebra where a matrix V is factorized into two matrices W and H, with the property that all three matrices have no negative elements(Sra and Dhillon, 2005). This non-negativity makes the resulting matrices easier to inspect.

In NMF, the goal is to find two non-negative matrices W and H whose product approximates the non-negative matrix X. This factorization can be used for example for dimensionality reduction, source separation or topic extraction.

2.2 Collaborative filtering based on neighbors

This category is based on using various similarity metrics on two vectors of the item-user matrix, thus calculating a similarity with which to find the nearest neighbors.

k-Nearest Neighbor

The k-nearest neighbors algorithm (k-NN) is a non-parametric supervised learning method(Cover and Hart, 1967). In k-NN, the input consists of the k closest training examples in a data set. The output depends on whether k-NN is used for classification or regression. In k-NN classification, the output is a class membership. An object is classified by a plurality vote of its neighbors, with the object being assigned to the class most common among its k nearest neighbors. If $k = 1$, then the object is simply assigned to the class of that single nearest neighbor. In k-NN regression, the output is the property value for the object. This value is the average of the values of k nearest neighbors. If $k = 1$, then the output is simply assigned to the value of that single nearest neighbor.

3 Methodology

The dataset has 4,543,369 ratings, made by a total of 1,944,316 users on 434,060 products. Not all of this data will be used to train the recommendation system, but only the ratings made by users with 10 or more ratings will be used, which leaves the dataset to be used with 1,351,025 ratings and 44,875 users. This filtering is carried out in order to try to avoid the cold start problem(Tey et al., 2020).

With the data prepared, and before starting to train the model, 20% of the ratings are extracted from the dataset, which will be used as a validation set to be able to check the good fit and performance of the model during training. This division occurs in a stratified manner(Ojala and Garriga, 2010). For training, 10-fold cross-validation is used in order to prevent overfitting. The average value of the 10 folds is used as the result.

To adjust the algorithms, we proceed in the same way for all of them, using the hyperparameter adjustment technique known as grid search(Yang and Shami, 2020). The values used are shown in table 1.

Table 1: Values used in the grid search technique by the different collaborative filtering methods. The meaning of each hyperparameter can be consulted in the documentation of the implementation used. ALS: Apache Spark - MLlib (2022), SVD: Nicolas Hug (2020c), NMF: Nicolas Hug (2020b) and k-NN: Nicolas Hug (2020a).

Method	Parameter	Values
ALS	MaxIter	5, 10, 15, 20
	Rank	10, 15, 25, 30, 50, 100
	RegParam	0.1, 0.3, 0.5, 0.7, 1.5
SVD	N_factors	50, 100, 200
	N_epochs	50, 100, 200
	Lr_all	0.005, 0.01
	Reg_all	0.02, 0.05, 0.1
NMF	N_factors	15, 50, 100, 200
	N_epochs	25, 50, 75, 100, 200
	Reg_pu	0.06, 0.1, 0.3
	Reg_qi	0.06, 0.1, 0.3
k-NN	K	10, 20, 25, 30, 40, 70, 100, 1000, 1163, 1500, 10000

Finally, the metrics used to measure the performance of the different methods were the following:

- **MAE**: Mean absolute error calculates the average of the absolute differences between the predicted ratings and the actual ratings.
- **RMSE**: The root mean square error measures, like the MAE, the difference between the predicted and actual valuations, with the difference that, instead of using absolute differences, it calculates the squared differences and takes the square root of the average of This differences. By squaring the differences, this metric penalizes larger errors in predictions more.
- **Precision**: Precision is defined as the ratio of relevant recommendations within the top k of recommendations over the total of recommended articles within the top k, taking in this specific case as relevant recommendations those that have a rating greater than 4.0, and 10 as the value of k.
- **Recall**: Recall is defined as the ratio of relevant recommendations within the top k of recommendations over the total of relevant articles, taking as relevant those ratings greater than 4.0 and 10 as the value of k, as in precision.
- **F1-Score**: Combination of precision and recall measures into a single value, allowing the combined performance of both to be compared.

4 Results

In order to know the performance achieved by each of the different methods, the best model obtained by each method will be used. Table 2 shows which hyperparameters of each algorithm have obtained the best results.

Table 2: Hyperparameters of each method that obtained the best results.

Method	Hyperparameters
ALS	maxIter=20, rank=10, regParam=0.3
SVD	n_factors=200, n_epochs=100, lr_all=0.005, reg_all=0.1
NMF	n_factors=50, n_epochs=50, reg_pu=0.1, reg_qi=0.06
k-NN	k=100

The results obtained by the different collaborative filtering methods are shown in table 3.

Table 3: Hyperparameters and execution time of each method that obtained the best results.

Method	RMSE	MAE	Precision	Recall	F1-Score	Execution time
ALS	0.97761	0.76676	0.23682	0.98334	0.38172	1h 45 min
SVD	0.83394	0.57884	0.85034	0.83183	0.84098	9h 30 min
NMF	0.92317	0.58730	0.85382	0.82332	0.83829	2h 50 min
k-NN	0.96502	0.74615	0.86629	0.94585	0.90432	30 min

5 Conclusions

The ALS algorithm is the one that obtains the worst results, having similar RMSE and MAE to KNN. This is surprising since ALS is an algorithm focused on predictions, while KNN focuses on recommendations, which can be seen at first glance since it is the algorithm that obtains the best F1-Score. Of course, the F1-Score of ALS, which is very low compared to the rest, is entirely due to the precision that this algorithm has, since its recall is the best of all, the reason for this being that the algorithm focuses in recommending the most popular products. On the other hand and as a point in favor, although the RMSE and MAE values obtained are the worst compared to the rest, they are not considered bad results, and combining this with the fact that the combination of execution time of this algorithm competes with Because NMF is the best, ALS can be considered a good option if you want to obtain acceptable recommendations in a very good time.

Regarding KNN, as already mentioned, it gives the best F1-Score, having slightly the highest precision plus a very good recall. This makes a lot of sense since it is an algorithm especially focused on recommendations, but the RMSE and MAE it obtains are also quite acceptable.

For the two remaining algorithms, and starting with SVD, it can be seen that of all the algorithms it is the one that obtains the best results overall.

On the other hand, it is also very clear that this algorithm is the slowest, and its use could be seen in cases where the best possible recommendations are wanted without the speed of obtaining them being a very important factor. Finally, the NMF algorithm is consolidated as a very solid option, since the difference in results with SVD is very small, even achieving better precision. It can be criticized that it can make some larger errors than SVD as can be seen in its RMSE values, but this is still a very good value. Also, its execution time is quite decent, more than good when compared to SVD.

Acknowledgements

CITIC is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS).

We are particularly grateful to CESGA (Galician Supercomputing Centre) for providing access to their infrastructure.

Bibliography

Apache Spark - MLlib. Als - pyspark 3.4.1 documentation. *https://spark.apache.org/docs/latest/api/python/reference/api/pyspark.ml.recommendation.ALS.html*, 2022. [Online; accessed 13-September-2023].

S. Banerjee and A. Roy. *Linear algebra and matrix analysis for statistics*. Chapman & Hall/CRC Texts in Statistical Science. Chapman & Hall/CRC, Philadelphia, PA, June 2014.

Billboard. The ledger: Are there really 100,000 new songs uploaded a day? maybe more. *https://www.riaa.com/u-s-sales-database/*, 2023. [Online; accessed 13-September-2023].

R. L. Branham. Sparse matrices. In *Scientific Data Analysis*, pages 34–66. Springer New York, 1990.

K. L. Clarkson and D. P. Woodruff. Low-rank approximation and regression in input sparsity time. *Journal of the ACM*, 63(6):1–45, Jan. 2017.

T. Cover and P. Hart. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1):21–27, 1967.

T. Hastie, R. Mazumder, J. D. Lee, and R. Zadeh. Matrix completion and low-rank SVD via fast alternating least squares. *Journal of Machine Learning Research*, 16:3367–3402, 2015.

Nicolas Hug. Surprise: A python library for recommender systems - k-nn inspired algorithms. *https://surprise.readthedocs.io/en/stable/knn_inspired.html*, 2020a. [Online; accessed 13-September-2023].

Nicolas Hug. Surprise: A python library for recommender systems - matrix factorization-based algorithms - nmf. *https://surprise.readthedocs.io/en/stable/matrix_factorization.html#surprise.prediction_algorithms.matrix_factorization.NMF*, 2020b. [Online; accessed 13-September-2023].

Nicolas Hug. Surprise: A python library for recommender systems - matrix factorization-based algorithms - svd. *https://surprise.readthedocs.io/en/stable/matrix_factorization.html#surprise.prediction_algorithms.matrix_factorization.SVD*, 2020c. [Online; accessed 13-September-2023].

M. Ojala and G. C. Garriga. Permutation tests for studying classifier performance. *Journal of machine learning research*, 11(6), 2010.

J. Rappaz, J. McAuley, and K. Aberer. Recommendation on live-streaming platforms: Dynamic availability and repeat consumption. In *Fifteenth ACM Conference on Recommender Systems*. ACM, Sept. 2021.

Recording Industry Association of America. U.s music revenue database. *https://www.riaa.com/u-s-sales-database/*, 2023. [Online; accessed 13-September-2023].

J. B. Schafer, D. Frankowski, J. Herlocker, and S. Sen. Collaborative filtering recommender systems. In *The Adaptive Web*, pages 291–324. Springer Berlin Heidelberg, 2007.

P. K. Singh, P. K. D. Pramanik, and P. Choudhury. Collaborative filtering in recommender systems: Technicalities, challenges, applications, and research trends. In *New Age Analytics*, pages 183–215. Apple Academic Press, May 2020.

S. Sra and I. Dhillon. Generalized nonnegative matrix approximations with bregman divergences. In Y. Weiss, B. Schölkopf, and J. Platt, editors, *Advances in Neural Information Processing Systems*, volume 18. MIT Press, 2005.

F. J. Tey, T.-Y. Wu, C.-L. Lin, and J.-L. Chen. Accuracy improvements for cold-start recommendation problem using indirect relations in social networks. *Journal of Big Data*, Dec. 2020.

L. Yang and A. Shami. On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing*, 415:295–316, Nov. 2020.

Forest Bathing Environments in Virtual Reality: Enhancing Well-being through Immersive Nature Experiences

Raquel Simões de Almeida, Paulo Veloso Gomes, António Marques, and Javier Pereira

LabRP, Center for Rehabilitation Research, School of Health, Polytechnic of Porto, Porto, Portugal
CITIC Research Center, University of A Coruña, A Coruña, Spain
Correspondence: velosogomes@ess.ipp.pt

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: The present study proposes an innovative approach to promote wellbeing by developing forest bathing environments in virtual reality (VR). Forest bathing, a practice rooted in traditional Japanese culture, involves immersing oneself in a forest atmosphere to enhance physical and mental health. VR technology enables the opportunity to recreate and simulate these natural environments, allowing individuals to experience the benefits of forest bathing regardless of their physical location and limitations of physical access to natural environments. This study aims to design and evaluate the effectiveness of a VR forest bathing environment in improving well-being, stress reduction, and overall relaxation.

1 Introduction

VR can be used in a wide range of domains, such as medicine, education, arts, entertainment, military, and recently it has been widely used in the health sector (Wang et al., 2019). The advantages of manipulating this technology as a therapeutic tool have been the subject of research. This type of immersive experience is now being used by several studies to complement other treatments, as an adjunctive distraction for both physical symptoms - pain, nausea, blood pressure, heart rate and respiratory rate - and psychological symptoms - anxiety, fear, and stress (Hsieh and Li, 2022). Researchers suggest that relaxation can be achieved, and symptoms of anxiety and depression can be reduced through the use of natural environments in VR (Chan and Qiu, 2021), and as this technology continues to be explored and studied, it may present an alternative to actual natural immersion by creating a simulated environment (Hsieh and Li, 2022). Psychological intervention based on exposure to nature can help alleviate anxiety and other negative emotions in chronically ill patients, supported by fundamentals such as the attention recovery theory that suggests that mental fatigue can be alleviated by exposure to nature (Hsieh and Li, 2022). The control of stimuli produced in the nature environment by VR allows precision in the implementation of therapeutic strategies (Freeman et al., 2017). Current literature shows that VR technology is growing rapidly and proving effective in reproducing natural forests for health and mental health applications (Alyan et al., 2021). The potential benefits of virtual nature are directly related to theories describing the effects of shinrin-yoku - also known as Forest Therapy (FT) (Zhang et al., 2022) - on health, including promoting relaxation, recovery and alertness, improving immune system function, and reducing exposure to air pollution and urbanity (Ross and Jones, 2022). FT is an alternative form of therapy that makes use

of physiological relaxation as a means to boost immunity and improve mental health (Alyan et al., 2021). Numerous studies have proven the effectiveness of FT in improving health, including reducing stress, improving mood, boosting immune function, and lowering blood pressure (Yi et al., 2022). To do this, the person or group travels to forests or nature spaces to participate in treatment activities that connect them to the landscape through sound, smell, touch, sight, and sometimes even taste (Park et al., 2022). Recent studies show that a program lasting one hour and forty-five minutes has a positive impact on the mental health of patients with affective disorders, corroborating the results obtained in previous studies also in other pathologies (Lee et al., 2017), (Poulsen et al., 2016). FT can be an asset in several situations, namely for people with depression, anxiety, hypertension and affective disorders, post-traumatic stress, among others (Lee et al., 2017), (Poulsen et al., 2016), (Song et al., 2016). In addition to the psychological benefits that this population could enjoy if they attended a VR FT program, its relevance is exacerbated due to recent changes in the structure of the hospital. The integration of this VR-based therapy into the healthcare system would compensate for the obstacles caused by the difficulty of access to forests (Jo et al., 2021). Thus, the potential for the transition and applicability of VR FT is promising. Furthermore, exposure to green spaces, which include vegetation in the form of foliage, trees and views, as in the case of FT, improves health outcomes, whether this exposure involves "live" and real nature or virtual nature (Ross and Jones, 2022). Realistic environments have been shown to be more effective in reducing stress levels (Alyan et al., 2021), and VR technology is the most capable of providing this, offering a sense of realism and presence that other devices simply cannot (Maples-Keller et al., 2017). Current findings suggest that virtual nature produces effects similar to those of real nature, making VR a safe alternative to experiencing nature (Chan and Qiu, 2021) and a suitable substitute for people who cannot visit natural areas for whatever reason (Alyan et al., 2021), (Chen et al., 2023). The immediate availability and increased frequency of VR treatments is also noteworthy, reflecting a great advantage and flexibility in service delivery. The increasing accessibility of high-quality VR devices on the market makes it feasible to transfer this technology from the laboratory to other diverse settings (Freeman et al., 2017). Thus, the potential to improve psychological and physiological health through virtual exposure to nature may become a flexible and easy-to-implement intervention method for vulnerable clinical populations (Hsieh and Li, 2022), ensuring that the best possible therapy is available to many more people (Freeman et al., 2017). This study aims to design and evaluate the effectiveness of a VR forest bathing environment in improving well-being, stress reduction, and overall relaxation.

2 Methods

The research will be conducted using a mixed-methods approach, integrating qualitative and quantitative data collection and analysis methods. To develop the environment, a multidisciplinary team of experts in psychology, multimedia, and nature sciences will collaborate. They will apply state-of-the-art VR technology and design principles to create realistic and interactive immersive forest environments, incorporating elements such as visual cues, sounds, scents, and interactive features. User-centered design methodologies will be employed, involving user feedback and iterative design iterations to ensure the authenticity and effectiveness of the virtual environments. The evaluation of the VR forest bathing environments will involve a sample of participants, who will be randomly assigned to either the VR forest bathing group or a control group. The VR forest bathing group will engage in immersive experiences within the virtual environments, while the control group will engage in non-immersive digital nature experiences. Pre and post-intervention assessments will be conducted using standardized measures to assess subjective well-being, stress levels, and relaxation. Additionally, qualitative interviews will be conducted to gather in-depth insights into participants' experiences and perceptions of the VR forest bathing environments. The variables we propose to assess and study are divided into psychological, physical and physiological symptomatology. As previously mentioned, VR FT

sessions have the potential to improve psychological symptomatology, namely reducing anxiety and stress levels. Therefore, we propose the use of the Depression, Anxiety and Stress scale - DASS 21. The 21-item version is quick to apply, making it feasible to assess before and after each session. The DASS is composed of a set of three self-completion scales, which are intended to measure clinically significant symptoms of emotional states of depression, anxiety and stress (Ross and Jones, 2022). It is also considered relevant to assess well-being and, thus, we propose the use of the Portuguese version of the Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) (Figueiredo et al., 2022). The WEMWBS is aimed at assessing subjective and psychological mental well-being for the Portuguese population (Figueiredo et al., 2022). In order to monitor relaxation throughout the session and the activities performed, we propose the use of an EEG device for the purpose of recording heart rate (Ahn et al., 2019).

3 Results

The findings of this study have the potential to revolutionize the field of well-being interventions by harnessing the immersive capabilities of virtual reality technology. If successful, the development of VR forest bathing environments can address the limitations of physical access to natural environments, particularly for individuals residing in urban areas or those with mobility constraints. This research can also inform the design of future VR applications for health and well-being promotion, highlighting the importance of incorporating natural elements in virtual experiences.

4 Discussion

An example of how cutting-edge technology and ancient practices can coexist harmoniously is the use of virtual reality to reproduce the feeling of woodland bathing (Murtiyoso et al., 2023). Through this integration, people are able to utilize the advantages of a practice that is deeply ingrained in Japanese culture while overcoming physical and geographical constraints. This confluence highlights how adaptable time-honored methods are in a digital environment that is continuously changing. The VR forest bathing environments also tackles the problem of access to outdoor areas. This novel approach offers a democratization of well-being by offering a practical substitute for those who might otherwise be excluded from the advantages of forest bathing (McEwan et al., 2023). Many people may not be able to access physical forests due to factors like urbanization, physical disabilities, or time constraints. The success of the VR forest bathing setting should be compared to conventional, real forest habitats, nevertheless (Reese et al., 2022). Are the results similar? Are there any subtle distinctions between the two experiences' well-being improvements? More research is required to fully understand the psychological and physiological factors at work during a VR forest bathing session, therefore highlighting the potential benefits and drawbacks of each strategy would help. How do these reactions compare to or differ from those had while actually taking a woodland bath? Investigating how both experiences affect brain activity, heart rate variability, and cortisol levels can provide insight into the underlying mechanisms behind the reported gains in wellbeing. There are undoubtedly individual variations in the VR forest bathing environment's effectiveness. Some people could find the virtual experience to be just as restorative, while others might be more drawn to the actual, physical feelings of a forest. By examining these preferences, individualized interventions will be developed and improve the design of next VR wellness applications. But as VR technology develops, moral concerns about the possible disconnection from reality and the natural world can surface. Researchers should examine the ecological effects of lessening the need for physical travel to natural settings, potentially aiding in conservation efforts, and address ethical issues associated to creating a connection with nature through a virtual environment.

5 Conclusion

Forest therapy is an alternative form of therapy that makes use of physiological relaxation with proven benefits for mental health and well-being. FT can be applied in several situations such depression, anxiety, hypertension and affective disorders, post-traumatic stress, among others. Forest bathing involves immersing oneself in a forest atmosphere to enhance physical and mental health. VR technology is capable to produce a sense of realism and presence and can simulate interactive and immersive natural environments, allowing individuals to experience the benefits of forest bathing, regardless of their physical location and limitations of physical access to natural environments. Several studies demonstrate that the use of immersive environments has been effective in complementing the results obtained by conventional therapies. The control of stimuli produced in the nature environment by VR allows precision in the implementation of therapeutic strategies.

Acknowledgments

CITIC is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS).

Bibliography

J. W. Ahn, Y. Ku, and H. C. Kim. A novel wearable eeg and ecg recording system for stress assessment. *Sensors (Switzerland)*, 9(19), 2019.

E. Alyan, T. Combe, D. R. Awang Rambli, S. Sulaiman, F. Merienne, and N. Diyana. The influence of virtual forest walk on physiological and psychological responses. *International journal of environmental research and public health*, 21(18):11420, 2021.

S. H. M. Chan and L. Qiu. Nature in virtual reality improves mood and reduces stress: evidence from young adults and senior citizens. *Virtual Reality*, 1(April), 2021.

R. Chen, Y. Gao, R. Zhang, Z. Zhang, W. Zhang, H. Meng, and T. Zhang. How does the experience of forest recreation spaces in different seasons affect the physical and mental recovery of users? *International Journal of Environmental Research and Public Health*, 3(20), 2023.

S. Figueiredo, M. J. Trigueiro, V. Simões-Silva, T. Coelho, R. S. d. Almeida, P. Portugal, S. Sousa, F. Campos, and A. Marques. Tradução, adaptação e contributo para a validação da escala warwick-edinburgh mental well-being scale para a população portuguesa. *RevSALUS - Revista Científica Internacional Da Rede Académica Das Ciências Da Saúde Da Lusofonia*, 2(4):2, 2022.

D. Freeman, S. Reeve, A. Robinson, A. Ehlers, D. Clark, B. Spanlang, and M. Slater. Virtual reality in the assessment, understanding, and treatment of mental health disorders. *Psychological Medicine*, 14(47):2393–2400, 2017.

C. H. Hsieh and D. Li. Understanding how virtual reality forest experience promote physiological and psychological health for patients undergoing hemodialysis. *Frontiers in Psychiatry*, 13(13):1–17, 2022.

S. H. Jo, J. S. Park, and P. S. Yeon. The effect of forest video using virtual reality on the stress reduction of university students focused on c university in korea. *International Journal of Environmental Research and Public Health*, 23(18), 2021.

I. Lee, H. Choi, K. S. Bang, S. Kim, M. K. Song, and B. Lee. Effects of forest therapy on depressive symptoms among adults: A systematic review. *International Journal of Environmental Research and Public Health*, 3(14), 2017.

J. L. Maples-Keller, B. E. Bunnell, S. J. Kim, and B. O. Rothbaum. The use of virtual reality technology in the treatment of anxiety and other psychiatric disorders. *Harvard Review of Psychiatry*, 3(25), 2017.

K. McEwan, K. Krogh, K. Dunlop, M. Khan, and A. Krogh. Virtual forest bathing programming as experienced by disabled adults with mobility impairments and/or low energy: A qualitative study. *Forests*, 5(14):1033, 2023.

A. Murtiyoso, S. Holm, H. Riihimäki, A. Krucher, H. Griess, V. Christiane Griess, and J. Schweier. Virtual forests: a review on emerging questions in the use and application of 3d data in forestry. *International Journal of Forest Engineering*, 2023.

B. J. Park, W. S. Shin, C. S. Shin, P. S. Yeon, C. Y. Chung, S. H. Lee, D. J. Kim, Y. H. Kim, and C. E. Park. Effects of forest therapy on psychological improvement in middle-aged women in korea. *Journal of Preventive Medicine and Public Health*, 5(55):492–497, 2022.

D. V. Poulsen, U. K. Stigsdotter, D. Djernis, and U. Sidenius. 'everything just seems much more right in nature': How veterans with post-traumatic stress disorder experience nature-based activities in a forest therapy garden. *Health Psychology Open*, 1(3), 2016.

G. Reese, J. Stahlberg, and C. Menzel. Digital shinrin-yoku: do nature experiences in virtual reality reduce stress and increase well-being as strongly as similar experiences in a physical forest? *Virtual Reality*, 26(26):1245–1255, 2022.

A. M. Ross and R. J. F. Jones. Simulated forest immersion therapy: Methods development. *International Journal of Environmental Research and Public Health*, 9(19), 2022.

C. Song, H. Ikei, and Y. Miyazaki. Physiological effects of nature therapy: A review of the research in japan. *International Journal of Environmental Research and Public Health*, 8(13), 2016.

X. Wang, Y. Shi, B. Zhang, and Y. Chiang. The influence of forest resting environments on stress using virtual reality. *International Journal of Environmental Research and Public Health*, 18(16), 2019.

Y. Yi, E. Seo, and J. An. Does forest therapy have physio-psychological benefits? a systematic review and meta-analysis of randomized controlled trials. *International Journal of Environmental Research and Public Health*, 17(19), 2022.

S. Zhang, J. Teng, Y. Zeng, H. Song, and Z. Gu. The effects of forest therapy on public mental health and circular economy: A policy support model in japan. *Frontiers in Public Health*, 10 (10):1–14, 2022.

m-Health System for the Self-management of Patients with Heart Failure

Patrícia Silva, Paulo Veloso Gomes, Jorge Condeço, António Marques, Henrique Curado, Rita Veloso, and Javier Pereira

LabRP, Center for Rehabilitation Research, School of Health, Polytechnic of Porto, Porto, Portugal
School of Health, Polytechnic of Porto, Portugal
Executive Board Member Centro Hospitalar Universitário Santo António
CITIC Research Center, University of A Coruña, A Coruña, Spain
Serviço de Cardiologia, Centro Hospitalar do Tâmega e Sousa, Penafiel, Portugal
Correspondence: velosogomes@ess.ipp.pt

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: Increasingly, the use of technology is emerging as a tool to support health care. The new information and telecommunications technologies have triggered a new paradigm in health: in access, availability and management of information. Telemonitoring interventions in patients with heart failure have been shown to be cost effective in reducing the rate of hospitalisations and mortality. The development of m-Health applications focused on the centralization of care in patients, considering the perspectives of different stakeholders, promises results not only in clinical (mortality and morbidity) but also in patient outcome reports (quality of life and self-care). This work presents an integrated information system for telemonitoring and self-management of patients with heart failure.

1 Introduction

The sustainability of the national health system is one of the biggest challenges for policy makers in the health sector. With the increase in life expectancy, the aging of the population and advances in medicine, the prevalence of chronic diseases has increased exponentially throughout the world, leading to an increase in the consumption of health resources. Digital health emerges as a support tool to respond to the increased demand for healthcare, particularly in chronic diseases, increasing efficiency and access to healthcare and reducing healthcare costs.
Heart failure (HF) is a syndrome with high incidence and prevalence, affecting approximately 1-2% of the adult population in developed countries, with a prevalence of more than 10% in individuals over 70 years of age (McDonagh et al., 2021).
Training for self-management of patients with heart failure is essential for managing the disease and allows for improved adherence to pharmacological and non- pharmacological treatments, thereby reducing mortality, reducing hospitalizations and improving quality of life.
Technological development has brought new tools to support the care process for chronically ill patients, namely telemonitoring using mHealth systems. mHealth systems can be defined as the use of portable communication devices (such as cell phones or smartphones, wireless devices, and monitoring devices) that provide health services and information to support clinical and public health practice (Nouri et al., 2018).
Although there are a large number of mHealth systems, their use, even after being downloaded

many times, is low. Furthermore, many of the applications are of low quality, developed without taking into account scientific evidence or the needs and expectations of their users (Nouri et al., 2018). Software development is characterized by the sequencing of activities necessary to specify, design and test the results of the software under development, which must comply with the requirements defined for its development and with the national and international standards defined for the area (Silva and Videira, 2008). Requirements gathering is the most important step in software development, dictating its success or not. To identify the requirements, it is essential to know the needs and expectations of all its users, so that its use can be an added value for their daily lives (Hussain et al., 2016).
Given this paradigm, the present work consists of identifying the requirements for an mHealth system in the self-management of patients with heart failure, taking into account the perspectives of the various stakeholders in order to respond to their objectives, expectations and priorities.

2 Methods

To define the requirements for the development of an mHealth system for the self-management of patients with HF, the problem was contextualized in the follow-up of patients with HF. Subsequently, a systematic review of the literature was carried out on current mHealth systems and the benefits of their application. The work stages are systematized in the (Figure 2)

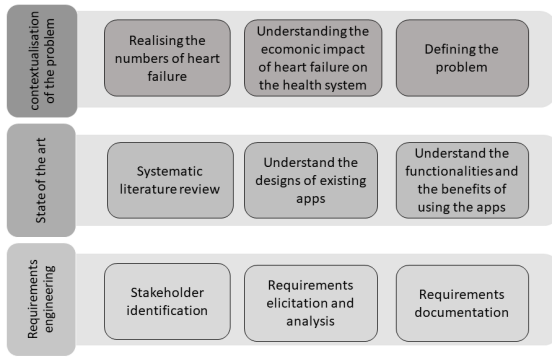


Figure 1: Study development phases

For requirements engineering, a qualitative methodological approach was chosen, which allowed exploring the perspectives of various stakeholders. The target population of the study were stakeholders involved in a care process for patients with heart failure in Portugal. The type of sampling was intentional – non-probability sampling, with the sample consisting of 25 elements. To this end, 4 focus groups were conducted:
- Group 1: Health professionals: 6 Cardiologists.
- Group 2: Health professionals: 5 Nurses.
- Group 3: 7 Patients.
- Group 4: 5 Informal Caregivers.
Interviews were also carried out with a health unit manager and a computer engineer.

3 Results

The systematic review of the literature allows to verify that the use of non-invasive telemonitoring using a smartphone is a useful tool in supporting the follow-up of patients with HF, demonstrating evidence of its effectiveness in improving the patient's self-care and quality of life, and in reduction in hospitalizations.
After analyzing the results from the perspectives of different stakeholders, the functional requirements of the VIVERcomIC app were divided into five main sections: Monitoring (daily surveillance); Alerts and feedbacks; Medication; Educational toturials and communication channels.
The application architecture is systematized in the 2)

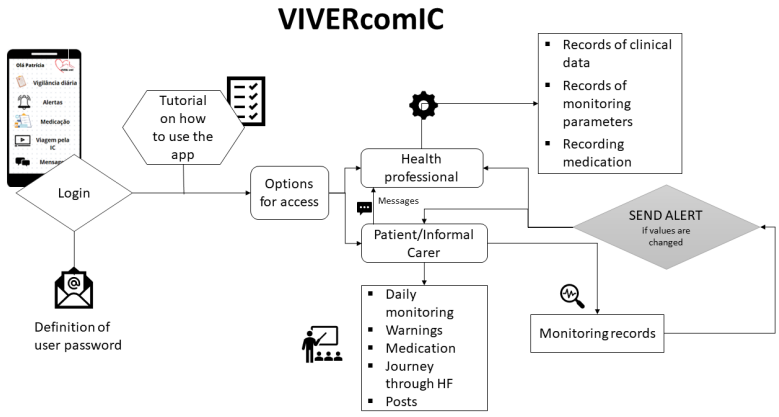


Figure 2: VIVERcomIC application architecture

The VIVERcomIC application aims to contribute to a strategy for recording the monitoring of signs and symptoms of the disease, allowing the patient access to feedback on their health status, and with this, more assertively resorting to healthcare.
In addition to monitoring, it provides educational tutorials on heart failure, enabling patients and informal caregivers to access information about their disease at their own pace and needs, and not just in specific moments of contact with healthcare professionals. Communication channels aim not only to facilitate contact with healthcare professionals, but also to provide feedback and alerts on health status, resulting in greater safety.

4 Conclusion

The emergence of new technologies has brought new models of care provision. The development of mHealth systems has proven to be an important tool in supporting patient training for self-management of heart failure. Reliability, accuracy of information and application interventions are the most discussed issues regarding the use of mHealth systems.
The functional requirements defined by stakeholders were divided into five main areas: Monitoring of signs and symptoms, alerts and feedbacks, recording of pharmacological therapy, educational tools and communication channels.
Therefore, the introduction of mHealth systems, more specifically VIVERcomIC, has the potential to positively promote changes in the health and well-being of patients, either by promoting health literacy, changing lifestyles, improving the self-management capacity of disease, namely by monitoring signs and symptoms of decompensation and early identification of warn-

ing signs, and by improving accessibility and communication between patients/informal caregivers and healthcare professionals.

Acknowledgments

CITIC is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS).

Bibliography

A. Hussain, E. Mkpojiogu, and F. Kamal. A systematic review on usability evaluation methods for m-commerce apps. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, 10(8):29–34, 2016.

T. A. McDonagh, M. Metra, R. S. Adamo, M.and Gardner, A. Baumbach, M. Böhm, H. Burri, J. Butler, J. Čelutkienė, O. Chioncel, J. G. F. Cleland, A. J. S. Coats, M. G. Crespo-Leiro, D. Farmakis, M. Gilard, S. Heymans, A. W. Hoes, T. Jaarsma, E. A. Jankowska, ..., and A. K. Skibelund. 2021 esc guidelines for the diagnosis and treatment of acute and chronic heart failure. *European Heart Journal*, 36(42):3599–3726, 2021.

R. Nouri, S. N. Kalhori, M. Ghazisaeedi, G. Marchand, and M. Yasini. Criteria for assessing the quality of mhealth apps: a systematic review. *Journal of the American Medical Informatics Association*, 25(8):1089–1098, 2018.

A. Silva and C. Videira. *ML, Metodologias e Ferramentas CASE, 2a Edição, Volume II*. Centro Atlântico. ISBN:978-989-615-061-7, Lisboa, 2008.

Indexing and Retrieval of Scores by Humming based on Extracted Features

Hilda Romero-Velo, Susana Ladra, José R. Paramá, and Fernando Silva-Coira

Universidade da Coruña, Laboratorio de Bases de Datos, Faculty of Computer Science, 15071 A Coruña, Spain
Universidade da Coruña, Centro de Investigación CITIC, 15071 A Coruña, Spain
Correspondence: h.rvelo@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: In order to be able to conduct searches over large collections of music scores with queries provided in audio format, this article considers recent literature in the field and proposes an implementation to extract specific features from music pieces. Afterwards, we index those features using modern Lempel-Ziv (LZ)-based data structures. These data structures take advantage of the intrinsic repetitiveness within music to reduce space consumption and, at the same time, to index the information optimizing the search time per query. Furthermore, taking advantage of this property-based representation framework, which does not depend on the way the music is portrayed, we enable the possibility to perform melodic searches by simply providing a query audio. This research branch is known as "query by humming" and has commonly been applied to audio sources. A preliminary study for its application in other forms of music representation is presented in this research.

1 Introduction

This paper presents a proposal to perform searches over large score collections using queries provided in audio format. Literature offers an extensive list of studies in the field of music information retrieval. Recently, Zhu et al. (2022) published an article that proposes a new approach: extracting three features describing a composition directly from the score and then, in their scenario, indexing them with the Elasticsearch search engine. However, this paper defines the theoretical framework for obtaining such characteristics, but does not specify the practical procedure of how to obtain them. In this article, we provide an implementation that allows such features to be derived. Then, those features are stored and indexed using modern Lempel-Ziv (LZ)-based data structures (Fariña et al., 2019; Kreft and Navarro, 2013). More specifically, from this family of auto-indexes, we use those based on LZ77 and LZ-End. These data structures take advantage of the intrinsic repetitiveness of music to save space consumption and, at the same time, to index the information so as to optimise the search time per query. Within this context, the ability to perform approximate searches is of utmost importance, since the query (audio excerpt) may not necessarily reflect exactly the same specifications as the original score does.

The most common approach is to take a specified audio query (analogue signal) and look at the stored audio data to find its equivalents. On the other hand, Zhu et al. (2022) depart from queries that specify the notes of the fragment to be compared on a score, which, nevertheless, is not very practical. This document explores the feasibility of offering a new querying method based on audio search. Contrary to widely used solutions such as Shazam (Wang,

2003), which requires the audio query to be a replica of a fragment of the original song's audio, or like Google's proposal (Kumar, 2020), which works with the fingerprint of each tune, our searches are independent of the format in which the original piece is collected, allowing queries over collections of scores. Therefore, by applying the principle of feature extraction on the query audio itself, score searches are enabled in the field referred to as "query by humming".

For this study, we work with 1,275 scores obtained from the Folkoteca Galega[1] under the MusicXML format. The corresponding melodic sequences are inferred from each one of them, and afterwards the proposed features are extracted, resulting in a set of 1,686 melodies to be indexed in our system. These three characteristics are the chromatic distance as well as the diatonic interval between notes, and the rhythm difference ratio between figures. In order to easily manipulate the scores, a conversion from MusicXML[2] format to **kern format is made using the Humlib library (Sapp, 2023), and then the previously mentioned features are extracted through the toolkit provided by Humdrum (Huron, 2022).

A similar approach is followed for queries generated from audio. First the audio is transformed into Musical Instrument Digital Interface (MIDI[3]) format. This task is a complex domain of study by itself, and even though we are dealing with a simple case (conversion of a monophonic sound of a single instrument), inaccuracies can be raised during the translation process between formats. For this entry, the *audio-to-midi* converter developed by Spotify is used (Bittner, 2022). Once the query is obtained in MIDI format, we transfer it to **kern representation using the Humdrum tools and, then again, we extract the features corresponding to that fragment to run the query against our index.

2 Previous Concepts

In this section, we will review the background knowledge and define certain aspects needed to understand the course of the case study. We will first describe the features to be extracted from the scores and then we will consider the musical representation formats involved.

2.1 Features

Three defining characteristics of a musical piece are extracted from the scores:

- **Chromatic Distance:** The first feature to be obtained is the chromatic distance between notes. This is calculated according to the chromatic scale, which consists of 12 notes. We start measuring from zero and proceed by counting the exact distance in semitones between the two notes involved. For instance, from C to high C there would be a distance of 12 semitones.
- **Diatonic Distance:** This second feature describes the distance between notes according to their position in the diatonic scale. The diatonic scale is formed by 7 notes. An example would be the natural scale of C major: C, D, E, F, G, A, B. In this case the distance is calculated starting at 1. Therefore, the unison, i.e. the distance between two consecutive notes which have the same pitch (for instance, a F followed by another F), in this scale would be represented by 1, whereas in the chromatic scale it would be marked as 0. Bringing the previous example, the diatonic distance from C to high C would be 8, which is commonly known as an octave.
- **Rhythm Ratio:** The rhythm of the piece is shown by indicating the ratio of difference in duration between a note and the preceding figure. In Figure 1, we observe that the value given to each figure doubles in length as we climb up the tree. Thus, a whole note is equivalent to 2 half notes, 4 quarter notes and 8 eighth notes.

1 *https://folkotecagalega.gal/pezas*
2 *https://www.musicxml.com/*
3 *https://www.midi.org/*

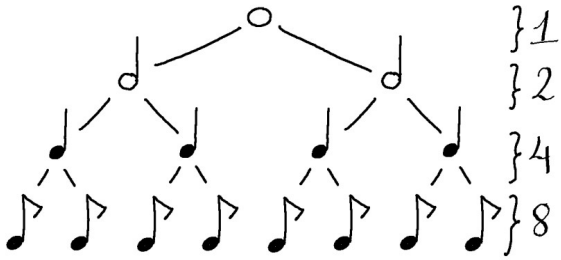


Figure 1: Rhythmic Figures Tree with Numbered Equivalences from 1 to 8

In the example of Figure 2, we distinguish 4 notes inside a score, accordingly, we can compute 3 distances. Those intervals within the notes are marked as 1, 2, 3 beneath them. Taking a piano as our perspective for the computation, the calculus of the chromatic distance considers the black notes and starts calculating at 0. Meanwhile, for the diatonic distance of this example we do not take into account the black notes, we only reckon the white ones and start counting at 1. The chromatic distance would be 1 - 4 - 3 respectively, yet the diatonic distance would be 2 - 3 - 3. We can verify, then, that, although the way of obtaining each feature is different, we can obtain the same values in certain cases, as in the last interval in the example. The rhythmic description according to the figure ratio is 1/2 - 2 - 1/2 since the second note is half as long as the first, the third note twice as long as the second and the fourth one half as long as the third (refer to Figure 1).

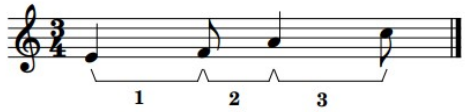


Figure 2: Score Excerpt with Highlighted Intervals 1, 2 and 3

2.2 Representation Formats

Both MusicXML and **kern (Huron, 2022) are schemas that allow the representation of scores with text notation. While the former focuses on rendering, the latter works with a functional format aimed at facilitating score manipulation and analysis. Thus, in this article we opted to rely on the **kern format in order to handle feature management and extraction.

On the other hand, MIDI files are designed to record events. Those events are messages that denote a specific occurrence regarding certain settings. Such files are useful for device-to-device communication and for synthesising the instructions into audio. However, they are not intended to describe scores. Some properties of these MIDI files that allow defining and setting the tempo, and hence defining the rhythmic figuration, may be misformed, mainly when performing a conversion from audio (due to the difficulty in inferring these parameters).

2.3 Lempel-Ziv (LZ)

The idea of Lempel-Ziv compression is based on looking for repeated sequences within the data. When found, this sequence is replaced by a reference to the location of the first sequence along with the length of the repeated pattern. If there are no such occurrences, each byte is written as a literal.

Example of LZ compression: Let the sequence to be compressed be [ABCBCBC]. LZ would phrase this stream as [ABC2:4]. The notation 2:4 indicates that starting from 2 positions behind, the next 4 characters are to be repeated. Therefore, the decompression process would be the following (the character to be written next in the string is highlighted in bold): [A**B**C] - [AB**C**B] - [ABC**B**C] - [ABCB**C**B] - [ABCBCBC].

In this scenario, however, there are different implementations depending on certain factors. A key point in compression is the "historical" window and the "future" window. In other terms, how much of the already processed data we store in memory to check if any future pattern is repeated, and how much ready-to-process data we keep in memory to seek those sequences in the already processed data. Another aspect to consider is the way a pattern is referenced and when should this happen, since the space needed for referencing should be less than the space required for storing the literals alone.

3 Implementation

When extracting features, we need to consider that there are two different source formats: the original score in MusicXML and the queries in audio format. For our particular scenario, we will assume only the Waveform Audio Format (WAV). In the following, we will discuss each of the cases.

Scores

The Folkoteca Galega is a web portal that gathers more than 1,200 pieces of traditional Galician music. It is classified according to the type of composition (jota, muiñeira, pasodoble, polka, etc.). For each song, it provides the score in MusicXML format, as well as in MIDI and PDF. Its main collaborator is PuntoGal, but any contribution is appreciated. Different associations and individuals related to the folk world have submitted scores to this collection.

Taking this corpus from the Folkoteca Galega in MusicXML, we use the *musicxml2hum* tool provided by Humlib (Sapp, 2023) to obtain those same scores in **kern format. We then proceed to remove information from the composition that is not required for defining the characteristics (e.g. dynamics). Afterwards, we extract the different melodic lines (voices). All of this is done by using the *extractx* command, also provided by Humlib. Thus, the 1,275 scores are fragmented into 1,686 melodic lines, i.e. 1,686 files.

After this preliminary preparation, the *humsed* command is used to remove the grace notes (ornaments). We use the *mint* command for the diatonic analysis and the *beat* command for the rhythmic analysis, both from the Humdrum toolkit. Meanwhile, to perform the chromatic analysis, we need to convert the **kern file into **semits, that is, into numerical semitone representations of the pitch (*semits* command). We must also remove rests so that they do not interfere with the calculation (once again with the *humsed* command). Finally, we can compute the semitone offsets with the *xdelta* command plus the *-s ^=* option to prevent the bar lines from interfering.

Having obtained all the features, for the melodic ones (chromatic and diatonic) we generate the files with the corresponding documents to be indexed according to the specification of the LZ implementation proposed by Fariña et al. (2019).

Audio Query

The Spotify basic-pitch converter is used to transform audio queries in WAV format to MIDI. However, even with a clear audio source, anomalies are generated during

the conversion process which disables exact searches. Merely to rectify this situation, it is therefore natural to conduct the research towards approximate searches. Nevertheless, for the simplicity of these first stages, after conversion, the MIDI files are manually polished (e.g. by removing phantom notes and harmonics). Even so, it is noted that the length of the notes is not properly transformed, as expected from the earlier observations regarding the misformation of MIDI files. It is precisely the inaccuracy of the transformation that causes some real notes to be perceived as grace notes which leads for them to not be taken into consideration during feature extraction.

Once the audios have been converted to MIDI, we use Humdrum's *mid2hum* command to get the **kern queries. Thereafter, feature extraction is continued in the same way as for the scores previously. Afterwards, they are adjusted to the requested format by the LZ implementation. The recovery of the expected number of occurrences is verified.

4 Experimental Evaluation

For this case study our aim is to test the proper indexing of the extracted features by running 20 exact search queries. Furthermore, we will verify the robustness of the implementation framework.

Evaluation Environment

The experiments were conducted on a computer with an Intel(R) Core(TM) i7-10750H and 16GB of DDR4 RAM with a Microsoft Windows 10 Pro operating system. The Docker version used was 4.17.0. For this preliminary research we took as a starting point the original corpus of 1,275 documents held in the Folkoteca Galega[4]. The features extracted from this collection yield 1,686 documents and occupy a total of 598.65 KB. Features have been adapted and transferred to the required format and layout in the test framework configuration provided in the form of Docker image with an implementation of LZ-based indexes by Fariña et al. (2019).

Each query is generated by recording in audio format a piano performance of the fragment to be searched. A total of 20 queries of varying lengths are collected. The displayed times reflect the average of 10 runs for the queries.

Results Analysis

Results indicate that the query time depends on the number of occurrences to be retrieved (Tables 2 and 3). For both chromatic and diatonic features, LZ-End offers a lower average recovery time per occurrence and a smaller standard deviation than LZ77, at detriment of a higher space requirement (Table 1). The difference between LZ77 and LZ-End lies in the segmentation of the repeated strands. While LZ77 opts for long repeating strands, LZ-End chooses to frame the repeats in smaller chains. Thus, LZ-End requires more space but offers faster substring extraction.

We also observe that the first four queries do not retrieve any occurrences even though they should. This is caused by the incorrect MIDI file formation during audio conversion. In this particular case, the necessity of moving towards an approximate

[4] *https://folkotecagalega.gal/pezas*

Table 1: Comparison of Sizes (KB) and Times per Occurrence (ms)

Index	Size	Feature	Average Time per Occurrence	Standard Deviation
LZ77	344.43	Chromatic	0.302	0.595
		Diatonic	0.467	0.907
LZ-End	408.43	Chromatic	0.205	0.379
		Diatonic	0.294	0.676

Table 2: Diatonic Results per Query ordered by Number of Occurrences

Number	Query Length	Occurrences	LZ77 Time (ms)	LZ-End Time (ms)
1	24	0	9.86	4.51
2	34	0	2.63	1.89
3	24	0	1.87	1.20
4	31	0	2.40	1.53
5	15	1	3.55	2.96
6	16	2	3.83	0.87
7	25	2	1.84	1.12
8	20	2	1.45	1.02
9	31	2	2.66	1.54
10	13	4	2.94	2.15
11	17	10	1.34	0.88
12	2	57	0.26	0.25
13	3	1,781	4.56	4.24
14	5	2,461	6.17	5.96
15	4	3,857	9.54	9.47
16	4	4,611	11.65	11.39
17	1	5,414	13.15	12.48
18	3	7,867	19.11	19.97
19	2	16,450	40.11	38.25
20	2	36,263	89.73	84.61

search is reflected. Lewenstein (2013) proposes an alternative to enable approximate searches on LZ indexes, which would overcome the loss of information in translations.

Altogether, both implementations, LZ77 and LZ-End, perform at a high level. They are compact, fast, and, consequently, efficient. The incidence is that they do not solve some queries properly due to the different casuistry involved. We, therefore, conclude that the approach should be directed towards approximate searches.

5 Conclusions

This paper introduces a preliminary study to offer score retrieval by humming. In order to take advantage of musical patterns, a feature extraction framework is established to connect the different representation formats. Thereby, the indexation takes place to exploit the intrinsic repetitiveness of the music.

After studying the causalities of some occurrences, it is concluded that the next step should be working with approximate searches in order to achieve the flexibility that

Table 3: Chromatic Results per Query ordered by Number of Occurrences

Number	Query Length	Occurrences	LZ77 Time (ms)	LZ-End Time (ms)
1	24	0	1.56	1.15
2	34	0	2.32	1.48
3	24	0	1.78	1.26
4	31	0	2.06	1.47
5	15	1	1.00	0.71
9	31	1	2.41	1.47
6	16	2	1.05	0.74
7	25	2	1.66	1.08
8	20	2	1.67	1.30
10	13	4	0.84	0.65
11	17	6	1.20	0.98
12	2	45	0.31	0.22
14	5	674	2.84	2.29
13	3	836	2.97	2.12
15	4	1,074	3.81	2.82
16	4	1,645	5.58	4.16
18	3	2,797	9.51	7.13
17	1	5,061	15.72	12.16
19	2	10,564	35.08	25.40
20	2	15,086	50.16	37.48

the system requires. The impact of the issues derived during the translation process from audio to MIDI could be minimize if we focus on approximate search indexes such as those proposed by Moshe Lewenstein in "*Orthogonal Range Searching for Text Indexing*" (Lewenstein, 2013).

Acknowledgements

Work funded by: CITIC is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS), 80% through FEDER funds, Galicia Operational Programme FEDER 2014-2020, and the remaining 20% by the "Secretaría Xeral de Universidades" (Grant ED431G 2019/01), Xunta de Galicia/FEDER-UE [ED431C 2021/53]; Ministry of Science and Innovation [PID2020-114635RB-I00; PDC2021-120917-C21; PDC2021-121239-C31; PID2019-105221RB-C41; TED2021-129245-C21].

Bibliography

R. Bittner. Meet basic pitch: Spotify's open source audio-to-midi converter, jun 2022. URL *https://engineering.atspotify.com/2022/06/meet-basic-pitch/*.

A. Fariña, M. A. Martínez-Prieto, F. Claude, G. Navarro, J. J. Lastra-Díaz, N. Prezza, and D. Seco. On the reproducibility of experiments of indexing repetitive document collections. *Information Systems*, 83:181–194, jul 2019.

D. Huron. The humdrum toolkit for computational music analysis, 2022. URL *https://www.humdrum.org/index.html*.

S. Kreft and G. Navarro. On compressing and indexing repetitive sequences. *Theoretical Computer Science*, 483:115–133, 2013. Special Issue Combinatorial Pattern Matching 2011.

K. Kumar. Song stuck in your head? just hum to search, oct 2020. URL *https://blog.google/products/search/hum-to-search/*.

M. Lewenstein. *Orthogonal Range Searching for Text Indexing*, pages 267–302. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

C. S. Sapp. Humlib: Humdrum data parsing library in c++, aug 2023. URL *https://humlib.humdrum.org/*.

A. Wang. An industrial-strength audio search algorithm. *Columbia.edu*, 2003. URL *https://www.ee.columbia.edu/~dpwe/papers/Wang03-shazam.pdf*.

T. Zhu, R. Fournier-S'niehotta, P. Rigaux, and N. Travers. A framework for content-based search in large music collections. *Big Data and Cognitive Computing*, 6(1), 2022.

Distributed Database Model for Mobile Health Telemonitoring Applications

Paulo Veloso Gomes, Henrique Curado, António Marques, Bárbara Gomes, and Javier Pereira

LabRP, Center for Rehabilitation Research, School of Health, Polytechnic of Porto, Porto, Portugal
CITIC Research Center, University of A Coruña, A Coruña, Spain
Correspondence: velosogomes@ess.ipp.pt

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: Telemonitoring interventions in patients allow to collect and transmit information to health professionals bringing recognized advantages in the personalization of health care services. Those applications deal with sensitive data, raising ethical and legal issues that must be considered. Guarantee the quality and accuracy of the data collected in telemonitoring and ensure confidentiality to protect the patient privacy is a fundamental requirement. The patient's informed consent implies that he is aware of the potential benefits and risks of the system, how telemonitoring will be used, what data will be collected and with whom it will be shared. This work proposes a database model ensuring that sensitive data is handled securely and accessible only to authorized health professionals.

1 Introduction

The evolution of communication and information technologies opens up new opportunities for improving the provision of healthcare. The evolution, portability and dissemination of new devices that allow sophisticated applications and facilitate the communication process through wireless networks, allows the creation of new networks and forms of communication between healthcare professionals, patients and informal caregivers.
The concept of telemonitoring opens new opportunities in the area of healthcare provision. Its application can have different purposes, counseling, monitoring, communication between the user and the healthcare professional, selfcare management, among others.
The development of telemonitoring applications focused on the centralization of care in patients must consider the perspectives of different stakeholders. Despite the advantages that telemonitoring can provide to patients, healthcare professionals and the healthcare system in general, there are also important challenges that must be carefully studied, so as not to compromise the quality of healthcare services provided, patient privacy and the protection of healthcare professionals.
Telemonitoring applications raise some important challenges, such as data protection, confidentiality, reliability, accuracy of information, ensuring that the patient is effectively and properly informed before giving informed consent and that the healthcare professional can prove afterwards that provided a quality service in a timely manner. Studies show that privacy and security are the main concerns of users of this type of systems (Houser et al., 2023), (Pool et al., 2022). Technology constraints and digital literacy are other important challenges mentioned (Houser et al., 2023).
In this work, we seek to propose a database model that guarantees that sensitive data are

treated, at the telemonitoring level, in a secure manner and accessible only to authorized health professionals, it is essential that the privacy and informed consent of the patient is guaranteed, the basis of any intervention. It is important to note that we are in the field of telemonitoring, through mobile applications that are not owned by healthcare providers, when there is already legal support for the collection of data in databases, in accordance with the General Data Protection Regulation (GDPR).

2 Trust, Confidence and Privacy

The doctor-patient relationship is based on trust. This, however, is not unilateral. Not only is it imperative that the patient trusts his doctor and understands the information he gives him, it is also essential that the doctor obtains all the necessary information, without omissions, to know how to advise the person. This double dimension is even more relevant in terms of telemonitoring, as it requires the patient to provide the necessary data in an accurate and timely manner. This presupposes that patients have confidence in the system and the guarantee of confidentiality of their data. On the other hand, it presupposes that the doctor is able to demonstrate that, given the available data, he did everything in terms of providing the best advice in accordance with the Clinical Guidance Standards and his objective duty of care. This scenario falls within the domain of medical, ethical, civil and criminal liability.
It is not intended to analyze issues of medical liability here, as they cannot be dissociated from telemonitoring. Indeed, if an error occurs during the process, this may be due to the doctor, his patient, but also to the mechanism used, as machines, programmed by humans, are also fallible (Lança, 2022). It is clear, therefore, that the health professional must have at his disposal means that allow him to prove his lack of guilt or negligence. However, this is only possible if the healthcare professional has access to their patients' sensitive data, under penalty of, in the face of an accusation, whether founded or not, being unable to obtain elements necessary for his defense.
This dimension rules out the possibility of telemonitoring mechanisms allowing access to health professionals only for a limited period, during which they would make a statement, leaving the database only with the holder of the same, the patient. For this, but also for another reason, which is the need for data analysis to presuppose the perception of clinical history, access to health data in telemonitoring must be complete. This aspect thus falls into the domain of health information security. It should be noted that at this level, not only data protection legislation is relevant, ethical questions also arise.

3 Informed Consent

Informed consent must exist at two levels. At the first level, the patient, when accessing the telemonitoring process, will have to give their consent for sensitive data to be processed (data concerning health, as defined in article 4 of the GDPR, is "personal data related to the physical or mental health of a natural person, including the provision of health care services, which reveal information about his or her health status"). Thus, consent is in the sense of the GDPR, which presupposes not only collection, but also access, and its transfer and deletion.
On the other hand, when the doctor in the monitoring process provides an intervention, which may constitute a medical act, it is necessary that, in the interests of his/her release from responsibility, he/she is aware that the patient, when carrying out the act (clinical, pharmacological, etc.) has was completely clear about the scope of what was advised. This is another level of consent, which, although normally called informed consent, also consists of a therapeutic clarification (Pereira, 2020), because if it is advised to take or change a drug, the patient will have to be duly informed of possible side effects or adverse reactions.
Finally, the ethical issue must be borne in mind. In effect, the doctor must, when taking on the telemonitoring process, in addition to the ethical obligation of confidentiality, which does not

have any incompatible relationships, namely, by providing services in private groups, or having any financial participation in them, for the which monitored data may have any economic interest, directly or indirectly (which will occur if the doctor works in a clinic owned by a health insurance group).
In the field of monitoring, whether of chronic patients or less complex processes, through electronic devices, usable remotely, making the privacy of health data compatible with the responsibility of the healthcare professional requires a careful and balanced approach. It is essential to prioritize the protection of sensitive patient data while allowing professionals to provide the best care possible. This can be achieved through secure technology in compliance with data privacy and medical liability legislation, through the use of mutual intervention mechanisms: the patient's double consent; the medical assumption of guaranteeing total confidentiality, through the non-consent agreement for third parties to access the information.
Despite its unquestionable contribution to individual and collective health, when dealing with health information, which is highly sensitive, the protection of privacy based on a system that merely refers to the GDPR, does not provide effective protection (Curado et al., 2023). A Distributed Database Model for Mobile Health Telemonitoring Applications (Figure 1) must consider a flexible use of data, ensuring that data is only accessed by those who need it and have the necessary access permission.

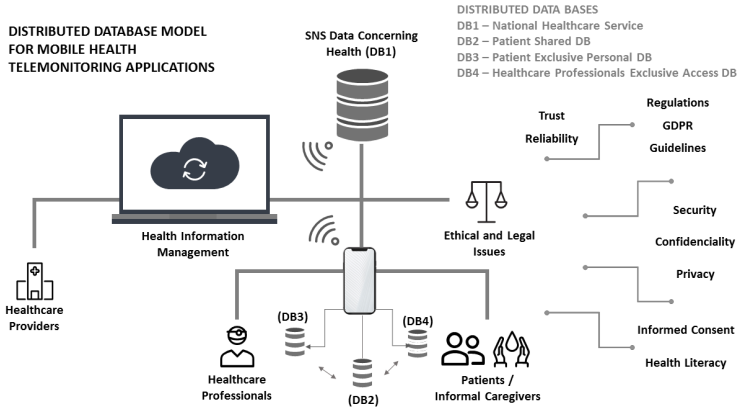


Figure 1: Distributed Database Model for Mobile Health Telemonitoring Applications

Therefore, the proposal to develop a "mobile application that allows users of the National Health Service, who are in Telemonitoring programs, to record measurements of biometric parameters and self-assessments", deserves a more careful analysis of the proposed terms and conditions of use recommended by Shared Services of the Ministry of Health (https://www.spms.min-saude.pt/telemonitorizacao-sns).

4 Conclusion

The new information and communications technologies have triggered a new paradigm in healthcare, allowing telemonitoring systems the access, provision and management health information. Telemonitoring applications focused on the centralization of care in patients have advantages in accessing and providing healthcare. But conditions that safeguard security, privacy and confidentiality must be guaranteed so that patients, informal caregivers and healthcare professionals can trust the system.
To guarantee data security and privacy, it is necessary that the data model used by applications

for mobile health telemonitoring ensures that data is only accessed by those who need it and have the necessary access permission.

Acknowledgements

CITIC is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS).

Bibliography

H. Curado, P. Veloso Gomes, M. Jacquinet, A. Marques, and J. Pereira. Strategy for data cybersecurity in european health data ecosystem. In *Proceedings of V XoveTIC Conference. XoveTIC 2022*, volume 14 of *Kalpa Publications in Computing*, pages 35–37. EasyChair, 2023.

S. H. Houser, C. A. Flite, and S. L. Foster. Privacy and security risk factors related to telehealth services - a systematic review. *Perspectives in Health Information Management*, 1(20), 2023.

H. Lança. *Inteligência Artificial e Tecnologia no Mundo da Medicina: Prolegómenos*. Editora D'Ideias. ISBN: 978-989-53817-0-8, In Inteligência Artificial, Tecnologia e Cuidados de Saúde, coord. Inês Fernandes Godinho, António Tavares. - 1ª ed. - Coimbra. 93-104, 2022.

A. G. D. Pereira. *Consentimento informado e bens jurídicos no direito penal e no direito civil. In Consentimento informado em Direito Civil e Penal. Lopes, E.9-32*. Centro de Estudos Judiciários. ISBN: 978-989-9018-27-3. https://cej.justica.gov.pt/LinkClick.aspx?fileticket=kBS6KaykToo

J. Pool, S. Akhlaghpour, F. Fatehi, and L. C. Gray. Data privacy concerns and use of telehealth in the aged care context: An integrative review and research agenda. *International Journal of Medical Informatics*, 160, 2022.

Introducing a Human Activity Recognition Dataset Gathered on Real-Life Conditions

Daniel Garcia-Gonzalez, Enrique Fernandez-Blanco, Daniel Rivero, and Miguel R. Luaces

Department of Computer Science and Information Technologies, University of A Coruna, CITIC, 15071 A Coruna, Spain
Correspondence: d.garcia2@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: Human activity recognition (HAR) has garnered significant scientific interest in recent years. The widespread use of smartphones enabled convenient and cost-effective data collection, eliminating the need for additional wearables. Given that, this paper introduces a novel HAR dataset in which participants had freedom in choosing smartphone orientation and placement during activities, ensuring data variability. It also includes contributions from diverse individuals, reflecting unique smartphone usage habits. Moreover, it comprises measurements from accelerometer, gyroscope, magnetometer, and GPS, corresponding to one of four activities: inactive, active, walking, or driving. Unlike other datasets, the collected data in this study were obtained from smartphones used in real-life scenarios.

Introduction

Emerging as a distinct field of study, human activity recognition (HAR) has gained substantial attention due to its precise classification of diverse human actions. This field focuses on classifying the activities executed by different individuals by analysing data obtained from an array of sensors (Aggarwal and Xia, 2014). These sensors capture information while subjects carry out predetermined movements like nodding, raising a hand, walking, running, or driving. Wearable devices like activity wristbands and smartphones have proven immensely valuable in generating such data. More specifically, smartphones, with their abundant sensors and convenient dimensions, offer a user-friendly means of gathering high-quality data. Additionally, the insights derived from people's behavioural patterns, as tracked by these sensors, provide for various domains such as healthcare, fitness, and home automation, thus enhancing the capabilities of these sectors (Zahin et al., 2019; Zhu et al., 2015). This fusion of widespread sensing facilitated by smartphones and the consequent model development has given rise to a flourishing research domain that has garnered escalating interest in recent years (Demrozi et al., 2020; Lara and Labrador, 2012).

However, there are two main challenges encountered in this field. Firstly, efficiently managing the vast amount of data produced by those devices, alongside their temporal interdependence, is a significant hurdle. Secondly, establishing the correlation between this data and predefined movements remains an ongoing enigma. Some methodologies have demonstrated impressive outcomes in extracting insights from sensor data (Attal et al., 2015; Hassan et al., 2018). Nonetheless, it is noteworthy that these studies often involve adapting devices for specific placement, like attaching them to various body parts such as the wrist or waist. As a result,

[1] This document is heavily based on the contents of Garcia-Gonzalez et al. (2020).

the effectiveness of these models might be skewed, given the controlled setting in which data is gathered, encompassing limited activities and particular device orientations.

Moreover, the controlled environment established for those purposes significantly deviates from real-world scenarios. Users, especially those who employ smartphones, exhibit diverse manners of carrying those devices. In addition, individuals may vary their clothing choices, resulting in discrepancies in the orientation and placement of the devices. This variability in user behaviours and situations clearly contrasts with the hypothetical ideal. Notably, the artificial intelligence (AI) models introduced so far are markedly influenced by factors like placement and orientation. This reliance restricts their widespread applicability and obstructs their seamless integration into practical situations. In this way, the lack of a smooth transition to real-life circumstances constitutes a significant disparity. Nowadays, the AI models developed for HAR are intricately linked to specific orientations and positions. As a result, these models lack the flexibility to adapt universally, thereby limiting their extension to diverse user types. Consequently, the effort to personalise AI models for human activity recognition across a large spectrum of individuals remains an ongoing avenue of investigation. In fact, this exploration has persisted for nearly ten years (Solis Castilla et al., 2020; Weiss and Lockhart, 2012).

For the previously mentioned reasons, a new dataset was gathered, looking to close the gap with the real-life application. Specifically, it was collected using the sensors of 19 individuals' smartphones, with almost complete freedom. In this way, the differences in the orientation and placement of the used devices are various, as well as the physical characteristics of each participant.

Data collection

The data collection process was made through a custom Android application developed by the authors, which streamlined the process of recording, categorising, and storing data. In this way, an initial data collection phase was undertaken, spanning approximately one month. The objective here was double: comprehending the nature of the acquired data and performing preliminary assessments. Then, a more intensive data gathering effort was conducted over a period of about one week. That served to rectify the imbalances and weaknesses identified in the preceding phase. The 19 participants in the study were instructed to specify the activity they were about to engage in using the previously mentioned Android application before commencing data collection. In such a manner, upon activity selection, data acquisition began automatically and stopped when the user indicated the conclusion of the activity. As a result, each recorded session corresponded to a distinct activity undertaken by a specific individual. The activities performed were classified into four categories:

- Inactive: not having the smartphone in motion. This consisted of any activity that involved not carrying the smartphone.
- Active: moving with the smartphone, without a specific destination. Activities such as preparing dinner, attending concerts, shopping for groceries, or doing household chores fell under the "active" category.
- Walking: any movement towards a defined location. Activities like running or jogging were categorised as "walking".
- Driving: every movement via motorised transportation, without the need to be the person driving. This included vehicles like cars, buses, motorbikes, trucks, and similar modes of conveyance.

Given that, data gathering originated from four distinct sensors: accelerometer, gyroscope, magnetometer, and GPS. The selection of accelerometer and gyroscope was predicated on their prevalence in the existing literature and their demonstrated efficacy. Furthermore, the magnetometer and GPS were also incorporated, proposing their utility in addressing this chal-

lenge. Specifically, GPS could play a crucial role in distinguishing those activities by detecting the user's movement velocity while carrying the smartphone. Anyhow, the data from the accelerometer, gyroscope, and magnetometer were stored with their tri-axial values. In the context of GPS data, the device's latitude, longitude, and altitude increments were stored alongside the measurement's bearing, velocity, and accuracy. In addition, the accelerometer data were refined by using the gravity sensor. That involved subtracting the last reading of the latter sensor from the observations of the former one, yielding refined accelerometer values (referred to as linear accelerometer values). These values remained unaltered by the smartphone's orientation, resulting in a dataset independent of the user's location and the device's orientation, as initially intended.

On the other hand, before local data storage, a series of filters were applied. Concerning the accelerometer and magnetometer, a low-pass filter was employed to mitigate excessive noise within the measurements from these sensors. In contrast, for the gyroscope, which confronts the well-recognised gyro drift issue, a high-pass filter was adopted as a workaround. Nevertheless, there was a challenge posed by Android, as each sensor could not be uniformly set to the same frequency. This situation proved particularly intricate in this context, with the need to merge data from sensors with highly disparate frequencies, such as the high-frequency accelerometer and the low-frequency GPS. While the accelerometer can yield as many as ten or even fifty measurements per second, the GPS provides new measurements approximately every ten seconds. Regrettably, Android's inherent constraints require the acceptance of receiving values as provided by the system, leading to potential data gaps. These gaps are particularly pronounced in the GPS data, wherein instances might arise when no new measurements are captured for over a minute (albeit this could be attributed to challenges related to enclosed environments). Similar gaps also emerge in accelerometer, gyroscope, and magnetometer data, notwithstanding their respective frequencies of around 10, 5, and 8 measurements per second under stable conditions. These gaps typically span 1 to 5 seconds, predominantly at the outset of each data collection session, although they occur less frequently than in GPS readings. Nevertheless, the average count of recordings per second for each sensor and activity is showcased in Table 1, along with the resulting mean frequency. A smaller font size beneath each average value outlines the corresponding standard deviation for each category. Notably, activities entailing movement, such as "active" or "walking", experience an elevation in these measurements, particularly noticeable with the accelerometer. The smartphone's sensors automatically heighten their frequency to derive maximum information from movements detected during these activities. This augmentation extends to the "driving" activity as well, possibly attributed to vehicular vibrations that the smartphone sensors might also detect. Furthermore, in instances of "walking" and "active" activities, intermittent periods of inactivity, such as waiting at traffic lights or moments of standing engagement, contributed to a moderation in these average frequencies.

Table 1: Mean recordings per second for each sensor and every measured activity.

Activity	Accelerometer Hz.	Gyroscope Hz.	Magnetometer Hz.	GPS Hz.
Inactive	11.00 $\pm$16.38	4.66 $\pm$0.74	7.91 $\pm$11.72	0.13 $\pm$0.35
Active	32.55 $\pm$24.80	4.46 $\pm$1.44	9.13 $\pm$13.64	0.06 $\pm$0.23
Walking	31.24 $\pm$27.47	6.24 $\pm$11.86	8.16 $\pm$12.05	0.06 $\pm$0.23
Driving	51.16 $\pm$31.59	4.66 $\pm$2.42	17.00 $\pm$20.01	0.04 $\pm$0.20

Table 2: Data distribution for each measured activity in the dataset.

Activity	Recorded Time (s)	Number of Recordings	Number of Observations	Percentage of Data
Inactive	292,213	147	7,064,757	24.25%
Active	178,806	99	8,918,021	30.62%
Walking	98,071	200	4,541,130	15.59%
Driving	112,226	128	8,602,902	29.54%
Overall	681,316	574	29,126,810	100%

As a result, the ultimate distribution of activities within the dataset is presented in Table 2. This table illustrates the total recorded time, the number of recordings, the number of observations, and the corresponding data percentage (the latter being related to the number of observations) for each designated activity. Here, a "recording" corresponds to an entire activity session, from the start of an action to its end. In contrast, an "observation" corresponds to an individual sensor measurement. Notably, a relatively lower number of observations is observed in "inactive" activities relative to the overall recorded time. This discrepancy arises from the sensors' increased frequency during activities involving more movement, a phenomenon explained earlier. In this way, the general data distribution, when measured by total percentage, might misperceive the actual scenario when sliding windows are introduced. That is because, with sliding windows employed for feature computation, the number of observations becomes secondary in importance, with total recorded time taking precedence. The more extensive the recorded time, the greater the number of computed sliding windows and resultant samples for a given class. Consequently, an imbalance becomes clearly evident in the dataset, wherein the "inactive" activity contains thrice as many samples as the "walking" category. As for the number of recordings made, a notable disparity exists, with the "walking" activity featuring significantly more recordings than the others. Regardless, the dataset is deemed valuable and feasible for developing models capable of discerning these activities. Furthermore, the study engaged 19 individuals, contributing a diverse array of behaviours that inherently enrich the potential models that could be crafted in the subsequent stages.

Furthermore, the data acquisition process sought to contain a spectrum of individuals with various attributes, encompassing differences in physical traits, usage routines for their smartphones, and the device models utilised. As a result, the study engaged 19 participants, spanning an age range of roughly 25 to 50 years. This approach was adopted to ensure the inclusion of an extensive array of behavioural patterns that could significantly contribute to developing subsequent models. However, it is worth noting that gender diversity remains limited, with merely two female participants. Nevertheless, the participants exhibit a wide array of physical attributes, habits, and smartphone preferences, with their manner of use and device placement. This scope of variation signifies that while there might be room for enhancement in terms of variability, a noteworthy level of diversity remains inherent within the dataset, as pursued initially.

Nonetheless, an additional difficulty arises within the Android framework, as not all devices are equipped with both a gyroscope and a magnetometer. Although an accelerometer and GPS are requisite, older Android versions do not mandate the inclusion of a gyroscope or magnetometer. Consequently, certain users stored measurements without the involvement of these sensors. Tables 3 and 4 provide an overview of the number of observations lacking a gyroscope or both a gyroscope and magnetometer concurrently. It is worth noting the contrast between the relationship of number of observations to recorded time depicted in these tables compared to the report in Table 2. Specifically, the number of observations is considerably higher compared to the recorded time. This peculiarity might account for the previously noted unusual data in Table 1, wherein the accelerometer's frequency potentially increases more extensively as it becomes the sole sensor detecting motion. In addition, the percentages displayed in these tables are derived from the entire dataset quantity, as outlined in Table 2. Fortunately, these

percentages remain pretty modest, and the dataset's integrity is relatively unharmed by this issue. Nevertheless, it is prudent to bear this in mind while preparing the data for its application in forthcoming AI models.

Table 3: Distribution of the dataset for each measured activity excluding the gyroscope.

Activity	Recorded Time (s)	Number of Recordings	Number of Observations	Percentage of Data
Inactive	11,523	8	668,536	2.29%
Active	13,866	7	619,913	2.13%
Walking	4,169	15	584,262	2.01%
Driving	25,718	23	3,776,468	12.97%
Overall	55,276	53	5,649,179	19.40%

Table 4: Distribution of the dataset for each measured activity without the gyroscope and magnetometer.

Activity	Recorded Time (s)	Number of Recordings	Number of Observations	Percentage of Data
Inactive	5,409	2	269,710	0.93%
Active	10,286	2	90,487	0.31%
Walking	0	0	0	0%
Driving	0	0	0	0%
Overall	25,695	4	360,197	1.24%

Actual data behaviour example

In this section, an example is provided to illustrate the behaviour of the previously collected data in a real-world scenario, wherein the data had undergone prior preprocessing, encompassing the removal of outliers and other observations that could be deemed corrupted.

Accordingly, without delving deeper into such preprocessing steps, Table 5 presents each sensor's mean and standard deviation values for each studied activity. Given that, to accurately comprehend the values in that table, it is worth explaining what each sensor measures. First, the accelerometer values correspond to the acceleration force applied to the smartphone along the three physical axes (x, y, z) in m/s^2. Then, the gyroscope measures the rotation speed of the smartphone around each of the three physical axes (x, y, z) in rad/s. Regarding the magnetometer, it measures the environmental geomagnetic field along the three physical axes (x, y, z) of the smartphone in μT. Concerning the GPS, its values include increments in longitude and latitude coordinates relative to the previous measurement, as well as increments in altitude in meters. Moreover, the values of speed, bearing, and accuracy were also considered. In this way, speed, measured in m/s, represents the smartphone's velocity. As for the bearing, it indicates the horizontal direction of the smartphone's travel in degrees. Finally, the accuracy values indicate the deviation from the actual smartphone location, expressed in meters, with smaller values indicating higher measurement accuracy. Back to Table 5, note that each cell contains the mean values at the top and the corresponding standard deviation values below in smaller font size. Each pair of values corresponds to each sensor set, where the accelerometer, gyroscope, and magnetometer refer to their respective axes (X, Y and Z). For GPS, the set includes latitude increments (*Lat.*), longitude increments (*Long.*), altitude increments (*Alt.*), speed (*Sp.*), bearing (*Bear.*), and accuracy (*Acc.*) measurements. There, it is worth noting some rare data, such as those associated with the GPS "inactive" activity, which exhibit unexpectedly high values. That can be attributed to the fact that such action was often performed indoors, which may limit GPS accessibility. Nevertheless, noticeable differences exist between the activities, indicating the potential for identification with future models.

Table 5: Mean and standard deviation values of sensors for each recorded activity.

		Activity			
		Inactive	**Active**	**Walking**	**Driving**
Accelerometer	X	0.11761 $\pm$0.45934	$-$0.01338 $\pm$1.30277	0.09425 $\pm$3.33422	$-$0.04747 $\pm$0.83290
	Y	0.06136 $\pm$0.26764	0.07598 $\pm$1.45440	$-$0.37604 $\pm$4.35808	$-$0.12936 $\pm$0.93828
	Z	0.84318 $\pm$2.66926	0.13008 $\pm$1.70294	0.07353 $\pm$4.09859	0.18127 $\pm$1.24042
Gyroscope	X	$-$0.00004 $\pm$0.03828	$-$0.00001 $\pm$0.36806	0.00760 $\pm$1.31125	0.00080 $\pm$0.19224
	Y	0.00004 $\pm$0.04719	$-$0.00102 $\pm$0.40959	$-$0.00020 $\pm$0.89244	0.00277 $\pm$0.19835
	Z	0.00001 $\pm$0.03526	0.00055 $\pm$0.24528	$-$0.00560 $\pm$0.53685	$-$0.00243 $\pm$0.16678
Magnetometer	X	25.93805 $\pm$56.45617	6.03153 $\pm$30.00980	$-$0.28182 $\pm$27.03210	$-$5.96356 $\pm$46.08005
	Y	$-$19.62683 $\pm$85.70343	$-$0.02890 $\pm$28.76398	18.73800 $\pm$29.63926	10.73609 $\pm$40.46829
	Z	$-$56.60425 $\pm$33.19593	9.56310 $\pm$39.76136	0.64541 $\pm$25.55331	$-$2.93043 $\pm$29.45994
GPS	**Lat.**	0.00075 $\pm$0.00166	0.00112 $\pm$0.00234	0.00047 $\pm$0.00220	0.00175 $\pm$0.00365
	Long.	0.00125 $\pm$0.00285	0.00118 $\pm$0.00314	0.00056 $\pm$0.00300	0.00204 $\pm$0.00420
	Alt.	32.59169 $\pm$53.06269	30.77538 $\pm$48.65634	34.06931 $\pm$42.51933	41.59391 $\pm$54.74934
	Sp.	0.37222 $\pm$0.82495	0.12109 $\pm$0.81007	0.79924 $\pm$0.71835	10.82191 $\pm$11.82733
	Bear.	57.25005 $\pm$105.49576	14.69719 $\pm$56.00693	124.85103 $\pm$119.80663	118.88108 $\pm$118.78510
	Acc.	265.44485 $\pm$494.66499	214.57640 $\pm$429.81169	75.54539 $\pm$259.59907	192.90736 $\pm$508.87285

In any case, to depict the actual distribution that may arise when processing the data to feed the relevant AI models, an illustrative example is presented in Table 6. In this instance, the data corresponds to the application of a 20-second window with a 19-second overlap, derived from the dataset resulting from the simultaneous utilisation of all sensors. As discernible, a bias towards the "inactive" activity is ultimately observed, as previously discussed, owing to the ease of collecting such data in comparison to the rest. Nonetheless, an adequate number of samples exists for all the studied activities. Consequently, their subsequent classification is viable, necessitating only an awareness of this issue and its ensuing mitigation.

Table 6: Total sample count generated using all sensors with a 20-second sliding window and 19-second overlap.

Activity				
Inactive	**Active**	**Walking**	**Driving**	**Overall**
214,130 (43%)	140,060 (28%)	83,376 (17%)	61,710 (12%)	499,276

Acknowledgements

This work was funded by CITIC is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS), Xunta de Galicia/FEDER-UE (ConectaPeme, GEMA: IN852A 2018/14), MINECO-AEI/FEDER-UE (Flatcity: TIN2016-77158-C4-3-R) and Xunta de Galicia/FEDER-UE (AXUDAS PARA A CONSOLIDACION E ESTRUTURACION DE UNIDADES DE INVESTIGACION COMPETITIVAS.GRC: ED431C 2017/58 and ED431C 2018/49).

Bibliography

J. K. Aggarwal and L. Xia. Human activity recognition from 3d data: A review. *Pattern Recognition Letters*, 48:70–80, 2014.

F. Attal, S. Mohammed, M. Dedabrishvili, F. Chamroukhi, L. Oukhellou, and Y. Amirat. Physical human activity recognition using wearable sensors. *Sensors*, 15(12):31314–31338, 2015.

F. Demrozi, G. Pravadelli, A. Bihorac, and P. Rashidi. Human activity recognition using inertial, physiological and environmental sensors: a comprehensive survey. *arXiv preprint arXiv:2004.08821*, 2020.

D. Garcia-Gonzalez, D. Rivero, E. Fernandez-Blanco, and M. R. Luaces. A public domain dataset for real-life human activity recognition using smartphone sensors. *Sensors*, 20(8):2200, 2020.

M. M. Hassan, M. Z. Uddin, A. Mohamed, and A. Almogren. A robust human activity recognition system using smartphone sensors and deep learning. *Future Generation Computer Systems*, 81:307–313, 2018.

O. D. Lara and M. A. Labrador. A survey on human activity recognition using wearable sensors. *IEEE communications surveys & tutorials*, 15(3):1192–1209, 2012.

R. Solis Castilla, A. Akbari, R. Jafari, and B. J. Mortazavi. Using intelligent personal annotations to improve human activity recognition for movements in natural environments. *IEEE Journal of Biomedical and Health Informatics*, pages 1–1, 2020.

G. Weiss and J. Lockhart. The impact of personalization on smartphone-based activity recognition. *AAAI Publications, Workshops at the Twenty-Sixth AAAI Conference on Artificial Intelligence*, 2012.

A. Zahin, R. Q. Hu, et al. Sensor-based human activity recognition for smart healthcare: A semi-supervised machine learning. In *International Conference on Artificial Intelligence for Communications and Networks*, pages 450–472. Springer, 2019.

N. Zhu, T. Diethe, M. Camplani, L. Tao, A. Burrows, N. Twomey, D. Kaleshi, M. Mirmehdi, P. Flach, and I. Craddock. Bridging e-health and the internet of things: The sphere project. *IEEE Intelligent Systems*, 30(4):39–46, 2015.

GUANIN: GUi-driven Analyzer for NanoString Interactive Normalization

Julián Montoto-Louzao, Alba Camino-Mera, María J Martín, and Antonio Salas

Genetics, Vaccines, Infections and Pediatrics Research Group (GENVIP) - Instituto de Investigación Sanitaria - Hospital Clínico Universitario de Santiago, USC, 15782, Santiago de Compostela, Spain
INCIFOR, Facultade de Medicina - GenPoB - Instituto de Investigación Sanitaria - Hospital Clínico Universitario de Santiago, USC, 15782, Santiago de Compostela, Spain
Grupo de Arquitectura de Computadores, Centro de Investigación CITIC, Universidade da Coruña, 15071 A Coruña, Spain
Correspondence: antonio.salas@usc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: Most tools for NanoString data normalization, aside from the default NanoString nCounter software, are R packages that focus on technical normalization but lack configurable parameters. However, content normalization is the most sensitive, experiment-specific, and relevant step to preprocess NanoString data. Currently, this step requires the use of multiple tools and deep management and understanding of all the normalization process by the researcher. To simplify this crucial step, we have developed GUANIN, a complete normalization tool that offers a wide variety of options to introduce, filter, choose, and evaluate reference genes for content normalization. GUANIN allows, among other features, introducing reference genes from an endogenous subset, a useful approach that addresses the problems associated with the selection of housekeeping genes only. GUANIN allows specific and straightforward normalization approach for each experiment, using a wide variety of parameters with suggested adjustments. GUANIN outperforms other available methods in terms of normalization, especially when comparison groups are defined beforehand, and allows the researcher to comprehensively interact with the preprocessing process without programming knowledge.

1 Introduction

NanoString Geiss et al. (2008) is a molecular barcoding platform for quantification of direct nucleic acid hybridization of RNA in tissue samples. It reports actual counts of sequences of interest through image analysis. As no amplification is needed, it avoids potential bias introduced by reverse transcription, striking a balance between the limitations of RNA-seq and microarrays. Due to its robust performance, NanoString is mainly used in experiments involving low quality samples, and/or tissue samples for the identification of nucleic acid presence, where proper quality control (QC) and normalization are crucial to maintain the accuracy of the experiments Gagnon-Bartsch and Speed (2012).

The NanoString nCounter platform nSolver offers a Graphical User Interface (GUI) addressing background correction, positive control (technical) normalization, and housekeeping normalization. However, nSolver lacks a comprehensive configuration for content normalization or the evaluation of normalization results, among other features.

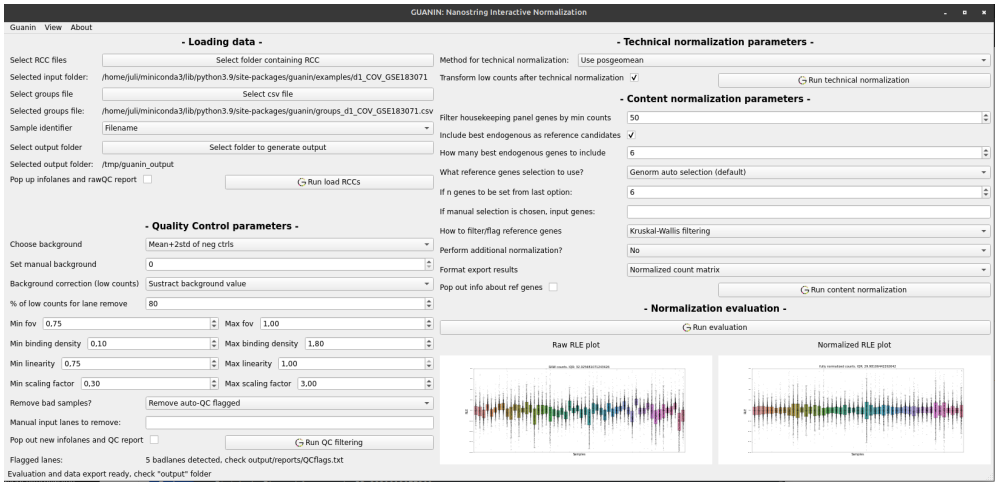


Figure 1: Main screen of GUANIN

Apart from nCounter NanoString native platform, there are several R packages that address the preprocessing of NanoString data, such as NACHO Canouil et al. (2020), NanoStringNorm Waggott et al. (2012), NanoStringDiff Wang et al. (2016), RUV-III Molania et al. (2019) or RCRNorm Jia et al. (2019). Most of them focus their robustness on a single issue, such as technical normalization, visualization, or differential expression. Therefore, a comprehensive normalization analysis would involve the use of several of these packages, which would require in-depth knowledge of the field, including data management and migration from one package to another, and where data handling problems usually arise.

Furthermore, previous approaches mainly focus on complete technical normalization (experiment-wise and sample-wise variability), while content normalization (gene-wise variability) approach is limited, usually offering selection among the housekeeping genes only.

To offer a comprehensive, easy to use, interactive pipeline for NanoString preprocessing we created GUANIN, a flexible and adaptable tool that allows the users to adjust the normalization process to the characteristics of their experiments.

GUANIN is implemented in Python and includes two user interfaces: a Command-Line Interface (CLI) and a very easy-to-use GUI. It is available through the official Python PyPI repository (*https://pypi.org/project/guanin*) and it can be installed with a single command (`pip install guanin`) in Linux, MacOS and Windows systems.

2 GUANIN overview

GUANIN enables users to easily detect and manage difficulties for normalization and QC problems within the experiment through wide flexible input of data, flexible QC, complete technical normalization, improved content normalization with an integration of several new and well-known methods, exhaustive evaluation of reference genes and easy evaluation through in-built plots and pdf reports.

Figure 1 shows the initial screen of the GUI of GUANIN. As can be seen in the figure, expert users can enter a large amount of parameters to fully adapt the normalization process to their experiments, but default values are also provided to help novice users to use the tool. The GUANIN GUI was built using PyQt6 Computing (2023), one of the most popular libraries for the development of Python graphical applications.

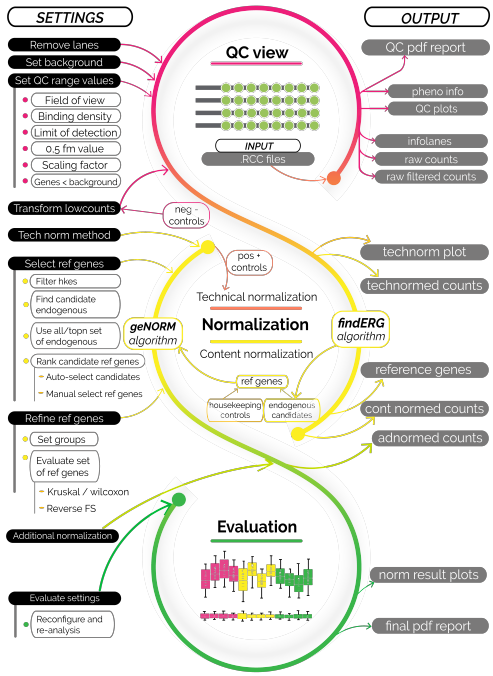


Figure 2: Scheme of GUANIN full workflow

3 GUANIN workflow

Figure 2 shows GUANIN workflow. It starts with load and inspection of input data (RCC files), continues with technical normalization (assessing experimental variations), background correction and content normalization (assessing biological variability). Additionally, it offers the possibility of performing additional normalization, formatting output data and evaluating the normalization process.

3.1 Step 1: Loading RCC files and generating a QC report

Right after loading input files, the first analysis reports the inherent information from the experiment needed to perform adequate normalization through a QC report. From there, QC parameters such as background, lanes to remove, or QC acceptance ranges can be recursively modified until an optimal QC status is achieved, allowing the normalization step to begin.

3.2 Step 2: Normalization (main step)

In contrast to the nSolver pipeline, GUANIN's default workflow performs technical normalization before background correction, as it has shown improved normalization results Lin et al. (2016). Several methods of background calculation and technical normalization are offered.

For content normalization, a set of reference genes needs to be chosen. In addition to default housekeeping genes, we introduce a new approach to select candidate reference genes among the endogenous ones, as it is a common issue that housekeeping genes are not suitable for the experiment. We utilized ERGene Zeng et al. (2020), a Python library for screening endogenous reference genes. The candidate reference genes, including n selected endogenous genes and panel housekeeping genes, are evaluated using a geNORM-based algorithm Vandesompele

	nSolver	NanoStringNorm	NACHO	RCRnorm	RUV-III	Guanin
Platform	GUI	R package	R package	R package	R package	GUI
Visualization	None	None	Shinyapp	None	None	QC pdf report
Background	Manual/auto	Manual/auto	Auto	Auto	Auto	Manual / 4 auto
Tech. norm.	Geomean normfactor	Geomean / sum normfactor	Geomean normfactor and glm	Regression model	Remove unwanted variation algorithm	Geomean/sum/med normfactor
Remove outliers	Manual/flagged	Manual	By parameter	None	Manual	Manual/auto
Housekeeping filtering	Manual	Manual	Manual	None	None	Throrough
Reference genes available	Default / manual housekeeping	Default / manual housekeeping	Default / manual housekeeping	Regression model housekeeping	None	FindERG endogenous + housekeeping
Reference genes selection	Manual	Manual	Manual	Manual	None	geNorm selection + filter and rank
Additional normalization	None	Quantile and z-score	None	None	None	Quantile and standarization
Unwanted variation removed	Low	Low	Medium	Medium	High	High
Evaluation	None	None	None	None	None	Mean IQR and RLE plots
User friendlyness	Easy	Medium	Medium	Hard	Hard	Easy

Figure 3: Comparative on usability amongst main NanoString normalizing tools and GUANIN

et al. (2002); Zhong (2019) to select n genes for driven content normalization. Indeed, the candidate reference genes are filtered or flagged by a three-way group-driven differential expression analysis among groups, employing the Kruskal-Wallis Kruskal and Wallis (1952), Wilcoxon rank sum test Wilcoxon (1945), and a reverse sequential feature-selection method Gelsema and Kanal (2014) that considers the combined effect of several genes. Alternatively, manual selection of reference genes is also available to the researcher. Once content normalization is performed, additional normalization options such as standardization or quantile normalization are available.

3.3 Step 3: Evaluation of normalization results

The normalization results are evaluated through computation of the interquartile range and graphical analysis using Relative Log Expression (RLE) plots Gandolfo and Speed (2018), which compare the raw data with the normalization results.

An example of RLE plots can be seen at the bottom right corner of Figure 1. It corresponds to the processing of the dataset GSE183071 available at the GEO database.

For more detailed information see the GUANIN User Guide at *https://github.com/julimontoto/guanin/blob/master/GUANIN_userguide_1.3.pdf* .

4 Results

The main objective of GUANIN is to provide a flexible and adaptable parametrization to allow the users to adjust the normalization process to the characteristics of their experiment. As experienced users of NanoString data normalization, we have implemented into GUANIN the next features to offer a better user experience and wide experiment compatibility:

- GUANIN can preprocess miRNA and RNAs experiments by default.
- Wide compatibility with different editions of RCC format, column names, and gene identifiers.
- Optional visualization of results for every step.
- Wide range of background choices, including brand new approaches relevant to specific experiments.
- Configurable and visual QC.

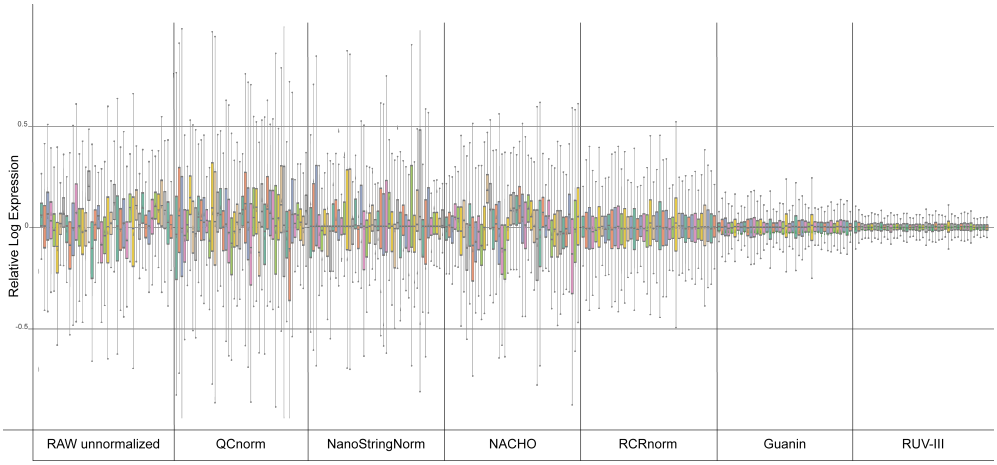


Figure 4: Comparison of RLE plots from normalized output data for the dataset GSE183071 and for several NanoString normalizing solutions

- Possibility to perform technical normalization after or before background correction.
- Adaptable content normalization, which includes algorithms to select, ponderate, and validate any number of endogenous genes.
- GUANIN incorporates a thorough evaluation of candidate reference genes, including the Kruskal-Wallis test, Wilcoxon test, and a machine learning reverse feature selection procedure to asses to assess the combined effect of multiple genes on the condition under study.
- PDF reports with main plots for QC and normalization results.
- In-built evaluation with RLE plots in the main window.

A comparison of usability with respect to other NanoString data normalization tools can be found in Figure 3.

In order to evaluate GUANIN, we have examined three studies, including one in-house dataset of a COVID-19 study (GSE183071) and two published datasets (GSE160208, GSE108395) that can assess several standard casuistic issues when analyzing NanoString data. All of them are available at the GEO database (*https://www.ncbi.nlm.nih.gov/geo/*). The User's Guide of GUANIN (*https://github.com/julimontoto/guanin/blob/master/GUANIN_userguide_1.3.pdf*) details how the normalization process of each of these dataset is performed and the results obtained.

While other software packages may be unable to address specific problems during data handling and normalization, GUANIN maintains good RLE plots and provides accurate normalized data for all the datasets. Figure 4 shows a comparison of the RLE plots for normalized output data for the dataset GSE183071 amongst main NanoString normalization tools. While RUV-III shows better RLE plots that could indicate a better normalization, it is less configurable, flexible to input data, adaptable to preprocessing problems, and user-friendly compared to GUANIN. Besides, it is important to note that centered and narrow RLE plots are not always indicative of a better normalization, as some relevant variability can be lost too. Additionally, GUANIN implements additional features like the geNorm selection algorithm, which is not possible to use with other approaches, making it a comprehensive tool for the normalization process.

5 Conclusions and future work

In a sequential interactive framework, it is necessary to properly address preprocessing step of every specific experiment properly. GUANIN provides the ability to conduct thorough analysis and adapt preprocessing to each experiment without an in-deep knowledge of any programming language, which is a useful advantage for health researchers.

Assuming NanoString pre-built-in housekeeping genes will work ideally as reference genes for an experiment is naive, especially when dealing with diverse tissues and metabolic processes. Therefore, including endogenous genes as candidate reference genes proves to be a great option. The geNorm intelligent evaluation and selection process implemented in GUANIN includes at least three endogenous genes in best 6-gene selection to be used as reference genes.

While no other tool provides as wide interactive parametrization as GUANIN does, NACHO can be useful for a smooth alternative visualization, and RUV-III offers excellent results in removing unwanted variation but it requires to include technical replicates, which are not typically available in most NanoString experimental designs or in public repository datasets. RUVSeq Risso et al. (2014) has been also probed to offer good normalization results on NanoString data Hafemeister and Satija (2019), although it was designed for RNA-Seq analysis.

Additionally to its flexibility, GUANIN generally exhibits better RLE plots than any previous NanoString normalizing tool, especially when groups are given.

Furthermore, GUANIN's results are particularly promising when the analysis can be refined; which is frequently the case on exploratory or confirmatory studies, main objective of NanoString experiments. It can address issues such as poor housekeeping performance, poor negative control, low general expression, and suboptimal experiment design. This is common, as most NanoString panels are preset with default housekeeping genes for a tissue or with a selection of endogenous genes that does not have to match specifically our experiment if panel is not custom (and even if it is custom these problems are usual). Because of this, poor QC is often encountered.

As for future work, although we introduced a regularized linear binomial regression model option for technical normalization, a single scaling factor does not effectively normalize both lowly and highly expressed genes Bhattacharya et al. (2021), thus we are working on improving the method and/or applying it to content normalization, as results are not as good as expected.

Finally, we think that GUANIN's excellent results, combined with its wide flexibility and easy-to-use interface, make it the best preprocessing tool for clinical scientists seeking a fast, reliable, and comprehensive method to preprocess their data and obtain visual reports of the results.

It can be also a useful tool for experienced scientists with programming experience, as it allows for an easy transition from RCCs to evaluated normalized data and provides intermediate data of the processes that can be easily accessed, facilitating the introduction of custom pipelines if desired.

Data Availability

GUANIN is open software distributed under the GPL v3 license. Source code, documentation and case studies are available at https://github.com/julimontoto/guanin.

Acknowledgements

CITIC is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS).

Bibliography

A. Bhattacharya, A. M. Hamilton, H. Furberg, E. Pietzak, M. P. Purdue, M. A. Troester, K. A. Hoadley, and M. I. Love. An approach for normalization and quality control for NanoString RNA expression data. *Briefings in bioinformatics*, 22(3):bbaa163, 2021.

M. Canouil, G. A. Bouland, A. Bonnefond, P. Froguel, L. M. t. Hart, and R. C. Slieker. NACHO: an R package for quality control of NanoString nCounter data. *Bioinformatics*, 36(3):970–971, 2020.

R. Computing. What is pyqt. *https://www.riverbankcomputing.com/software/pyqt/*, 2023.

J. A. Gagnon-Bartsch and T. P. Speed. Using control genes to correct for unwanted variation in microarray data. *Biostatistics*, 13(3):539–552, 2012.

L. C. Gandolfo and T. P. Speed. RLE plots: Visualizing unwanted variation in high dimensional data. *PloS one*, 13(2):e0191629, 2018.

G. K. Geiss, R. E. Bumgarner, B. Birditt, T. Dahl, N. Dowidar, D. L. Dunaway, H. P. Fell, S. Ferree, R. D. George, T. Grogan, et al. Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nature biotechnology*, 26(3):317–325, 2008.

E. S. Gelsema and L. N. Kanal. *Pattern Recognition in Practice IV: multiple paradigms, comparative studies and hybrid systems*. Elsevier, 2014.

C. Hafemeister and R. Satija. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome biology*, 20(1):296, 2019.

G. Jia, X. Wang, Q. Li, W. Lu, X. Tang, I. Wistuba, and Y. Xie. RCRnorm: an integrated system of random-coefficient hierarchical regression models for normalizing NanoString nCounter data. *The annals of applied statistics*, 13(3):1617, 2019.

W. H. Kruskal and W. A. Wallis. Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association*, 47(260):583–621, 1952.

Y. Lin, K. Golovnina, Z.-X. Chen, H. N. Lee, Y. L. S. Negron, H. Sultana, B. Oliver, and S. T. Harbison. Comparison of normalization and differential expression analyses using RNA-Seq data from 726 individual Drosophila melanogaster. *BMC genomics*, 17(28), 2016.

R. Molania, J. A. Gagnon-Bartsch, A. Dobrovic, and T. P. Speed. A new normalization for Nanostring nCounter gene expression data. *Nucleic Acids Research*, 47(12):6073–6083, 2019.

D. Risso, J. Ngai, T. P. Speed, and S. Dudoit. Normalization of RNA-seq data using factor analysis of control genes or samples. *Nature biotechnology*, 32(9):896–902, 2014.

J. Vandesompele, K. De Preter, F. Pattyn, B. Poppe, N. Van Roy, A. De Paepe, and F. Speleman. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome biology*, 3(research0034.1), 2002.

D. Waggott, K. Chu, S. Yin, B. G. Wouters, F.-F. Liu, and P. C. Boutros. NanoStringNorm: an extensible R package for the pre-processing of NanoString mRNA and miRNA data. *Bioinformatics*, 28(11):1546–1548, 2012.

H. Wang, C. Horbinski, H. Wu, Y. Liu, S. Sheng, J. Liu, H. Weiss, A. J. Stromberg, and C. Wang. NanoStringDiff: a novel statistical method for differential expression analysis based on NanoString nCounter data. *Nucleic acids research*, 44(20):e151, 2016.

F. Wilcoxon. Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6):80–83, 1945.

Z. Zeng, Y. Xiong, W. Guo, and H. Du. ERgene: Python library for screening endogenous reference genes. *Scientific Reports*, 10(18557), 2020.

S. Zhong. ctrlGene: assess the stability of candidate housekeeping genes. R Package, version 1.0.1, 2019.

Extended Reality in the Operating Room: Robot-Assisted Orthopedics Surgery with Live and Interactive Streaming for Medical Students

Renato Magalhães, Rita Veloso, Paulo Veloso Gomes, António Marques, and Javier Pereira

LabRP, Center for Rehabilitation Research, School of Health, Polytechnic of Porto, Porto, Portugal
Executive Board Member Centro Hospitalar Universitário Santo António
CITIC Research Center, University of A Coruña, A Coruña, Spain
Academic Clinical Center ICBAS. CHUdSA, Porto, Portugal
Correspondence: rfm@ess.ipp.pt

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: Traditionally, medical education comprises both theoretical learning in classrooms and clinical training in hospitals where students can gain clinical experience. This is mostly done on face-to-face teaching models, focused on the educational philosophy of "see one, do one, teach one", was the standard teaching methodology in medical education. Medical education is transforming thanks to medical schools adopting innovations to new clinicians, such as immersive prepare techniques (extended reality): virtual reality, augmented reality and virtual reality. Immersive learning technologies, such as extended reality, can provide an engaging and interactive platform to generate a stimulating learning environment and with the recent development and increased accessibility of immersive technologies, educators have the potential to make simulation-based training more effective. By using holographic devices, such as Microsoft HoloLens 2®, and 5G wireless communications we intent to explore the innovative experience of a robot-assisted orthopaedic surgery, where the procedures were transmitted live stream to Pregraduate Medical Students using the Microsoft Remote Assist®. In addition, students had the opportunity to interact directly from a classroom to the operating room, asking to the surgeon about the procedures performed during surgery and get involved in the surgery, even remotely. At the end, students completed a questionnaire to evaluate the experience and the preliminary results made possible to assess the effectiveness of this experience and identify areas for improvement for future surgery transmission, revolutionizing the teaching and practice of surgery.

1 Introduction

Medical knowledge is changing rapidly: medical knowledge doubles every 6–8 years, with new medical procedures appearing every day. The rapid change in medical knowledge calls for innovative learning tools for medical practice and education. Immersive technology seems to be one of them (Tang et al., 2022). Traditionally, medical education comprises both theoretical learning in classrooms and clinical training in hospitals where students can gain clinical experience (Van Way, 2017). This is mostly done on face-to-face teaching models, focused on the educational philosophy of "see one, do one, teach one", was the standard teaching methodology in medical education (Viglialoro et al., 2021).

Medical training is a long and demanding process, in which the first stages are usually based on two-dimensional, static, and unrealistic content. Conversely, advances in preoperative imaging have made it an essential part of any successful surgical procedure (Sánchez-Margallo et al.,

2021). E.g., although cadavers constitute the gold standard for teaching anatomy to medical and biomedical science students, previous benefits have also been reported through the use of tablet-based software and new immersive technology (Sánchez-Margallo et al., 2021). Simulation-based training, relying on "see one, simulate many deliberately, do one" principle, has been proposed as an excellent adjunct method to traditional medical education (Stefanidis et al., 2015), (Vozenilek et al., 2004). However, due to COVID-19 pandemic there has been a recent shift to the greater adoption of technology in medical education (Haroon et al., 2020) and medical education has benefitted from the introduction of new technology within recent years such as immersive devices (Jacobs and Rigby, 2022). Medical education is transforming thanks to medical schools adopting innovations to new clinicians, such as immersive prepare techniques: virtual reality, augmented reality and virtual reality (Yasser, 2019).

Mixed Reality (MR) was first mentioned in 1994 by Paul Milgram, and is a blend of physical and digital worlds, unlocking the links between human, computer, and environment interaction, is based on advancements in computer vision, graphical processing power, display technology, and input systems (Microsoft, 2021). By using holographic devices, such as Microsoft HoloLens 2, the participants could take advantage of the ability to place digital content in the real world as if it were there. This technology allows participants to see through display and see the physical environment while wearing the headset and allows a full six-degrees-of-freedom movement, both rotation and translation. The participants can hold "hands-free" and "heads-up" teams video calls with experts anywhere in the world, with all of benefits in this kind of experience (Microsoft, 2021), (Gallagher and Alford, 2020).

Virtual reality (VR), augmented reality (AR), mixed reality (MR), and extended reality (XR) are examples of immersive technologies that have the potential to improve medical practice and education (Tang et al., 2022). In surgical training, most simulation-based approaches have focused on traditional VR and AR technologies, which offer different degrees of immersive experience but are generally unable to interact with 3D information combined with the real-world environment (F.M. et al., 2018), (Sappenfield et al., 2018). MR techniques have replaced these traditional technologies intending to combine the real working environment with virtual content so that users can interact with both simultaneously. MR surgical simulators and medical training applications are becoming an important part of the training process for physicians, as they allow for a training environment appropriate for recreating realistic and reproducible scenarios without putting the patient at risk (Amparore et al., 2022).

Recent research has also shown that with the use of advanced technological solutions such as MR, the spaces of the operating rooms tend to decrease and the number of professionals and students present during a surgery too (Anjali and David, 2018), improving the efficiency in the use of resources in a hospital, whether they are human, of space, or technicians and materials. MR is a concept which provide an "ideal virtual space with [sufficient] reality essential for communication" (Milgram and Kishino, 1994) . The combination of computer processing, human input, and environmental input sets the stage for creating true MR experiences. Movement through the physical world translates to movement in the digital world and improve the experience and better outcomes of the participants and tasks (Flavián et al., 2019), once it did not blind doctors' original view of the real world, but shows a new vision with mutual correction function, which improved the safety of surgery (Hu et al., 2019). This kind of technology is also a powerful for better training and improving education on surgical tactics and methods (Zhu et al., 2014).Medical teaching, due to the limitation of medical environment and ethics, as well as the teaching cost has been increasing, is limited (Hu et al., 2019). MR technology helps students to understand more intuitively the complex anatomy of the human body, which enables students to get more effective training. MR provides new learning models in medical education, which will transform the pedagogy from using two-dimensional images and video to promoting learning through interactive mobile environments (Sappenfield et al., 2018).

Surgical training outside the operating room (OR) using simulation has widely spread this last decade, especially in laparoscopic surgery (Reznick and MacRae, 2006) and training out of the OR has proven its positive effect in basic skills during real laparoscopic procedures in

patients. Such training out of OT may reduce learning curves and improve patients' safety in the OR. Indeed, juniors surgeons have limited access to these complex procedures as primary operator (Miskovic et al., 2010). In conclusion, medical education is transforming thanks to medical schools adopting innovations to new clinicians, such as immersive prepare techniques (extended reality): virtual reality, augmented reality and virtual reality.

2 Methods

To assess the utility of Mixed Reality (MR) in surgical medical education by the using of holographic devices, such as Microsoft HoloLens 2®, and 5G wireless communications a MR set up during an orthopaedic surgical lesson was prepared to explore the innovative experience of a robot-assisted orthopaedic surgery, where the procedures were transmitted live stream to Pre-graduate Medical Students. Instead of regular teaching (Figure 1), in a classroom distant from the real operating room with the share of slideshows and anatomical parts of the body, a surgeon who is also a professor of surgery, guided the students throughout the surgery using mixed reality glasses and the application Microsoft Dymanics 365 Remote Assist® (Figure 2). In addition, students had the opportunity to interact directly from a classroom to the operating room, asking to the surgeon about the procedures performed during surgery and get involved in the surgery, even remotely (Figure 2). At the end, students completed a questionnaire to evaluate the experience and the preliminary results made possible to assess the effectiveness of this experience and identify areas for improvement for future surgery transmission, revolutionizing the teaching and practice of surgery.

Figure 1: Regular Lesson.

Figure 2: Mixed-Reality Lesson.

3 Results

The state of art review of the literature allows to verify that the growth of RM-assisted medical education and surgical practice are driven by advances in hardware and software, but they are yet underexplored. Real-time intra-operative guidance is a key for surgical precision.

Immersive learning technologies can provide an engaging and interactive platform to generate a stimulating learning environment (Barrie et al., 2019) and with the recent development and increased accessibility of immersive technologies, educators have the potential to make simulation-based training more effective. After analyzing the results from the perspectives of

the medical students it is possible to conclude that:
1. The quality of the video allowed a clear and detailed visualization of the surgery;
2. The quality of the sound allowed clear hearing of explanations and instructions during surgery;
3. The transmission of surgery through Hololens 2 and the Remote Assistant application facilitated the understanding of surgical procedures;
4. The ability to ask questions during the surgery live transmission made surgical procedures easier to understand;
5. The real-time interaction with the assistant surgeon through the remote assistant application was effective in answering students' questions;
6. The transmission of surgery using Hololens 2 and the Remote Assistant application promoted a greater interest and involvement of the medical students compared to traditional teaching methods;
7. The experience of watching the surgery through the Hololens 2 and the Remote Assistant application was very satisfactory for the medical students.

4 Conclusion

The emergence of new technologies is bringing new models of medical education. The development of MR systems has proven to be an important tool in supporting medical teaching and training. In fact, MR-assisted surgical education and practice is likely to undertake a greater role in the near future. As MR is now affordable, usable, acceptable, and increasingly well-validated, more future studies focus on efficacy of MR in enhancing surgical education and practice are needed and the MR availability needs to be increased. Moreover, MR expands the capabilities and effectiveness of remote learning, which was normalized during the COVID-19 pandemic, to ensure effective student and patient education. MR-based lessons, or even select modules, provide a unique opportunity to exchange experiences inside and outside the medical community.

Acknowledgements

CITIC is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS).

Bibliography

D. Amparore, A. Pecoraro, E. Checcucci, S. DE Cillis, F. Piramide, G. Volpi, A. Piana, P. Verri, S. Granato, M. Sica, M. Manfredi, C. Fiori, R. Autorino, and F. Porpiglia. 3d imaging technologies in minimally invasive kidney and prostate cancer surgery: which is the urologists' perception? *Minerva urology and nephrology*, 74(2):178–185, 2022.

J. Anjali and A. David. Designing a safer or, 2018. URL *https://healthcaredesignmagazine.com/trends/research-theory/designing-a-safer-or*.

M. Barrie, J. Socha, L. Mansour, and E. S. Patterson. Mixed Reality in Medical Education: A Narrative Literature Review. *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care*, 8(1), 2019.

C. Flavián, S. Ibáñez-Sánchez, and C. Orús. The impact of virtual, augmented and mixed reality technologies on the customer experience. *Journal of Business Research*, 100:547–560, 2019.

S.-M. F.M., S.-M. J.A., C. A., R. A., and S. M. Application of mixed reality technology for surgical training in urology. *Surgical Endoscopy*, 32, 2018.

L. Gallagher and J. Alford. Mixed-reality headsets in hospitals help protect doctors and reduce need for ppe, 2020. URL *https://www.imperial.ac.uk/news/197617/mixed-reality-headsets-hospitals-help-protect-doctors/*.

Z. Haroon, A. A. Azad, M. Sharif, A. Aslam, K. Arshad, and S. Rafiq. Covid-19 era: Challenges and solutions in dental education. *Journal of the College of Physicians and Surgeons Pakistan*, 30 (2):S129–S131, 2020.

H. Hu, X. Feng, Z. Shao, M. Xie, S. Xu, X. Wu, and Z. Ye. Application and prospect of mixed reality technology in medical field. *Current Medical Science*, 39(1):1–6, 2019.

C. Jacobs and J. M. Rigby. Developing measures of immersion and motivation for learning technologies in healthcare simulation: a pilot study. *Journal of Advances in Medical Education and Professionalism*, 10(3), 2022.

Microsoft. What is Mixed Reality?, 2021. URL *https://docs.microsoft.com/en-us/windows/mixed-reality/discover/mixed-reality*.

P. Milgram and F. Kishino. Taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems*, E77-D(12), 1994.

D. Miskovic, S. M. Wyles, M. Ni, A. W. Darzi, and G. B. Hanna. Systematic review on mentoring and simulation in laparoscopic colorectal surgery, 2010.

R. K. Reznick and H. MacRae. Teaching surgical skills – changes in the wind. *New England Journal of Medicine*, 355(25), 2006.

J. Sánchez-Margallo, C. Plaza de Miguel, R. A. Fernández Anzules, and F. Sánchez-Margallo. Application of mixed reality in medical training and surgical planning focused on minimally invasive surgery. *Frontiers in Virtual Reality*, 2, 2021.

J. W. Sappenfield, W. B. Smith, L. A. Cooper, D. Lizdas, D. B. Gonsalves, N. Gravenstein, S. Lampotang, and A. R. Robinson. Visualization improves supraclavicular access to the subclavian vein in a mixed reality simulator. *Anesthesia and Analgesia*, 127(1), 2018.

D. Stefanidis, N. Sevdalis, J. Paige, B. Zevin, R. Aggarwal, T. Grantcharov, and D. B. Jones. Simulation in surgery: What's needed next?, 2015.

Y. M. Tang, K. Y. Chau, A. P. K. Kwok, T. Zhu, and X. Ma. A systematic review of immersive technology applications for medical practice and education - trends, application areas, recipients, teaching contents, evaluation methods, and performance. *Educational Research Review*, 35, 2022.

C. W. Van Way. Mixed Reality in Medical Education: A Narrative Literature Review. *Missouri medicine*, 114(6), 2017.

R. M. Viglialoro, S. Condino, G. Turini, M. Carbone, V. Ferrari, and M. Gesi. Augmented reality, mixed reality, and hybrid approach in healthcare simulation: A systematic review. *Applied Sciences*, 11(5), 2021.

J. Vozenilek, J. S. Huff, M. Reznek, and J. A. Gordon. See one, do one, teach one: Advanced technology in medical education. *Academic Emergency Medicine*, 11(11), 2004.

E. M. Yasser. Virtual reality and augmented reality bt. In *Rheumatology Teaching: The Art and Science of Medical Education*, pages 403–427. Springer International Publishing, 2019. ISBN 978-3-319-98213-7.

E. Zhu, A. Hadadgar, I. Masiello, and N. Zary. Augmented reality in healthcare education: An integrative review. *PeerJ*, 2014.

Applicability of Mixed Reality in Operating Room

Renato Magalhães, Rita Veloso, Adélio Vilaça, Ana Oliveira, and David Terroso

LabRP, Center for Rehabilitation Research, School of Health, Polytechnic of Porto, Porto, Portugal
Executive Board Member, Centro Hospitalar Universitário Santo António, Porto, Portugal
Orthopedic Department, Centro Hospitalar Universitário Santo António, Porto, Portugal
Porto School of Engineering, Polytechnic of Porto, Portugal
Academic Clinical Center ICBAS { CHUdSA, Porto, Portugal
Correspondence: rfm@ess.ipp.pt

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: Mixed Reality (MR) technology has emerged as a promising tool for enhancing surgical procedures and training in the operating room (OR). The applicability of MR in the OR has potential benefits and challenges. In this study, we conducted a systematic review to comprehensively evaluate the use of MR in the OR, specifically focusing on training, remote support, and surgical planning. Additionally, we analysed available tools, assessing their usability and integration into existing workflows. The analysis highlights MR solutions that improve surgical precision, visualization, decision-making, and educational experiences. The conclusions derived from the systematic review are presented, emphasizing the applicability of MR in OR and its potential contributions to training and surgical planning.

1 Introduction

Over the years, surgical practise (and consequently how it is taught) and the intraoperative space have changed as technology has evolved following the great advances in the field, and seeking safer, faster, and easier methods to succeed in the OR (Roberts et al., 2006). Recently, with the increasing development of computers and related technologies, such as smartphones and tablets, the drive to use them in the OR has followed the trend to take advantage of the main benefits that these devices bring: easily accessible information (Tahamtan et al., 2017).

MR was first mentioned in 1994 by Paul Milgram, and is a blend of physical and digital worlds, unlocking the links between human, computer, and environment interaction (Milgram and Kishino, 1994) and allows clinicians to do something that was previously unthinkable: be in two places at once, without ever leaving one location.

Extended reality can be briefly defined as the use of technology that expands what we feel in the real world with cues from a virtual environment. These cues can be visual, auditory and, in some cases, even haptic or olfactory (Alizadehsalehi et al., 2020).

Traditionally, doctors and students develop theoretical knowledge by participating in real-life clinical scenarios and classroom simulations using mannequins. However, these forms of training have some limitations. The OR could use this kind of technology as a teaching or support tool to give the surgeon access to information, in a completely manipulative way, that

would allow him to consult annotations of the surgical technique, contact with other specialists via online, and have access to the patient's information, all within his point of view and without the contact with the device, using gestures, voice commands or even the retina movement. This would allow a better preparation and a flattening of the learning curve for the youngest surgeons, through an immersive experience, filling the void of information access and contact with other specialists with ease, that exists in the modern OR.

2 Methods

To support the development of our work, a systematic review was initially carried out on the use of MR in the OR.

The review goals were defined following the PICO methodology. PICO stands for patient/population, intervention, comparison and outcomes. The PICO methodology is a technique used in evidence-based practice to contextualize and answer clinical questions in terms of specific patient problems that are clinically relevant to evidence in the literature (Roever, 2018).

The first topic we wanted to be answered was in which areas MR applications could be used and be helpful in the OR. The second was what solutions are already implemented and the third was what technical challenges and limitations the MR users face. The fourth topic was what are the effects of MR and its costs and the fifth and final question was what effectiveness was demonstrated to have been achieved.

After this work, applications were developed with the aim of demonstrating the use of the tools in the context of training and surgical planning.

2.1 Transmission to Microsoft HoloLens 2® x-rays obtained during surgery, in real time, into the field of vision

The developed application places the image present in the x-ray monitor, connected to the c-arm in the operating room, but also supports the professional with help exterior to the referred room, with theoretical support and communication via wi-fi, recording or livestreaming the surgeon's point of view, creating more interactive and ludic content for the medical students. To show the x-ray images present on the monitors tower in the surgeon's point of view (Figure 1), it was used a DVI to HDMI converter, where the DVI was connected to the monitors tower or the C- arm, and the HDMI cable would be connected to the video converter from HDMI to USB-C that connected to the computer (Figure 2). The computer would then have a Microsoft Teams® account set up, different from the one logged in the HoloLens 2®, where the video camera exposed on the call was setup to the USB-C port. This will live stream the screen where the DVI cable is connected to the call and show the x-rays from the c-arm monitor on the HoloLens 2®, in real time.

Since Microsoft Teams® was already being used, another functionality was added: implement communication. Since it uses wi-fi connection, the user could communicate with other professionals and technical support, while using the HoloLens 2®. This ability was also used for livestreaming the procedure: allowing two-ways communication between the surgeon and the students, who may want to ask some questions, and the visualization of the professional's point of view. In case the livestreaming function is not possible due to any reason, the HoloLens´ recording functionalities would be used. Although it is not as interactive as the livestreaming, it still shows the professional's point of view.

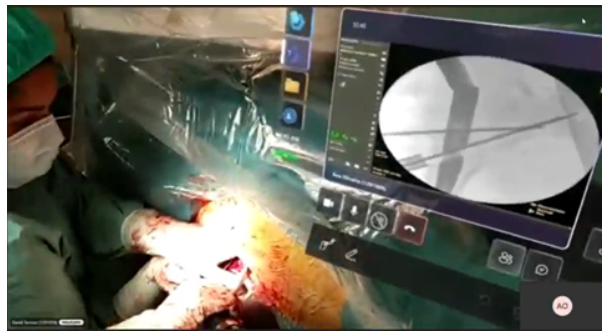


Figure 1: Example of the use of HoloLens to visualize x-rays, intraoperatively.

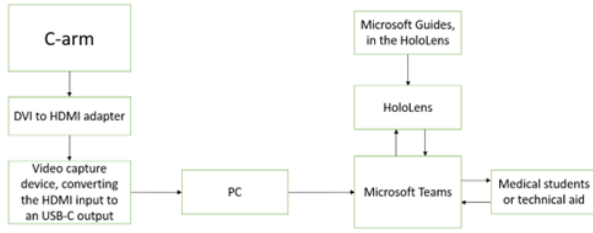


Figure 2: - Relationship between the components of this project.

2.2 Microsoft Dynamics Guides® and HoloLens 2® to create new teaching and training tools for medical education

Microsoft Dynamics 365 Guides® was used to create new teaching and training tools for medical education.

It was created a MR tutorial on Dynamics 365 Guides® about the origin, insertion, nerve supply and actions of the obturator muscles. The 3D model of the pelvis and the muscles was manually segmented from a real patient DICOM image using 3D Slicer next, the image was optimized using Blender and upload to Dynamics 365 Guides®.

The operator starts by scanning the QR code (Figure 3), which functions as an anchor, previously placed by the author in the physical skeleton model. Instructions of this teaching session content showed in the HoloLens 2® in front of the operator whilst an embedded image of the step is shown to the left of the instructions and 3D model of the muscles is displayed (Figure 4). The operator can go to the next step by gazing at the arrow button. The important muscles in the 3D hologram are highlighted in blue.

During the session, students can also interact with healthcare professionals, professors, and colleagues around the world anytime.

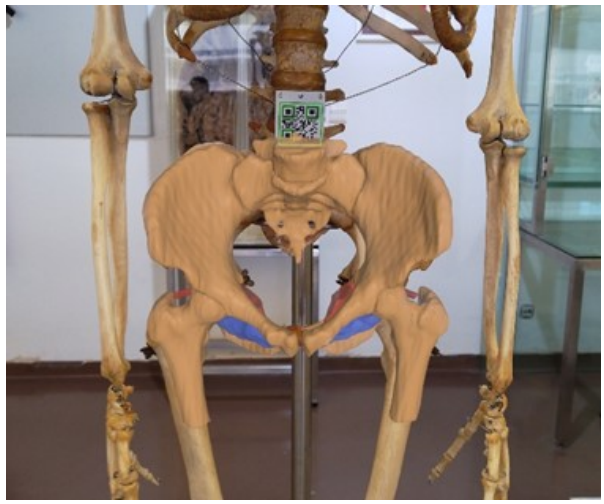


Figure 3: Example of QR Code and position

Figure 4: Example Microsoft Dynamics 365 Guides® preview

3 Results

3.1 Transmission to Microsoft HoloLens 2® x-rays obtained during surgery, in real time, into the field of vision

When implementing the described application, the aim is to diminish the strain and the risk of injury resulting from the unnatural movement present in the OR when the surgeon wants to visualize the image shown in the x-rays monitor. This action can occur multiple times during the intervention, which could lead to errors related to the performance, since redirecting the action may cause loss of concentration in the procedure and parallax errors.

This implementation also seeks flatten the learning curve of the surgical procedure, supporting the learning medics with theoretical support during surgery, allowing the HoloLens 2® user to visualize, while conducting the surgery, theoretical contents, or annotations regarding the procedure. While conducting the first procedures, the surgeons have support of a more experienced professional who needs to be there. However, with this implementation the sup-

port does not need to be present, extending the possibilities, but also diminishing the number of people present in the room.

In the field of education, this use case pursuits to create contents of better quality for the medical students. While being present in the OR may be a physical experience to the students, it is not very clear to watch the surgery. Occupying a lot of space in the room, the students can barely see what the surgeon is doing. By recording or livestreaming the procedure from the surgeon´s point of view, it is expected to supply the students with a much more perceptive experience, understanding what is, in fact, happening. The streaming functionality is more interesting than recording, allowing two-ways communication between the students and the surgeon, giving a better insight into the procedure.

The communication between the team members, the team's level of experience, equipment failures, the interaction between the team, the machines and the equipment, and team stress management were the main issues appoint to surgery scenario. Technology can help solve these problems according to the experts that have also agreed that the maintenance of the equipment could be beneficial, as well as easier access to patient information.

3.2 Microsoft Dynamics Guides® and HoloLens 2® to create new teaching and training tools for medical education

The interactive tutorial had good feedback from the professors and doctors that tested it.

A key limitation of this study was provided by the technology itself which remains in relative incipiency. The velocity of the device and other technical aspects of its operation were principal concerns. In a future investigation, it would be interesting to use the tracking function instead of anchors so that the holograms can follow the physical skeleton.

4 Conclusion

Holographic technologies in healthcare emerge as promising solutions and may be alternatives to traditional solutions.

The developments presented received very positive feedback from professionals who tried them. Although more tests must be done in this field, to evaluate the advantages and disadvantages that may arise.

The surgeons who tested it were satisfied and enthusiastic with the application developed for viewing x-rays, as well as the interactive use of the holographic image and the script with Microsoft 365 Guides®.

As future work, an integrated application will be designed in a specific area with options to better satisfy the needs of the surgeon and medical student without compromising surgical performance. After its use in a simulation environment, a tool usability questionnaire will be presented to these professionals to evaluate the impacts of using these tools on training and surgical planning.

Bibliography

S. Alizadehsalehi, A. Hadavi, and J. C. Huang. From BIM to extended reality in AEC industry. *Automation in Construction*, 116, 2020.

P. Milgram and F. Kishino. Taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems*, 1994.

K. E. Roberts, R. L. Bell, and A. J. Duffy. Evolution of surgical skills training. *World Journal of Gastroenterology*, 12(20), 2006.

L. Roever. PICO: Model for clinical questions. *Evidence-Based Medicine*, 3(2), 2018.

I. Tahamtan, S. Pajouhanfar, S. Sedghi, M. Azad, and M. Roudbari. Factors affecting smartphone adoption for accessing information in medical settings. *Health Information and Libraries Journal*, 34(2), 2017.

Self-adaptive Cooperation Scheme in a Parallel ACO Algorithm for Binary Combinatorial Problems

Roberto Prado-Rodríguez, Patricia González, Julio R. Banga, and Ramón Doallo

Computer Arquitecture Group, Faculty of Computer Science, Universidade da Coruña, 15071 A Coruña, Spain
Computer Arquitecture Group, Faculty of Computer Science, Universidade da Coruña, 15071 A Coruña, Spain
Computational Biology Lab, MBG-CSIC, Spanish National Research Council, Pontevedra, Spain
Computer Arquitecture Group, Faculty of Computer Science, Universidade da Coruña, 15071 A Coruña, Spain
Correspondence: roberto.prado@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: The ant colony optimization (ACO) is widely used for combinatorial optimization problems, although it can suffer from fast convergence to local minima. In order to provide a versatile implementation of ACO, we present a parallel multicolony strategy with an improved cooperation scheme for binary combinatorial problems. Our proposal is based on a self-adaptive method, which assigns appropriate run-time cooperation levels to each problem based on its size and available computational resources. We evaluate this proposal with problems with different levels of cooperation and number of processes. All these configurations combined show its flexibility as a versatile solver for this type of problems.

1 Introduction

One of the most popular metaheuristic used for general combinatorial optimization problem is Ant Colony Optimization (ACO) (Dorigo and Stützle, 2019). It is inspired by the social behavior of ant colonies, specifically in the deposition of pheromones along the explored paths during the search for food sources. ACO has been found to be robust and easily tailored to a wide range of optimization problems, and it has been applied to a number of binary combinatorial instances (Jang et al., 2011; Kashef and Nezamabadi-pour, 2015; Kong and Tian, 2005).

The basic ACO is a general-purpose algorithm, easy to understand and implement. ACO achieves good results in unimodal problems, that is, those defined by the fact that all solutions are guided towards the same optimal result without local minima. However, when tackling problems in which local minima abound, its convergence quickly suffers, easily stagnating in one of the local solutions.

In this work we explore an extension of a parallel multicolony ACO implementation to handle challenging binary combinatorial problems, that incorporates a self-adaptive mechanism for colony cooperation.

The structure of the paper is as follows. Section 2 describes the ACO algorithm adapted to handle binary combinatorial problems and its cooperative parallel scheme proposed is ex-

plained. In section 3, we present the experiments carried out and discuss the results. Finally, in Section 4 we summarize the conclusions of this work.

2 Parallel ACO for binary optimization problems

ACO is often used for problems that can be reduced to finding routes in graphs, such as the Traveling Salesman Problem (TSP) (Stützle et al., 1999). This problem is based on discovering the best route for a traveller who has to visit many cities. In a classical TSP problem, the objective is to visit all the cities and return to the origin covering the shortest possible distance. When it comes to binary combinatorial problems, this can be reduced to finding the optimal path that goes from one node to another by choosing between two possible paths: 0 or 1. Pheromones will be deposited on paths 0 or 1 in each of the N steps. Ants in the subsequent iterations will be influenced by the previously deposited pheromones.

As mentioned before, ACO offers good results for solving unimodal problems. However, when it comes to solving problems with many local minima, ACO tends to get stuck easily. To avoid premature convergence to local minima and, thus, the stagnation of metaheuristics, previous studies indicate the need of increasing the diversity in the search. A good way to achieve this is the use of parallel strategies. The following subsections describe the solution used in this work and the enhancement introduced in the cooperation strategy.

Different parallel strategies can be applied to metaheuristics in general (Alba, 2005), and ACO in particular (González et al., 2022). Most of them can be classified into fine-grained and coarse-grained strategies. A fined-grained parallelization attempts to find parallelism in the sequential algorithm.

A different solution is a coarse-grained approach, which involves looking for a parallel variant of the sequential algorithm. The most popular coarse-grained solution consists of implementing an island-based model. In these models, different distributed colonies exist where the original algorithm is executed in isolation and, from time to time, these colonies exchange information that allow them update their results with the information received from the rest. This parallel implementation is usually known as multicolony model.

Multicolony approaches aim to take advantage of distributed resources to extend the search for solutions. The most trivial multicolony solution consists of a parallel search on multiple non-cooperating colonies. Although this solution was found to yield good results, results usually stand out for approaches that include colony cooperation.

In the cooperative scheme proposed in (González et al., 2022), when a promising new solution arrives at a colony and improves the *best-solution-so-far*, the latter is always replaced by the former. Therefore, a colony that receives a better solution is diverted from its own search. All colonies converge to the same local solution, reinforcing the same path, and eventually getting stuck at the same local minimum.

In this work an efficient selective cooperative scheme is explored. When a colony obtains a promising solution, this solution is spread to the rest and all processes receive the promising solutions. However, to avoid the problem mentioned, only a few processes introduce these solutions into their colony, modifying the pheromone matrix. To determine if a solution that has just arrived in a process should be included in the colony, two aspects are taken into account.

First, although all the colonies cooperate by spreading their promising solutions, some colonies keep their execution outside the influence of the rest, for which they never use the solutions received from outside the colony. This ensures the desired diversity in the ACO progression. The number of colonies that remain independent depends on an integer parameter called $cfreq$.

Second, to further avoid the danger of premature convergence due to early cooperation, the processes will only use the solutions received from other colonies once they have been stalled for a certain number of iterations. This number of iterations are defined by the *cstall* parameter.

As a general rule, it is very complicated to know the optimal level of cooperation for a prob-

lem before dealing with it. To address this situation, a self-adaptive approach to automatically determine the tier of cooperation is also proposed in this paper. To do this, the problem size and the number of resources to be used are taken as a basis, and the parameters are tuned at runtime. Note that the hardness of the problem, which also influences the level of cooperation, is impossible to determine beforehand. The cooperation-index is the ratio between the size of the problem and the number of processes to be used. The higher this index is, the more intensive the cooperation between the colonies should be. Ranges of this index are established to set an upper limit to the *cstall* parameter. The minimum *cstall* is always 0 (full cooperation). At the start of the execution, *cstall* takes its maximum value. The cooperation at the beginning of the execution is scarce. However, each time the algorithm gets stuck and a restart is triggered, *cstall* is reduced by 10% of restart-iterations size, so the algorithm increases their rely on incoming solutions after being reinitialized. If reboots continue to occur, the *cstall* will continue to drop and, thus, the algorithm increases the cooperation between colonies. When the minimum *cstall* is reached (0 iterations), the algorithm returns to the maximum value. In this way, even in multimodal problems, different cooperation degrees are explored in a round-robin fashion.

3 Experimental results

In this section, a series of experiments are shown to assess the value of the strategies proposed in this work.

All the benchmarks used to carry out the experiments reported in this paper are obtained from the W-Model (Weise et al., 2020). The W-Model is a tunable black-box discrete optimization benchmarking problem (BB-DOB) that uses a bit-string representation of the data. The W-Model framework creates different benchmarks by means of different input parameters that modulate different features for the problems. Six challenging benchmarks labeled as B1 to B6 have been defined here. The W-model parameters used to define these challenging problems can be seen in Table 1.

Table 1: W-Model parameters for benchmarks B1 to B6.

Benchmark	Problem size	Neutrality	Epistasis	Ruggedness/Deceptiveness
B1	640	High	130 (81%)	10000 (78%)
B2	720	High	150 (83%)	12000 (75%)
B3	1000	High	200 (80%)	25000 (80%)
B4	640	High	130 (81%)	0
B5	720	High	150 (83%)	0
B6	1000	High	200 (80%)	0

All the experiments were performed at the Galicia Supercomputing Center (CESGA) using the FinisTerrae-III supercomputer. Each FinisTerrae-III node is composed of two Intel Xeon Ice Lake 8352Y CPUs running at 2.2 GHz, with 32 cores per processor (64 cores per node), and 256 GB of RAM. The nodes are connected using an Mellanox InfiniBand HDR 100 Gbps interconnect using a fat-tree topology.

Table 2 lists the different cooperation configurations that have been compared in this work, being n the problem size. Note that M1 configuration corresponds to a non-cooperative parallel solution, while M6 corresponds to the cooperation scheme proposed in (González et al., 2022), that is, a full cooperation between colonies. Besides, M7 configuration corresponds to the self-adaptive solution proposed in this paper.

Table 3 shows the preliminary results of the average execution time for each experiment. Some conclusions can be extracted based on the results. A medium-term cooperation like the M4 configuration overcomes the other configurations for 4 colonies in benchmark B1. In a harder problem like B2, there is a need for increasing the cooperation, and results obtained

Table 2: Configurations with different cooperation degrees.

Configuration	$cfreq$	$cstall$	Comments
M1	∞	-	No cooperation at all.
M2	2	$n/25$	1 of 2 receive after stagnation $\geqslant$ problem size/25.
M3	1	$n/10$	All receive after stagnation $\geqslant$ problem size/10.
M4	1	$n/25$	All receive after stagnation $\geqslant$ problem size/25.
M5	1	$n/50$	All receive after stagnation $\geqslant$ problem size/50.
M6	1	0	Full cooperation, all receive all the time.
M7	1	$self-adapted$	Self-adapted at runtime.

for M5 configuration are better than the others when 4 colonies are used. This behavior is restated in B3 problem, bigger and even more harder than B2. The degree of cooperation that achieves the best performance with the same processes is M6. In brief, the larger and harder to solve the problem, the larger the need for cooperation between the colonies. However, these three problems need less cooperation when you increase the processes to 12. In this case, the optimal level decreases from M4 to M2, from M5 to M4 and from M5 to M4, respectively. In other words, the more processes, the lower the level of cooperation required.

Table 3: Comparison of different cooperative configurations in BiPCACO. Average time (in seconds) achieved on experiments. Highlighted in green and red the best and worst times by configuration, respectively.

		Average time (s)						
Benchmark	**#PROC**	**M1**	**M2**	**M3**	**M4**	**M5**	**M6**	**M7**
B1	**4**	322,5	282,4	328,8	**271,1**	348,2	**423,5**	384,9
	12	109,2	**97,3**	102,4	118,3	183,5	**322,7**	135,3
	64	36,4	39,5	39,6	34,9	79,5	**228,9**	**34,1**
B2	**4**	852	719	858,7	690,6	**630,9**	**865,5**	728,6
	12	328,6	230,1	269,8	**212,1**	330	**562,1**	290,5
	64	68,8	55,9	64,4	**53,2**	92,7	**376,6**	64
B3	**4**	**666,6**	662,2	669,3	572,4	322,3	**223,7**	236,5
	12	206,2	**218,8**	206,5	167,4	**108,5**	141,7	152,8
	64	62,2	59,4	62	61,9	**44,6**	**109,4**	63,6
B4	**4**	**47,1**	44	46,1	45	**39,6**	46,2	46
	12	34,3	32,8	33	32,2	**31,2**	**37,4**	33,3
	64	31,4	31,2	31,1	30,2	**26,9**	32,7	**37,9**
B5	**4**	65,4	64,5	**70,3**	62,8	**42**	61	56,6
	12	37,8	37	39,1	37,4	**34,1**	**47,2**	36,8
	64	33,9	31,1	33,9	33,5	**30,3**	34	**34,5**
B6	**4**	214	207,7	**350,6**	219,7	134,9	**77,9**	98,5
	12	82,4	78,9	**150,8**	71,8	**57,1**	60,7	73,5
	64	44,5	44,5	52,7	45,1	**42**	**56,2**	45

These results also prove that an intense cooperation like M6 configuration is effective when using few colonies to solve very complicated problems. In those benchmarks, the possibilities of finding a great solution are low, so the need to share potentially successful solutions as soon as possible is high in order to speed up the progress of the search. However, when the number of processes increases, the chances for one of them to find a good solution on its own increases, and cooperation can interfere with this search and damage diversity. If a process frequently accepts foreign solutions, the colony deviates from their own search and ends converging to the same local minima as the rest of the colonies.

These custom-defined benchmarks are not only big in size, but also three of them (B1 to B3) show characteristics that make them very hard to solve. They are *multimodal* and are defined by owing many local minima. In those cases, the colonies cooperation benefits the convergence ratio of the algorithm. Problems B4 to B6 are unimodal, i.e. the search is oriented smoothly to

the global minimum. Since the algorithm does not get stuck due to the absence of local minima, it does not need cooperation, and therefore in problems B4 and B5 there is almost no significant difference in times among all methods. In practice, all behave in a non-cooperative fashion. Problem B6, even unimodal, is too big in size and cooperation becomes highly necessary with few processes.

Based on previous experiments, it can be concluded that there are essentially three features of the problem at hand that will determine the degree of cooperation that benefits BiPCACO execution: (1) amount of processes: the smaller the number of processes, the higher the cooperation between colonies must be; (2) multimodality: The higher the bias towards the multimodal landscape, the larger the probability that the algorithm will get stuck; and (3) problem size: the larger the problem, the greater the need for cooperation.

The problem of the previous conclusions is the difficulty for the user to know the features of the problem in advance, and therefore the difficulty of adjusting the configuration parameters before the execution. It is at this point where the self-tuned approach is especially appealing. Moreover, results of Table 3 evidence that it is also competitive when compared with the solution obtained with the best configuration in each problem.

To better illustrate the behavior of the proposed M7 self-tuned cooperation method, logarithmic scale plots for benchmarks B1 and B3 are shown in Figure 1. Those subfigures show the cumulative probability of reaching the optimum related to the execution time, for the 100 runs of each experiment. We choose benchmarks B1 and B3 to show how, despite being both large and difficult problems, in one case cooperation is more beneficial than in the other, however, this is somewhat difficult to know beforehand.

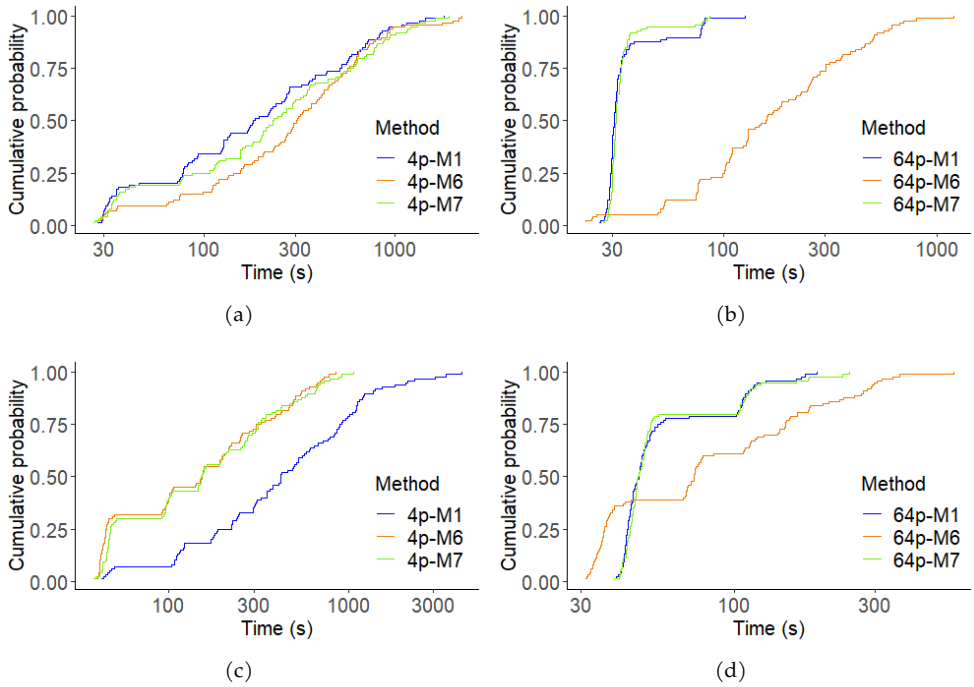


Figure 1: Cumulative probability of reaching the optimum for (a) benchmark B1 using 4 processes, (b) benchmark B1 using 64 processes, (c) benchmark B3 using 4 processes and (d) benchmark B3 using 64 processes, where M1-7 indicates different cooperation schemes.

As it can be seen, for benchmark B1, a configuration with no cooperation (M1) is better than

an intensive cooperation (M6). However, a self-tuned cooperation (M7) adapts during runtime and offers a competitive result versus the non-cooperation solution.

On the other hand, for benchmark B3, the best solution turns out to be a configuration with an intensive cooperation (M6) compared to lack of cooperation (M1) for 4 processes. And, on the contrary, the absence of cooperation (M1) is better compared to an intense cooperation (M6) for 64 processes. Besides, the self-tuned solution (M7) is always competitive when compared with the best solution in each situation.

These previous results show that the self-tuned solution, although it may not improve the superior configuration, allows the user to get rid of the responsibility of choosing the most appropriate parameters, making the algorithm reconfigure itself, at execution time, depending on the progress of the search. Results of the self-tuned solution stands out as a competitive alternative.

4 Conclusions

An improved multicolony ACO for binary combinatorial problems is explored in this paper. The goal is to compensate for the main drawback of the algorithm: its trend to get stuck in local minima. Here, a parallel cooperative ACO strategy is evaluated, in which all colonies share promising routes with each other, but only use them for their own search if certain conditions are met. That is, they collaborate only when necessary. This manages to maintain diversity while avoiding rapid convergence to the same local minimum of all colonies. An implementation that adjust the parameters of the cooperative algorithm at runtime has been also evaluated. This improvement allows us to find a good solution for each problem without the need to know in advance the characteristics of the problem in consideration.

This new self-adapted cooperative approach prevents the user from knowing in advance which configuration is the most appropriate for the problem in consideration and saves them the time needed to set the parameters of the new algorithm.

As future work, we consider to improve and refine the cooperative self-adaptation, as well as to explore the hybridization with other metaheuristics.

Acknowledgments

RPR, PGG and RDB acknowledges funding from Grants PID2019-104184RB-I00 and PID2022-136435NB-I00, funded by MCIN/AEI/ 10.13039/501100011033, PID2022 also funded by "ERDF A way of making Europe", EU; Xunta de Galicia and FEDER funds of the EU (Centro de Investigación de Galicia accreditation 2019–2022, ref. ED431G 2019/01; Consolidation Program of Competitive Reference Groups, ref. ED431C 2021/30).

JRB acknowledges funding from the Ministry of Science and Innovation of Spain MCIN / AEI / 10.13039/501100011033 through grant PID2020-117271RB- C22 (BIODYNAMICS).

Authors also acknowledge the Galician Supercomputing Center (CESGA) for the access to its facilities.

Bibliography

E. Alba. *Parallel metaheuristics: a new class of algorithms*, chapter 5–14, pages 105–346. John Wiley & Sons, 2005.

M. Dorigo and T. Stützle. Ant colony optimization: overview and recent advances. *Handbook of metaheuristics*, pages 311–351, 2019.

P. González, R. R. Osorio, X. C. Pardo, J. R. Banga, and R. Doallo. An efficient ant colony optimization framework for HPC environments. *Applied Soft Computing*, 114:108058, 2022.

S.-H. Jang, J.-H. Roh, W. Kim, T. Sherpa, J.-H. Kim, and J.-B. Park. A novel binary ant colony optimization: Application to the unit commitment problem of power systems. *Journal of Electrical Engineering and Technology*, 6(2):174–181, 2011.

S. Kashef and H. Nezamabadi-pour. An advanced aco algorithm for feature subset selection. *Neurocomputing*, 147:271–279, 2015.

M. Kong and P. Tian. A binary ant colony optimization for the unconstrained function optimization problem. In *International Conference on Computational and Information Science*, pages 682–687. Springer, 2005.

T. Stützle, M. Dorigo, et al. Aco algorithms for the traveling salesman problem. *Evolutionary algorithms in engineering and computer science*, 4:163–183, 1999.

T. Weise, Y. Chen, X. Li, and Z. Wu. Selecting a diverse set of benchmark instances from a tunable model problem for black-box discrete optimization algorithms. *Applied Soft Computing*, 92: 106269, 04 2020.

Self-management Web Application for Persons with Schizophrenia

José Domingues, Paulo Veloso Gomes, Raquel Simões de Almeida, Constantino Martins, Nuno Fonseca, Joaquim Santos, António Marques, and Javier Pereira

Institute of Engineering Oporto Polytechnic Institute Porto, Portugal
LabRP, Center for Rehabilitation Research, School of Health, Polytechnic of Porto, Porto, Portugal
CITIC Research Center, University of A Coruña, A Coruña, Spain
Correspondence: velosogomes@ess.ipp.pt

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: This paper describes a web-based application with the purpose of improving the self-management of people with schizophrenia. The main objective with the creation of a web application for mental health management was to provide easy and convenient access to resources, information, and tools that can help individuals manage their mental health issues more effectively and with greater autonomy. The application integrates different functionalities, such as stress management system, problem solving tools, medication control, symptoms monitoring and social interaction. In addition to self-management, the application facilitates interaction between the person with schizophrenia and the health professional through telemonitoring systems, reducing risks and increasing the patient's quality of life.

1 Introduction

Schizophrenia is a chronic mental disorder that affects the way a person thinks, feels, and behaves. It is characterized by symptoms such as delusions, hallucinations, difficulty concentrating, disorganized thinking, and emotional changes. Schizophrenia can have a significant impact on daily life and social relationships (National Institute of Mental Health, 2023). The prevalence of schizophrenia can vary depending on various factors, including geographical region, diagnostic criteria used and the populations studied2. The World Health Organization (WHO) estimates that the prevalence of schizophrenia ranges from 0.3% to 0.7% of the global population. This estimate encompasses different studies and research conducted in various regions of the world (National Institute of Mental Health, 2023). The association of schizophrenia with technology has been explored in various areas, including the use of digital technologies to assist in the diagnosis, treatment and continuous support of individuals with schizophrenia (Hoşgelen et al., 2020). The association of web applications with schizophrenia has been investigated in several studies, focusing on the potential benefits of web-based interventions for individuals with schizophrenia (Rotondi et al., 2010), (Almeida and Marques, 2022). These sources highlight the potential of web applications as a mean to provide support, psychoeducation and interventions for individuals with schizophrenia. They discuss the development and implementation of web-based platforms to improve access to information, services and social support for this population.

2 Methods

The collaboration between mental health professionals and informatics engineers drove the development of weCope, a web-based application. From an abstract perspective, the weCope system can be viewed as a collection of tools that can be used in psychotherapy through interactions between various stakeholders and interfaces. The weCope web platform uses a centralized server, enabling the registration and monitoring of patients' therapeutic progress through the therapist's interface. Furthermore, the weCope application aligns with the project's core philosophy of supporting the rehabilitation process for individuals with schizophrenia, offering a platform that aids in computer-assisted treatment. In other words, weCope provides a mean for patients to engage with their therapy course using computer-assisted tools, supporting their rehabilitation journey. The web platform consists of three modules, namely the administrator's interface, the therapist's interface and the patient´s interface, each serving different functions. The administrator interface provides administrators with tools to manage user accounts. This includes tasks such as registering new user accounts, updating incorrect or outdated information and managing the association between therapists and patients. The therapist's interface is designed for professional psychotherapists to interact with their patients. Through this interface, therapists can monitor and review their patients' inputs in the Daily Log, Objectives and Problem Solving modules. This allows therapists to gain a deeper understanding of each patient's challenges and goals, enabling them to tailor the therapy accordingly and improve the patient's overall quality of life. The patient´s interface provides individuals with schizophrenia with several useful functionalities for self-management of their condition, which can be accessed from the Main Menu (Figure 1). This modules include a Daily Log (Figure 2), Goals and Progress, Problem Solving (Figure 3), Chillout and sending messages to the Patient´s Therapist (Figure 4).

Figure 1: Patient interface main menu.

Figure 2: Symptom monitoring module exemple.

Figure 3: Problem solving module exemple.

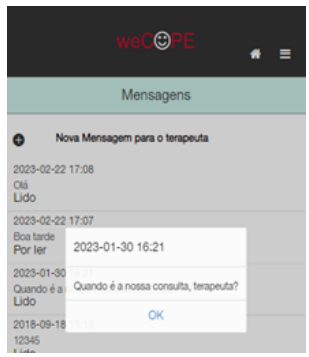


Figure 4: Message exchange from Patient to Therapist module example.

3 Results

In a previous study carried out on the weCope application, it was found regarding usability that the majority of participants (45%) used the application two or three times a week and recognized the modules “Goals”, “Relaxation” and “Messages with the therapist” being the most used by them. The participants’ view regarding the browsing experience, visual aesthetics, and learning process with weCope was favorable. Participants were satisfied with the weCope system, considering the application useful for disease management (89%) and 78% said they

would continue using the application (Almeida et al., 2019). In another previous study analyzing the weCope application, the results achieved with the use of a usability survey for therapists and patients showed general approval of the user interfaces (Almeida et al., 2017).

4 Conclusion

Empowering patients to manage their illness is crucial for regaining control and fostering a collaborative relationship with healthcare professionals. New self-management technologies have the potential to transform the healthcare experience, including health symptoms and recovery progress. The positive feedback from participants highlights weCope's user-friendliness and empowerment potential. Notably, 89% found it useful for managing their condition, with 78% intending to continue usage. In a separate study involving therapists and patients, the application received widespread approval, reflecting the broader acceptance of self-management technologies like weCope in healthcare. These findings underscore their role as valuable tools for enhancing patient-provider collaboration.

Acknowledgements

CITIC is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS).

Bibliography

R. Almeida, C. Martins, A. Marques, D. Benevides, A. Costa, C. Queirós, T. Sousa, A. Almeida, and L. Fonseca, N.and Faria. Development of a hybrid application for psychotic disorders self-management. *Advances in Intelligent Systems and Computing*, vol 615. Springer, Cham., 2017.

R. Almeida, T. Sousa, A. Couto, A. Marques, C. Queirós, and C. Martins. Development of wecope, a mobile app for illness self-management in schizophrenia. *Arch Clin Psychiatry*, 1 (46):1–4, 2019.

R. S. d. Almeida and A. Marques. User engagement in mobile apps for people with schizophrenia: A scoping review. *Frontiers in Digital Health*, 4, 2022.

E. I. Hoşgelen, B. B. Akdede, and K. Alptekin. Prevalence and use of technological devices and internet among patients diagnosed with schizophrenia and schizoaffective disorder. *Schizophrenia Bulletin*, (Supplement 1)(46):S173–S174, 2020.

National Institute of Mental Health. Schizophrenia. nimh.nih.gov/health/topics/schizophrenia/, 2023. [Online; accessed 7-August-2023].

A. J. Rotondi, G. L. Haas, C. M. Anderson, C. E. Newhill, M. B. Spring, R. Ganguli, and J. B. Rosenstock. Web-based psychoeducational intervention for persons with schizophrenia and their supporters: One-year outcomes. *Psychiatric Services*, 12(56):1621–1628, 2010.

Unveiling the Dark Side of Social Media: Developing the First Galician Corpus for Misogyny Detection on Twitter and Mastodon

Lucía Álvarez-Crespo, and Laura M. Castro

Models and Applications of Distributed Systems (MADS), Faculty of Computer Science, Universidade da Coruña, 15071 A Coruña, Spain
Correspondence: lucia.maria.alvarez.crespo@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: This work aims to develop the first Galician corpus for the detection of misogyny on Twitter and Mastodon. We collect and analyze linguistic data in Galician on these social media platforms, identifying manifestations of misogyny in digital communication. The process involves data collection, text selection, and normalization, followed by thorough cleaning. We apply machine learning techniques to train accurate models for classifying the presence of misogyny. The resulting corpus facilitates analysis and study by research teams interested in misogyny in Galician. This scientific advancement contributes to the understanding and prevention of misogyny in the Galician-speaking community, promoting equality and respect in digital communication in Galician.

1 Introduction

Social media's rapid growth has exposed a rise in misogyny. Studies reveal a concerning prevalence of harassment, particularly of sexual nature, aimed at women, online (Emily A. Vogels, 2021). Despite this, natural language processing (NLP) research often overlooks minority languages like Galician, impeding our understanding and combating online misogyny in specific linguistic contexts.

Misogyny takes various forms online, including discrimination, threats, and sexual objectification (Fersini et al., 2018; Siapera, 2019). NLP emerges as a promising tool for understanding this discourse, but identifying subtle misogyny can be challenging due to cultural and contextual differences. However, automated misogyny detection is increasingly important in addressing this social issue.

This work introduces the GalMisoCorpus[1], a first approach of a Galician-language dataset, harvested from Twitter[2] and Mastodon, and develops and evaluates machine learning models for misogyny detection.

2 Related work

Detecting misogynistic discourse on social media is complex. This review explores innovative approaches in sentiment analysis for understanding abusive behavior.

[1] *https://github.com/luciamariaalvarezcrespo/GalMisoCorpus2023/*

[2] As of the summer of 2023, Twitter changed its name to 'X'. However, for the sake of readability and common usage, we will continue to use the term 'Twitter' throughout this text.

Research teams have been using sentiment analysis for some time, especially on Twitter. This technology helps understand public attitudes in various contexts worldwide. For example, it was used to study opinions on COVID-19 (Manguri et al., 2020). In Japan, computational tools analyzed misogynistic hate speech towards female politicians on Twitter, combining quantitative and qualitative methods (Fuchs and Schäfer, 2021).

Addressing online misogyny is crucial due to its prevalence. Automation can assist in rapidly identifying and combating abusive content. Challenges include distinguishing between active and passive misogyny. Tasks like IberEval's AMI and EvalIta's AMI approached misogyny as a binary classification problem (Fersini et al., 2018, 2020).

Mastodon, a decentralized network, lacks sentiment analysis research. Recent studies are emerging (Cerisara et al., 2018; Monachelis et al., 2022). However, the scarcity of academic literature hinders the understanding of misogyny in Mastodon communities.

This project aims to optimize a misogyny detection system in Galician on Twitter and Mastodon. The goal is creating accurate machine-learning models for identifying misogynistic messages. Limited literature exists in Galician on this topic, making this research a valuable contribution to understanding misogyny in Galician on these platforms.

3 Corpus

The data collection phase is crucial, especially for our dataset focused on Galician language texts mirroring common online language and topics, including misogyny. We opt for automated methods following a binary classification approach (Fersini et al., 2018; García-Díaz et al., 2021) for misogyny detection, requiring a dataset with Galician samples from both Twitter and Mastodon, classified as misogynistic or not.

We begin by acquiring a non-misogynistic class via toots from the Galician Mastodon instance. Using tailored data scraping tools, we retrieve toots in Galician from this instance, assured by stringent moderation guidelines. This allows us to select non-misogynistic toots without exhaustive content review. The Mastodon API facilitates this process, ensuring efficient and systematic data collection from May 2022 to February 2023.

Given the lack of prior work in Galician misogyny detection, we turn to the Spanish MisoCorpus-2020 (García-Díaz et al., 2021). This corpus, focused on Spanish misogyny, serves as a valuable resource. Despite being in Spanish, we leverage the CIXUG[3] translator to convert the texts to Galician, ensuring alignment with our study language. To circumvent translation issues, only tweets labeled as "Spanish from Spain" will be translated.

Merging data from the Galician Mastodon instance (representing a non-misogynistic class) with misogynistic tweets from Twitter, we create a balanced dataset. This enables accurate training and evaluation of our binary classification model. This dual approach ensures an effective distinction between misogynistic and non-misogynistic content in Galician on social media. This positions us to tackle the challenge of misogyny detection on digital platforms.

The resulting corpus comprises labeled toots and tweets, divided into two subsets. The `toots.csv`[4] contains a substantial 19,387 non-misogynistic samples. In contrast, `tweets.csv` is smaller, with 1,307 samples, as tweets lacking geolocation info in Spain were removed.

We've shared the compiled corpus with the community, though Twitter guidelines only allow sharing tweet IDs to respect content creators' rights online[5]. Also, the GalMisoCorpus is released under an open license to promote its use and the advancement of this type of research.

[3] *https://tradutor.cixug.gal/*

[4] *https://github.com/luciamariaalvarezcrespo/GalMisoCorpus2023/*

[5] *https://github.com/luciamariaalvarezcrespo/GalMisoCorpus2023/*

4 Evaluation

The model development process included data preprocessing, algorithm selection, and performance evaluation. Preprocessing, following the pipeline of García-Díaz et al. (2021), involved: (1) Converting text to lowercase, (2) Removing HTML tags and blank lines, (3) Eliminating mentions and hashtags, (4) Spelling error correction (not done due to limited resources), (5) Reducing continuously repeated symbols (see figure 1).

Figure 1: Proposed pipeline.

After preprocessing, fastText (Joulin et al., 2016) was used to generate sentence embeddings. Algorithms like Random Forest (RF), Support Vector Machine (SVM), and Linear Support Vector Machine (LSVM) were employed. SVMs used manual hyperparameters (polynomial kernel and C=1), RF retained default settings, and LSVM employed L1 penalty and squared hinge loss. Data was split 70-30 for training and testing, facilitating model evaluation and generalization testing. The evaluation used 10-fold cross-validation, dividing data into 10 parts for 10 iterations, measuring model performance with F1-score, ideal for imbalanced datasets. For evaluation, a Bag of Words (BoW) text representation was applied, with TF-IDF for unigram relevance calculation. The chi-squared method selected the most relevant unigrams for distinguishing misogynistic and non-misogynistic texts.

In summary, the pipeline included text conversion, relevance calculation, feature selection, and classification. This approach aimed to accurately identify misogynistic content in Galician on Twitter and Mastodon.

5 Results

The results observed in table 1 reveal that machine learning models, including Random Forest (RF), Support Vector Machine (SVM), and Linear Support Vector Machine (LSVM), exhibit very similar performance in this iteration. The F1-score, a metric that balances precision and recall, is high for all three models, approximately 0.90. This indicates that they all achieve a good balance between accurately classifying positive cases and finding all positive cases. Precision is high for all three models, with values above 0.86, indicating a minimization of false positives. Recall, which assesses the ability to find all positive cases, is also high, with values around 0.93. Precision and recall align with the accuracy metric, which is approximately 0.93 for all three models, indicating a high proportion of correct predictions overall.

	RF	**SVM**	**LSVM**
F1-score	0.9038	0.9101	0.8975
Precision	0.9390	0.9428	0.8664
Recall	0.9348	0.9391	0.9308
Accuracy	0.9348	0.9391	0.9308

Table 1: Metrics

Among the three models, SVM stands out as the best choice in the first iteration due to its higher F1-score, approximately 0.9101. It also has high precision (approximately 0.9428) and

recall (approximately 0.9391). This means that the SVM model has an excellent ability to correctly classify positive cases, minimize false positives, and find the most positive cases in the dataset. Considering these factors, SVM emerges as the strongest choice due to its combination of a high F1-score, high precision, and high recall.

In summary, the results of the first iteration are promising and indicate that the models perform well in the task of classifying misogyny in the Galician language corpus.

6 Discussion

Despite sentiment analysis' popularity, few solutions focus on misogyny detection, especially in minority languages like Galician. Detecting misogyny on platforms like Twitter has the potential to foster respectful online communities and protect human rights. The results, with high precision, recall, and F1-scores, show promise in identifying misogynistic content in Galician on Twitter and Mastodon.

Bibliography

C. Cerisara, S. Jafaritazehjani, A. Oluokun, and H. Le. Multi-task dialog act and sentiment recognition on mastodon. *arXiv preprint arXiv:1807.05013*, 2018.

Emily A. Vogels. The state of online harassment. *https://www.pewresearch.org/internet/2021/01/13/the-state-of-online-harassment/*, 2021. [Online; accessed 13-September-2023].

E. Fersini, P. Rosso, and M. Anzovino. Overview of the task on automatic misogyny identification at ibereval 2018. *Ibereval@ sepln*, 2150:214–228, 2018.

E. Fersini, D. Nozza, and P. Rosso. Ami @ evalita2020: Automatic misogyny identification. In *Proceedings of the Seventh Evaluation Campaign of Natural Language Processing and Speech Tools for Italian*, pages 21–28, 2020. doi: 10.4000/books.aaccademia.6764.

T. Fuchs and F. Schäfer. Normalizing misogyny: hate speech and verbal abuse of female politicians on japanese twitter. In *Japan forum*, volume 33, pages 553–579. Taylor & Francis, 2021.

J. A. García-Díaz, M. Cánovas-García, R. Colomo-Palacios, and R. Valencia-García. Detecting misogyny in spanish tweets. an approach based on linguistics features and word embeddings. *Future Generation Computer Systems*, 114:506–518, 2021.

A. Joulin, E. Grave, P. Bojanowski, and T. Mikolov. Bag of tricks for efficient text classification. *arXiv preprint arXiv:1607.01759*, 2016.

K. H. Manguri, R. N. Ramadhan, and P. R. M. Amin. Twitter sentiment analysis on worldwide covid-19 outbreaks. *Kurdistan Journal of Applied Research*, pages 54–65, 2020.

P. Monachelis, P. Kasnesis, L. Toumanidis, C. Patrikakis, and P. Papadopoulos. Evaluation and visualization of trustworthiness in social media – eunomia's approach. In *2022 IEEE 46th Annual Computers, Software, and Applications Conference (COMPSAC)*, pages 217–222, 2022.

E. Siapera. *Online Misogyny as Witch Hunt: Primitive Accumulation in the Age of Techno-capitalism*, pages 21–43. Springer International Publishing, Cham, 2019. ISBN 978-3-319-96226-9.

Predicting Inflow Flow in Hydraulic Dams Using Artificial Neural Networks

Alberto Fernandez Sanchez, Juan Ramon Rabuñal Dopico, Daniel Rivero Cebrian, Alejandro Celestino Pazos Sierra, Marcos Gestal Pose, and Luis Cea Gomez

RNAsa-IMEDIR, Faculty of Computer Science, Universidade da Coruña, 15071 A Coruña, Spain
CITEEC Research Center, Universidade da Coruña, 15071 A Coruña, Spain
Correspondence: alberto.fsanchez@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: Accurate prediction of inflow in dams plays a crucial role in water resource management Kim et al. (2019); Vargas-Garay et al. (2018); Zhong et al. (2018) and risk mitigation Costabile et al. (2020); Rabuñal et al. (2007). This study focuses on the Portodemouros dam (located between the provinces of A Coruña and Pontevedra), where a model based on a Long Short-Term Memory (LSTM) artificial neural network has been implemented to predict dam inflow.
The results demonstrate the well-established effectiveness of the LSTM network in flow prediction Dongkyuna and Seokkoob (2021); Jo and Jung (2023); Li et al. (2020) applied to the Portodemouros dam compared to other models. This comparison has already been performed in other studies with both mathematical models Amirreza et al. (2022); Ansori and Anwar (2022); A.R1 et al. (2018); Beck et al. (2017); Ciabatta et al. (2016); Costabile et al. (2020); Fan et al. (2013); Heřmanovský et al. (2017); Kim et al. (2019); Vargas-Garay et al. (2018); Zhong et al. (2018), genetic programming Aytek et al. (2008); Havlíček et al. (2013); Heřmanovský et al. (2017); Rabuñal et al. (2007) and other machine learning algorithms Jo and Jung (2023). Combining precipitation data from multiple regions and meteorological forecasts significantly enhances the model's ability to anticipate variations in dam inflow. This improved accuracy is essential for early flood detection and informed decision-making in dam operation.
This study forms part of the Marine Science programme (ThinkInAzul) supported by Ministerio de Ciencia e Innovación and Xunta de Galicia with funding from European Union NextGenerationEU (PRTR-C17.I1) and European Maritime and Fisheries Fund.

1 Introduction

Predicting the variation in surface water flows (runoff) based on local precipitation (rainfall-runoff models) has been addressed from the 19th century to the present day using mathematical models and later machine learning models separately or with comparative studies.

2 Dataset

2.1 Sources

For the training of the neural networks, a dataset of parameters from the Portodemouros dam, located in the Ulla River basin (42°50′47″N 8°08′26″W) with a total capacity of 297 hm^3, which includes a set of 4554 daily records of dam inflow, precipitation, and 1, 2, and 3-day predictions provided by meteogalicia, was used. See Figure 1.

Figure 1: Dam and pluviometer stations situation map

To prepare the complete dataset, several sources were used, including records from the dam, which include filling level, dam inflow, and dam flow outputs. These data will allow us to link this study to a subsequent one on human behavior modeling.

2.2 Preliminary Analysis: Discharge Time

To establish a rainfall-runoff model, it is necessary to estimate the time it takes for precipitation at each measuring station to affect the dam inflow. To accomplish this, the correlation between precipitation in several areas and changes in the stream was assessed. Four pluviometer station records were used to measure the aforementioned correlation: Arzua, Melide, Serradofaro, and Olveda. These stations were selected due to their proximity to the upper part of the Ulla River, demonstrating that all basins take less than 24 hours (1 day) to reach the dam, regardless of the soil water saturation state.. See Figure 1.

Table 1: Maximum correlation between precipitation and dam inflow (in days)

Soil Moisture	Arzua	Olveda	Serradofaro	Melide
Dry	1	1	1	1
Half-wet	1	1	1	1
Wet	1	1	1	1
Very wet	1	1	1	1
Saturated	1	1	1	1

2.3 Cross-Validation

The dataset has been split into subsets based on the year of each sample, resulting in 13 subsets (spanning from 2010 to 2022). Once these subsets were separated, 13 different sets of training, validation, and test data were prepared, such that the test set covers a complete year, the validation set the following complete year (or the first year if the test set includes the last year), and the training set includes the remaining years. This approach helps mitigate the impact of year-to-year variability on both model training and evaluation metrics. Subsequently, each of these sets was normalized by subtracting the mean of each column and dividing the result by the standard deviation of that column Li et al. (2020). The measurements were calculated by

averaging across all k-fold subsets.

To compare the model's performance, other machine learning models and a naive model have been used:

- Linear Regression Tallarida and Murray (1987)
- SVM (Support Vector Machine for Regression) C. and V. (1995)
- Naive model (using the last known dam inflow value as the prediction for each sample)

3 Results

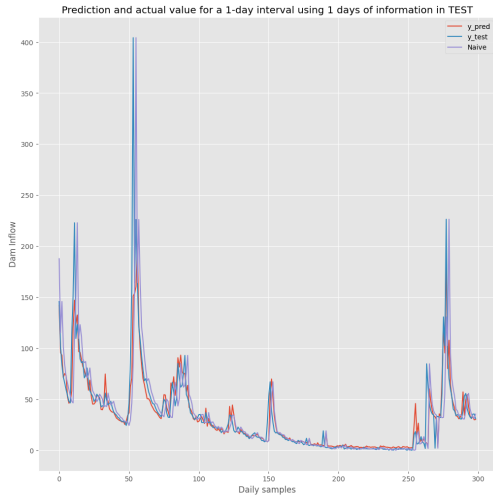


Figure 2: LSTM predictions (y_pred), Naive predictions (Naive) and real values (y_test) of dam inflow

Below are the results using different metrics commonly used as a basis in most rainfall-runoff studies, such as the coefficient of determination (R2) Ansori and Anwar (2022); Aytek et al. (2008); Jo and Jung (2023); Zhong et al. (2018) and the Nash-Sutcliffe Efficiency coefficient (NS or NSE) Ansori and Anwar (2022); A.R1 et al. (2018); Ciabatta et al. (2016); Dongkyuna and Seokkoob (2021); Havlíček et al. (2013); Heřmanovský et al. (2017); Jo and Jung (2023); Li et al. (2020).

Nash-Sutcliffe Efficiency (NS) is a widely used metric in hydrology to assess the accuracy of runoff models or hydrological forecasts in comparison to observed data McCuen et al. (2006). It was proposed by John R. Nash and James C. Sutcliffe in 1970 and is considered an effective measure to evaluate the overall performance of a model in simulating runoff events over time.

Model output metrics

Table 2: Prediction Evaluation

Model	R2 train	R2 test	NS train	NS test
LR	0.849	**0.816**	0.822	**0.782**
SVM	0.877	0.76	0.795	0.494
Naive	0.777	0.717	0.776	0.715
LSTM	**0.881**	0.8	**0.845**	0.709

These metrics were obtained using a 13-fold cross-validation with complete years as subsets. Each value is computed by a mean of each subset metric giving a better estimation of each measure avoiding any bias by extreme climatic difference between dataset years range.

4 Conclusions

As can be observed, the LSTM network achieves a training score of 0.845 and a testing score of 0.709 for NS, which measures the goodness of rainfall-runoff models, regardless of the dataset used. These NS values are very similar to those found in other studies, such as 0.73 in training by Ansori and Anwar (2022) or 0.782 in testing by Jo and Jung (2023).

LR has also demonstrated strong performance in the test dataset, achieving the highest values of R2 and NS in a subset of 365 samples, which makes it a reliable estimator for the rainfall-runoff model in this specific dataset.

This results confirms the robustness of LSTM algorithm in this dataset and establishes it as a valid foundation for further fine-tuning or complementation with other types of algorithms. For example, it can be used to model dam operator behavior, allowing predictions of dam overflow based on inflow flow forecasted by an algorithm like LSTM, along with another model predicting the orderly dam release by the operator.

In addition to the rainfall-runoff model, analyzing the behavior of the dam operator when managing dam releases is crucial to enhance the robustness of an early flood and overflow alarm system. This helps minimize operational risks while maximizing effectiveness in complex situations. To model their behavior, we have access to a historical record of actions based on the variables previously used in the rainfall-runoff model.

Bibliography

M. Amirreza, D. Amirhossein, S. Gerrit, and T. Massoud. Daily reservoir inflow forecasting using weather forecast downscaling and rainfall-runoff modeling. *Journal of Hydrology: Regional Studies*, 44(101228):1–20, 2022.

M. B. Ansori and N. Anwar. The trmm rainfall-runoff transformation model using gr4j as a prediction of the tugu dam reservoir inflow. *GEOMATE Journal*, 23(97):45–52, 2022.

A. R. A.R1, S. L.M, B. H., A. J.L., J. K., and S. S.K. Reservoir sediment inflow prediction using integrated rainfall-runoff and discharge–sediment model. *International Journal of Engineering and Technology*, 7(4):917,923, 2018.

A. Aytek, M. Asce1, and M. Alp. An application of artificial intelligence for rainfall–runoff modeling. *Journal of Earth System Science*, 117(2):145–155, 2008.

H. E. Beck, A. I. J. M. van Dijk, A. de Roo, E. Dutra, G. Fink, R. Orth, and J. Schellekens. Global evaluation of runoff from 10 state-of-the-art hydrological models. *Hydrology and Earth System Sciences*, 21(6):2881–2903, 2017.

C. C. and V. V. Support-vector networks. *Machine Learning*, 1(20):273–297, 1995.

L. Ciabatta, L. Brocca, C. Massari, T. Moramarco, S. Gabellani, S. Puca, and W. Wagner. Rainfall-runoff modelling by using sm2rain-derived and state-of-the-art satellite rainfall products over italy. *International Journal of Applied Earth Observation and Geoinformation*, 48:163–173, 2016.

P. Costabile, C. Costanzo, D. Ferraro, F. Macchione, and G. Petaccia. Performances of the new hec-ras version 5 for 2-d hydrodynamic-based rainfall-runoff simulations at basin scale: Comparison with a state-of-the art model. *Water*, 12(9), 2020.

K. Dongkyuna and K. Seokkoob. Data collection strategy for building rainfall-runoff lstm model predicting daily runoff. *Journal of Korea Water Resources Association*, 54(10):795–805, 2021.

F. Fan, Y. Deng, X. Hu, and Q. Weng. Estimating composite curve number using an improved scs-cn method with remotely sensed variables in guangzhou, china. *Remote Sensing*, 5(3): 1425–1438, 2013.

V. Havlíček, M. Hanel, P. Máca, M. Kuráž, and P. Pech. Incorporating basic hydrological concepts into genetic programming for rainfall-runoff forecasting. *Springer-Verlag*, 95(1):363–380, 2013.

M. Heřmanovský, V. Havlíček, M. Hanel, and P. Pech. Regionalization of runoff models derived by genetic programming. *Journal of Hydrology*, 547:544–556, 2017.

Y. Jo and K. Jung. Comparative study of machine learning and deep learning models applied to data preprocessing methods for dam inflow prediction. *GeoAI Data Society*, 5(2):92–102, 2023.

S.-H. N. Kim, W.-S. Bae, and Deg-Hyoc. An analysis of effects of seasonal weather forecasting on dam reservoir inflow prediction. *Journal of Korea Water Resources Association*, 52(7):451–461, 2019.

W. Li, A. Kiaghadi, and C. N. Dawson. High Temporal Resolution Rainfall Runoff Modelling Using Long-Short-Term-Memory (LSTM) Networks. *arXiv e-prints*, art. arXiv:2002.02568, 2020.

R. H. McCuen, Z. Knight, and A. G. Cutter. Evaluation of the nash-sutcliffe efficiency index. *Journal of Hydrologic Engineering*, 11(6):597–602, 2006.

J. R. Rabuñal, J. Puertas, J. Suárez, and D. Rivero. Determination of the unit hydrograph of a typical urban basin using genetic programming and artificial neural networks. *Hydrological Processes*, 21(4):476–485, 2007.

R. J. Tallarida and R. B. Murray. *Linear Regression I*. Springer, New York, NY, 1987.

L. Vargas-Garay, O. D. Torres-Goyeneche, and G. A. Carrillo-Soto. Evaluation of scs - unit hydrograph model to estimate peak flows in watersheds of norte de santander. *Respuestas journal of engineering sciences*, 24(1):6–15, 2018.

W. Zhong, R. Li, Y. Q. Liu, and J. Xu. Effect of different areal precipitation estimation methods on the accuracy of a reservoir runoff inflow forecast model. *IOP Conference Series: Earth and Environmental Science*, 208(1):012043, dec 2018.

Ultra Wideband for Shipyard 5.0 Indoor Gantry Crane High-Precision Positioning

Ángel Niebla-Montero, Paula Fraga-Lamas, José Varela-Barbeito, and Tiago M. Fernández-Caramés

Centro Mixto de Investigación UDC-Navantia, Universidade da Coruña, Edificio de Batallones, s/n, 15403 Ferrol, España
Centro de Investigación CITIC, Universidade da Coruña, 15071 A Coruña, España
Navantia S. A., Unidad de Producción de Ferrol, Taxonera, 15403 Ferrol, España
Correspondence: angel.niebla@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: The shipyard of the future, called Shipyard 5.0, is a highly technological environment where real-time monitoring of products, by-products and transport vehicles is crucial. Among the technologies able to locate such elements indoors, Ultra Wideband (UWB) is a good option for providing accurate positioning. However, the use of UWB in shipyards faces challenges due to interference from metallic objects, which impacts its accuracy. To validate the use of UWB in a shipyard, this paper presents tests that were carried out in workshops that Navantia owns in Ferrol's estuary, where UWB tags were placed on a gantry crane hook. The presented results show the performance of the system when locating the hook in 3D and the impact of attaching multiple tags to the hook to harness signal diversity. Specifically, a relatively low error is obtained when estimating only the height of the gantry crane hook (approximately 1 m), while the three-dimensional positioning error reached an error of between 2 and 3 m for the z-axis.

1 Introduction

Industry 5.0 is the natural evolution of Industry 4.0. While Industry 4.0 is primarily concerned with the automation and digitization of industrial processes, Industry 5.0 seeks greater integration of human and technological capabilities (Directorate-General for Research and Innovation, 2021). Furthermore, Industry 5.0 emphasizes sustainability and social responsibility, attempting to balance technological progress with human well-being.

Between 2015 and 2021, Navantia, Spain's largest shipbuilding company, dedicated itself to researching the application of various technologies to shipyards and ships through the so-called Joint Research Unit (JRU) "The Shipyard of the Future". One of the research lines of the JRU concentrated on the automatic identification and traceability of shipbuilding components, tools, and products throughout their life cycle. Throughout the life of this line, numerous studies on auto-identification technologies (Auto-ID) were conducted (P. Fraga-Lamas et al., 2016, 2017). In line with these developments, the goal of the work presented in this paper was to assess the performance of Ultra-WideBand (UWB) technology for 3D object localization in Navantia's workshops when using a gantry crane to move them.

2 Design, implementation and evaluation of the system

2.1 Used hardware and software

To validate the use of UWB in crane hook positioning, two independent systems were designed and built. Then, their performance was compared when determining the height of the crane.

The first system consisted of the following components:

- A single DWM1001 module (Qorvo, 2023) attached to the crane hook. Such a module acted as a tag (i.e., as a mobile node).
- Several DWM1001 modules were deployed along the workshop. Such modules acted as anchors (i.e., as fixed nodes).

Every DWM1001 module embeds a Nordic nRF52832 System-on-Chip (SoC), which provides Bluetooth connectivity and that allows the tag position to be transmitted to a remote smartphone (i.e., to an Android application). Moreover, each DWM1001 module includes a Decawave DW1000 UWB transceiver, which handles bidirectional range exchanges between tags and anchors. The manufacturer provides a default firmware for the DWM1001 modules (Decawave, 2017) that makes it simple to implement a basic RTLS (Real-Time Location System).

The second developed system is made up of only two DWM1001 modules: One module serves Initiator, while the other one acts as Responder. The Initiator sends a frame, waits for a response from the Responder, and then calculates the distance based on the TOF (Time of Flight). The output is sent via the provided UART port, to which a Raspberry Pi Zero is connected to process and to save the position data.

Although the first and second system employ a one-way TWR (Two Way Ranging) scheme (SS-TWR), they are different on the fact that the RTLS created by the first system requires at least three anchors to calculate their location based on beacon message anchor positions.

2.2 Experimental setup

The experiments presented in this paper were carried out in a workshop at Navantia's shipyard of Fene (Galicia, Spain), where seven anchors of varying heights were installed. To ensure a wide coverage area, three anchors were placed on one side of the workshop, and three more on the other side, with each side anchors separated by 15 m. The positioning of the anchors in the workshop is shown in Figures 1 and 2.

During the tests, two tags were used simultaneously to verify the use of UWB for 3D gantry crane hook positioning, as shown in Figure 16.3(a). Such tags were placed on top of the crane hook as it is shown in Figure 16.3(b). While the anchors located throughout the workshop helped one tag determine its 3D position, the other tag was only responsible for measuring the crane height. This was achieved by placing another tag on top of the crane (as shown in Figure 4), using a Raspberry Pi Zero powered by a battery to store the collected data. Moreover, the Raspberry Pi Zero can be used to transmit the information opportunistically to a remote smartphone or computer.

2.3 Results

Initial tests

Some measurements were taken during the initial placement of the modules in the crane to determine the initial error for the two tested systems. Figure 16.3(b) shows the initial moment when the modules were placed and the first measurement was taken; the hook was at a height of 1 m. At that moment the system that only measured the height yielded a result of 0.8 m while the 3D positioning system was unable to maintain a constant result because in this first position, due to obstacles in the environment, it only received information from three or less anchors.

Figure 1: Test scenario (right side of the workshop) with the anchors (inside yellow squares).

Figure 2: Test scenario (left side of the workshop) with the anchors (inside yellow squares).

Tests at a height of 7.5 meters

In the first actual set of tests, the crane hook was placed at a height of 7 m. In such a position the height positioning system gave a result of 7.2 m (i.e., a 0.3 m error was obtained), which remained constant throughout the movement of the crane (i.e., a constant offset was observed).

On the other hand, the 3D positioning system gave a more variable result. It determined a

(a) One of the test moments in which the crane was used.

(b) Placement of the modules.

Figure 3: The crane hook with the modules attached.

height between 5 and 7 m, depending on the distance to the surrounding anchors. This variation in the results of several meters is due to the movement of the crane when traveling through the workshop.

The results of this first test are shown in Figure 5 between iterations 0 and 400. This substantial inaccuracy in the calculation of the 3D positioning might be attributed to the fact that the tag height was higher than the anchors, which were all between 3 and 4 m high, so the tag did not have a direct reference at that height.

Tests at a height of 3.5 meters

In a second set of tests the gantry crane hook was raised to a height similar to the one of the surround anchors (4 m). In such a scenario it was observed how the height positioning system determined a 3.3 m height, with a constant error of 0.2 m. The 3D system, on the other hand, began marking between 4 and 5 m, which represents an error of less than 0.5 m, but as the crane moved from one side of the workshop to the other, the error increased, reaching 1.5 m. The results of such tests are reflected in Figure 5 between iterations 400 and 1000.

Tests at a height of 5 meters

Finally, in a third set of tests, the gantry crane hook was lowered to a height of 5 m. The obtained results are shown in Figure 5 between iterations 1000 and 1200. It can be observed how the height positioning system marks constantly 5 m (i.e., there was no error), while the 3D positioning system determined a height of around 3 m, with an error of 2 m.

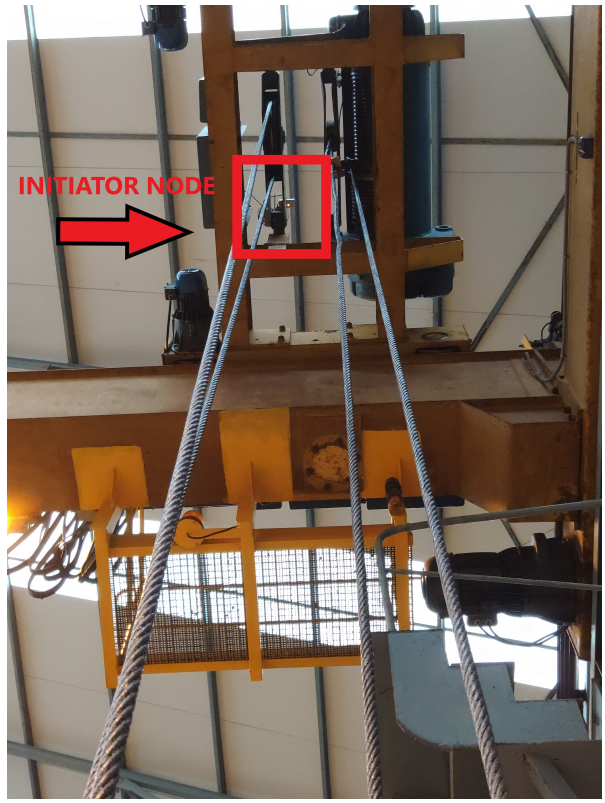


Figure 4: Placement of the Initiator node at the top of the crane.

3 Conclusions

This paper described an empirical evaluation of the precision of UWB tag positioning in a real-world environment (a workshop of Navantia in Fene's shipyard) and when locating the hook of a gantry crane. The differences in position accuracy have been shown for two independent systems: one that just determined the crane hook height and a second one that obtained the position in three dimensions. The results show that the height positioning system always manages to achieve a low error that remains constant and that can be easily corrected. The 3D positioning system, on the other hand, produces a larger error that varies as the crane moves through the workshop, sometimes producing errors of more than 2 m that would need to be corrected by modifying the system (e.g., by increasing the number of anchors or by placing such anchors in selected positions). In any case, the obtained results demonstrate the feasibility of using UWB technology to locate the crane hook in difficult industrial workplaces such as the one chosen, where numerous metallic objects are present.

Acknowledgment

This work has been funded by CITIC that is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS), by the Xunta de Galicia (by grant ED431C 2020/15), and by grant PID2020-118857RA-100 (ORBALLO) funded by MCIN/AEI/10.13039/501100011033.

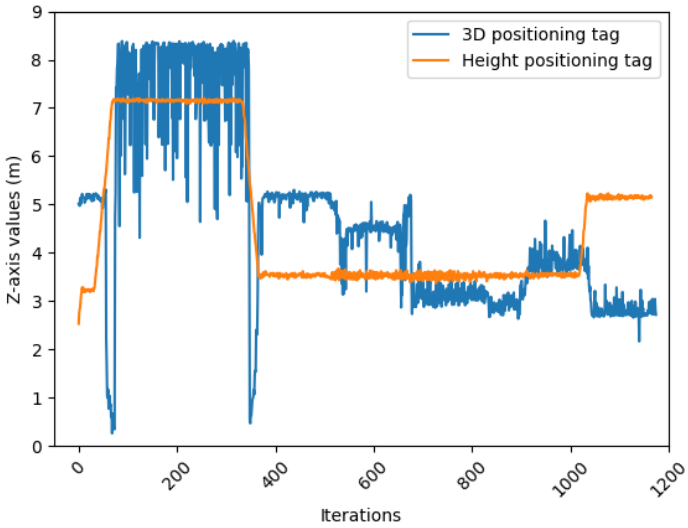


Figure 5: Height results obtained.

In addition, this work has been supported by Xunta de Galicia through Axencia Galega de Innovación (GAIN) by grant IN853C 2022/01, Centro Mixto de Investigación UDC-NAVANTIA "O estaleiro do futuro", which is ongoing until the end of September 2025. The support was inherited from both the starting and consolidation stages of the same project throughout 2015-2018 and 2018-2021, respectively. This stage is also co-funded by ERDF funds from the EU in the framework of program FEDER Galicia 2021-2027.

Bibliography

Decawave. Dwm1001 firmware user guide. *https://www.qorvo.com/products/d/da007972*, 2017. [Online; accessed 6-September-2023].

Directorate-General for Research and Innovation. Industry 5.0 - towards a sustainable, human-centric and resilient european industry. *https://rb.gy/rlg88*, 2021. [Online; accessed 8-September-2023].

P. Fraga-Lamas, D. Noceda-Davila, T. Fernández-Caramés, M. Díaz-Bouza, and M. Vilar-Montesinos. Smart pipe system for a Shipyard 4.0. *Sensors*, 12(16), 2016.

P. Fraga-Lamas, T. Fernández-Caramés, D. Noceda, and M. Vilar-Montesinos. RSS stabilization techniques for a real-time passive UHF RFID pipe monitoring system for smart shipyards. *2017 IEEE International Conference on RFID (RFID)*, pages 161–166, 2017.

Qorvo. Ultra wideband (uwb) transceiver development board dwm1001-dev. *https://www.qorvo.com/products/p/DWM1001-DEV#documents*, 2023. [Online; accessed 6-September-2023].

A Validation System for Academic Records based on Blockchain

Gabriel Fernández-Blanco, and Tiago M. Fernández-Caramés

GTEC, Faculty of Computer Science, Universidade da Coruña, 15071 A Coruña, Spain
Centro de Investigación CITIC, Universidade da Coruña, 15071 A Coruña, Spain
Correspondence: g.fblanco@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: Academic certificate forgeries and the lack of resources to tackle them is a problem that implies big costs to society. It harms the figure of the certificates and impairs the trust of academic institutions. The solution proposed in this paper is aimed at recording and verifying the academic merits through a decentralized application or Dapp. Such an application is supported by a smart contract deployed in the Ethereum blockchain with a simple frontend or interface. The application makes use of a decentralized storing system based on IPFS in order to consume the data that are not managed by the blockchain. To assess the performance of the developed system, the latency of its most common operation (read operations) is measured as the number of requests per second increase. The obtained results show that the system is really fast, being some nodes able to respond to more than 1,000 requests in less than one millisecond.

1 Introduction

The falsification of academic degrees continues to be a widespread and recurring problem that affects even the most developed countries. Traditional technological solutions suffer from vulnerabilities inherent to their design, such as the presence of centralized elements in their architecture (introducing single points of failure) or their incapability to detect unauthorized modifications. Although there are many solutions to verify the authenticity of physical or digital documents (A. Rustemi et al., 2023), they are not aimed at preventing frauds in the academic field. For instance, there are cases in which a student may have passed some exams or subjects due to a favorable treatment, which is a type of fraud committed even by political leaders.

To avoid such a kind of situations, the application proposed in this paper provides a solution for preventing academic title fraud by fully applying decentralized and distributed registry technologies. Thanks to the security and privacy provided by blockchain networks such as Ethereum, academic files become resistant to falsification or fraudulent modifications, since any change is recorded permanently and can be validated by each blockchain node.

Moreover, the proposed application stores the student records on a blockchain in a secure way. Such files consist of the marks received during a degree for the different subjects (or, ideally, for all the performed activities). There is a subset of administrators that are responsible for creating the transactions that update the blockchain. At any time, students can download and send their academic file to any interested entity, which can be easily verified through the developed application.

The proposed system is similar to the one proposed by the University of Zurich (J. Gresch et al., 2018). In addition, there are few solutions to this problem that fully decentralize its operation (A. Rustemi et al., 2023) A. Pfefferling and Kehling (2021). Working entirely on a decentralized architecture guarantees persistence on the academic information that is supported by the

nodes of a P2P (Peer-to-Peer) network, which avoids inappropriate modifications to the files or reliance on central servers (T. M. Fernández-Caramés and Fraga-Lamas, 2023). To provide such a feature, the system presented in this paper makes use of a database based on InterPlanetary File-System (IPFS), a purely decentralized protocol.

2 Design and Implementation

The proposed application consists of the subsystems shown in Figure 1, which are the following:

- **Blockchain**: record data are represented by their hash code. Such a hash is stored in the blockchain, supported by the rest of the data that are stored outside the blockchain (in the decentralized database). With this strategy, the size of the information stored in the blockchain is significantly reduced, but requires keeping the raw information outside the chain (off-chain). This increases transaction speed and cost savings, since operational costs decrease (i.e., the gas spent in Ethereum-based applications).
- **Decentralized Database (OrbitDB)**: This database stores the private information of teachers/professors and students (i.e., each academic record, AR), as well as their public keys. The application interacts with this database in order to login students and administrators, as well as for detecting new changes in the ARs. The information stored in OrbitDB is transparent by default, so developers have to take privacy-protection actions to protect users' data privacy.

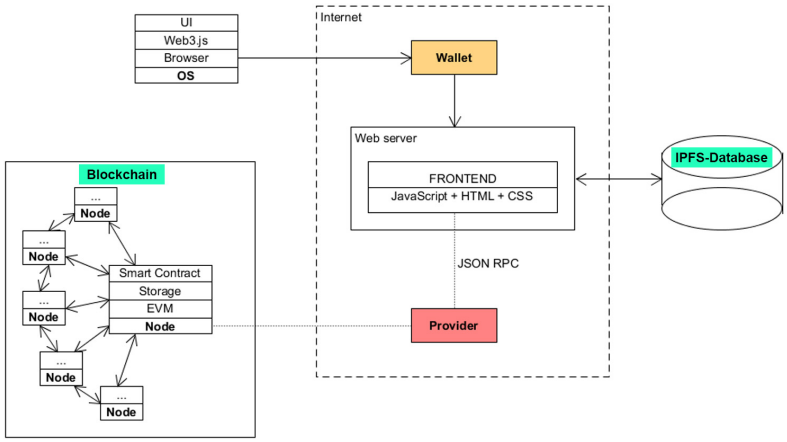


Figure 1: High-level view of the different subsystems.

The proposed application consists of the following actors (whose main interactions are shown in Figure 2):

- **Students**. Once registered, each student is able to see the constant evolution on his AR through the periodic exams and courses. Whenever he/she wants, he/she can download his/her AR (typically in a PDF file) and share it with any third-party, like an organization interested in hiring the student and in validating the student academic merits.
- **Teachers/Professors/Administrators**. They can be seen as permissioned teachers/professors, who are responsible for updating the ARs. For example, after the exams, teachers/professors will register the achieved marks of the students in the decentralized database. After such changes, the validity of these modifications needs to be approved by an administrator (e.g., by the head of the department or a person of the university/school administration). These updated records are introduced in a 'pending record

list' in which are stored all the new ARs and their changes. After confirming these last changes, the validator (i.e., the person that approves the registered changes) will perform a transaction on the blockchain to update the ARs. This update can be performed individually for each AR or in batch (i.e., for several ARs at the same time) to accelerate the process.

- **Third-party**. Any external user can verify the validity of an AR. This is simply performed by introducing the provided AR in the app, which will show almost instantly whether the introduced AR is part of the blockchain or not (by comparing its hash with all of the hashes stored in the blockchain).

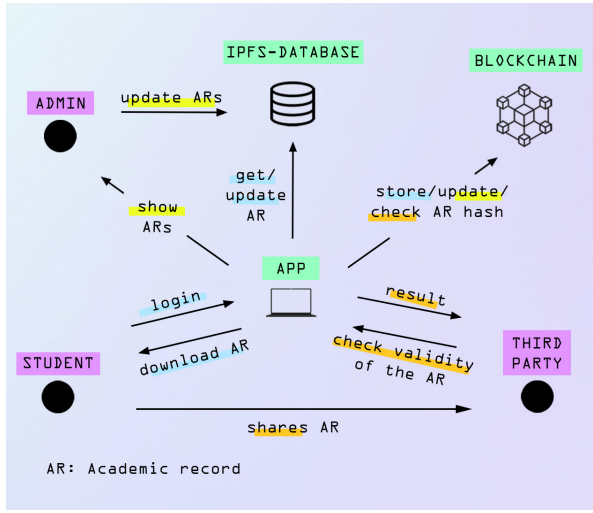


Figure 2: User interaction.

3 Results

Reading data from the blockchain is one of the most common operations in blockchain applications. Hence, it is important to see if the system maintains acceptable response times with high loads of read-operations. For this matter, a novel tool [3] was used to 'flood' the developed Ethereum-based application with read-only JSON-RPC requests such as `eth_call` (it executes a new message call immediately without creating a transaction on the blockchain) or `eth_getBalance` (it returns the balance of a specific account address).

Figure 3 illustrates the results of an `eth_call` test, measuring the evolution of the latency (the time it takes to receive the response from the call) as the number of requests per second (rps) grows. A customized network was created as a test environment with 12 nodes, 4 of them acting as validators/miners. However, note that node quantity is irrelevant when reading data from the blockchain, as transactions are actually not emitted.

As it can be observed in Figure 3, for the worst case evaluated (i.e., for a maximum of 2,000 requests per second), latency is really low (in the order of milliseconds). Obviously, higher loads above 2,000 rps will derive into a relevant increase in latency, but such loads are not usual in a typical academic data verification environment. Thus, the proposed provides all the advantages of decentralized systems in terms of security and trust with an equally acceptable operating time.

It is worth noting that Figure 3 shows a U-shape for each node curve, having more latency at 10 rps than for 100 rps. This is essentially due to the inner workings of the client cache and

the 'warming-up' effect of the nodes. However, from 100 rps the nodes behaves as expected, having a smooth increase of the latency as more rps are being executed. Since such an initial behaviour is inherent to the selected tool, it has been preserved in the Figure so that future researchers understand the obtained outputs.

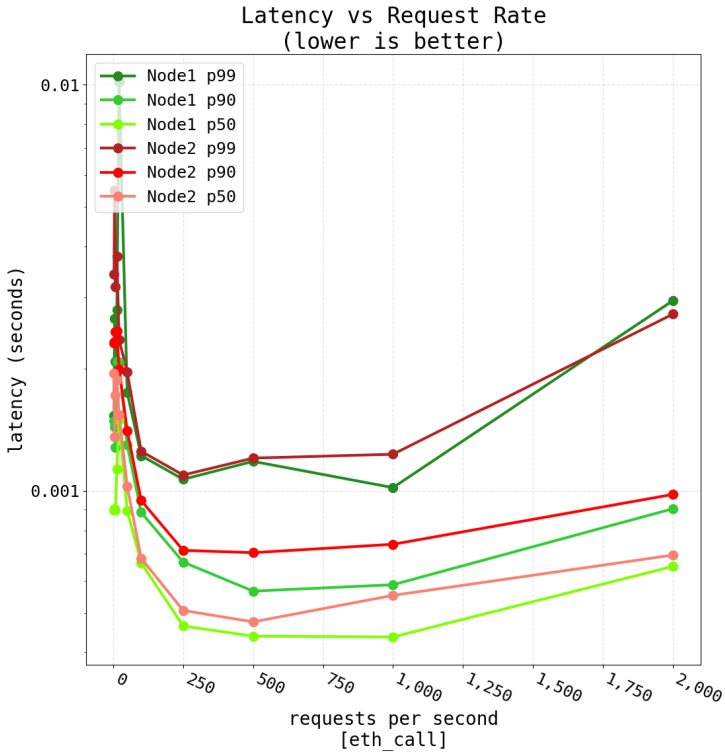


Figure 3: Latency results for the tested blockchain network.

4 Future work

The obtained results allow for concluding that the developed system provides a robust solution through smart contracts that reveals a great potential. For instance, the system can be used for automating a large number bureaucratic procedures such as title issuance or payments. Also, it could bring other new types of learning, as pay-as-you-go classes that employ smart contracts. Even, it could allow organizations to have greater collaborations with other institutions, so the verifiable records could be recognized as complete Curriculum Vitaes.

However, this technology is not without its challenges. It requires a careful design by specialised professionals, as its peculiarities involve new security vulnerabilities. Moreover, it is up to see its potential adoption, and how it would complaint with current data-protection laws as GDPR (General Data Protection Regulation) (i.e., the 'right to be forgotten' it is still hard to address due to its immutable nature). Nevertheless, there are more and more approaches to solve these problems such as the use of cryptography to separate the user's data from its physical-sensitive information.

In addition, the system would need to be improved to increase its energy-efficiency, since ecology is one of the foremost social debates in developed countries. Since the Bitcoin en-

ergy consumption has been widely criticized, there exists more interest in the literature that includes strategies and techniques to create green blockchains and to reduce blockchain energy consumption (N. Lei et al., 2021). Therefore, it is possible to use consensus protocols that improve massively electrical consumption over Proof-of-Work such as Proof-of-Authority. Also, it is necessary to study the feasibility of making use of an edge/fog-computing network of low-consuming devices (M. Sánchez-de la Rosa et al., 2023) in the developed Ethereum-based network, which may further decrease energy costs.

5 Conclusion

This paper has presented a decentralized tamper-proof academic record validation system that makes use of a blockchain (Ethereum) and a decentralized storage system (IPFS). The performance of the system has been tested and the obtained results show that the latency of the system is suitable for most of the potential applications that want to secure data or prevent fraud, while harnessing the benefits of decentralized technology (i.e., data privacy, security and immutability).

Acknowledgment

This work has been funded by CITIC that is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS), by the Xunta de Galicia (by grant ED431C 2020/15), and by grants PID2020-118857RA-100 (ORBALLO) and TED2021-129433A-C22 (HELENE) funded by MCIN/AEI/10.13039/501100011033 and the European Union NextGenerationEU/PRTR.

Bibliography

A. Pfefferling and P. Kehling. Design disclosure for blockchain-based application used in public education certificates with electronic hashes. *Hochschule Mittweida*, 2021.

A. Rustemi, F. Dalipi, V. Atanasovski, and A. Risteski. A systematic literature review on blockchain-based systems for academic certificate verification. *IEEE Access*, 11:64679–64696, 2023.

Flood:Github-Repository. Load testing tool for benchmarking EVM nodes over RPC, 2023. URL *https://github.com/paradigmxyz/flood*. [Online; accessed 4-September-2023].

J. Gresch, B. Rodrigues, E. John-Scheid, S. Kanhere, and B. Stiller. The proposal of a blockchain-based architecture for transparent certificate handling. *Springer International Publishing*, pages 185–196, 2018.

M. Sánchez-de la Rosa, C. Núñez-Gómez, M. Caminero, and C. Carrión. Exploring the use of blockchain in resource-constrained fog computing environments. *Software: Practice and Experience*, 53:971–987, 2023.

N. Lei, E. Masanet, and J. Koomey. Best practices for analyzing the direct energy use of blockchain technology systems: Review and policy recommendations. *Energy Policy*, 156: 112422, 2021.

T. M. Fernández-Caramés and P. Fraga-Lamas. Design of a fog computing, blockchain and IoT-based continuous glucose monitoring system for crowdsourcing mHealth. *Multidisciplinary Digital Publishing Institute Proceedings*, 4:5757, 2023.

Mapping the Poverty Proportion in Small Areas under Random Regression Coefficient Poisson Models

Naomi Diz-Rosales, María José Lombardía, and Domingo Morales

Centro de Investigación CITIC, Universidade da Coruña, 15071 A Coruña, Spain
Instituto Centro de Investigación Operativa IUCIO, Universitas Miguel Hernández, 03202 Elche, Spain
Correspondence: naomi.diz.rosales@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: In a complex socio-economic context, policy makers need highly disaggregated poverty indicators. In this work, we develop a methodology in small area estimation to derive predictors of poverty proportions under a random regression coefficient Poisson model, introducing bootstrap estimators of mean squared errors. Maximum likelihood estimators of model parameters and random effects mode predictors are calculated using a Laplace approximation algorithm. Simulation experiments are conducted to investigate the behaviour of the fitting algorithm, the predictors and the mean squared error estimator. The new statistical methodology is applied to data from the Spanish survey of living conditions to map poverty proportions by province and sex, developing a tool to support policy decision making.

1 Introduction

One out of every five people is at risk of poverty or social exclusion in the European Union (EU), with the figure rising to 26.4 % in Spain (Eurostat, 2022).

In order to reduce this number, the EU has set national targets for all its members. Among these measures, the SLC or Survey on living conditions, whose main objective is to provide comparable statistics on income distribution and social exclusion in order to support policy makers in the distribution of their policy packages, stands out. Thus, most European countries use the SLC to estimate poverty indicators. This survey in Spain, the Spanish survey of living conditions (SLCS), provides information on household income received during the year prior to the year of the interview, the domains foreseen being the Autonomous Communities (CCAA).

With these data, different tools can be developed, especially poverty maps, which are widely popularised as they provide a clear visual representation of the geographical distribution of poverty and the degree of inequality between territories in a country, becoming a key support for political decision-making. In fact, the estimation of these indicators and the use of the results for poverty mapping is of great international interest, initiated and sponsored in many cases by the United Nations and the World Bank, demanding their collection at increasingly lower levels of aggregation (see Molina and Rao (2010) for an exhaustive review).

This is a challenge in view of the small or even zero sample size, which can be assumed by Small Area Estimation (SAE) and which responds to the need to produce precise estimates on indicators of interest in areas or domains where the sample size is smaller than planned by the surveys from which the information on the target variable is extracted. Since its definition, with

the study by Fay III and Herriot (1979), research has followed one after the other, increasing significantly in recent years, driven by the Sustainable Development Goals (SDGs).

Thus, SAE methods are shown as potential alternatives, with the use of linear mixed models (LMMs) and generalised linear mixed models (GLMMs) standing out in recent years. However, all existing work so far in SAE considers the regression coefficients as fixed effects. This approach may be too rigid when the relationship between the target variable and the auxiliary variables is not constant across domains. This is common in the socio-economic context, covariates such as age, employment status, citizenship or education may have a different rate of influence on the poverty rate, depending on the region of study. In fact, we have as an example the Spanish case, where in terms of disparity between provinces, the different realities in the peninsular geography have been the subject of numerous investigations insinuating that the current structural inequalities were already established in the 1930s, with a pattern of increasing poverty from the north-east to the south-west of Spain (Tirado et al., 2016).

To provide more flexibility to the modelling process, Dempster et al. (1981), defined random slope mixed models for the first time in SAE. Despite an early initial development phase, with works such as those of Prasad and Rao (1990) to derive the empirical best linear unbiased predictor (EBLUP) and the analytical estimator of the mean squared error (MSE), research is limited, with the works of Hobza and Morales (2013) and Morales et al. (2021) standing out, and always in the context of LMMs. In fact, in SAE, to our current knowledge, GLMMs with random slope have not been defined.

Therefore, and due to the potential gain in flexibility of these approaches, in this work we present our current research, (Diz-Rosales et al., 2023, accepted for publication), aimed at developing new statistical methodology in SAE for poverty mapping by exploring the usefulness of area-level random slope Poisson GLMMs in the inference of poverty indicators. We introduce bootstrap estimators of the mean squared error. The maximum likelihood estimators of the parameters and the modal predictors of the random effects are calculated using a Laplace approximation algorithm. The behaviour of this fitting algorithm, as well as that of the predictors and the mean squared error estimators, is investigated by simulation experiments. Finally, the new statistical methodology is applied to the Spanish living conditions survey (SLCS) data. The objective is to estimate and map, by provinces, the proportions of women and men below the poverty threshold, thus developing a tool of social interest aimed at supporting social policy-making.

2 Data

The dataset selected for this study corresponds to the 2008 SLCS dataset, with a sample size of 35967. It should be noted that, despite its temporal distance with respect to the current year, the choice of this dataset lies in the fact that it has been widely used in other methodological approaches in SAE. In this way, it is possible to make comparisons of the results obtained using different procedures and, therefore, to make the necessary evaluations to improve the model for use with current data.

For the elaboration of the database, we constructed an aggregate data file with the 104 domains defined above. For each domain, the target variable of the Poisson model is the count of persons with an annual equivalised net incomes below a predetermined threshold established at 60 % of the median income per consumption unit, indicated in euros. The auxiliary variables are taken from the Spanish labour force survey (SLFS) of 2008, which provides information on the labour market participation of the population by relating it to characteristics of living conditions. Specifically, we consider the following categories where each category is the proportion of people who meet the defined condition:

- *Age*, with five categories: $\leqslant$15 years old (age0), 16 - 24 years old (age1), 25 - 49 years old (age2), 50 - 64 years old (age3), $\geqslant$ 65 years old (age4), with age0 as the reference category.

- *Education*, with four categories: $\leqslant$16 years old (edu0); illiterate persons with incomplete or complete primary education and/or lower secondary education (edu1); persons with complete secondary education and/or post-secondary education such as baccalaureate or vocational training (edu2); persons with university studies (edu3), with edu0 as the reference category.

- *Nationality*, with two categories: Spanish citizens including those with dual nationality (cit0); foreign citizenship (cit1), with cit0 as the reference category.

- *Labour status*, with four categories: $\leqslant$16 years (lab0), employed (lab1), unemployed (lab2), inactive (lab3), with lab0 as the reference category

For each factor at unit level, the sum of the proportions of its categories is one. Therefore, it is necessary to select a reference category and remove it from the dataset.

Finally, it is worth mentioning that, due to the socio-economic divergence across provinces (Tirado et al., 2016), we have created the *income group* variable, which classifies provinces into five categories ($K = 5$), $k = 1, \ldots, K$, based on an ascending order of the aggregate sum of the average income per household unit for men and women within each province. Thus, after an exploratory analysis of the data, we find that the relationship between the target variable and the covariates is different depending on the category k, and in the definition of the model, we introduce the random slopes by income groups of the provinces.

3 The area level random regression coefficient Poisson model

This section defines the basic principles of the methodology developed to define the area level random regression coefficient Poisson, which we refer to as the ARRCP model, and the derived predictors. A more in-depth development can be found in Diz-Rosales et al. (2023, accepted for publication).

Let us consider a count variable y_{ij} taking values on $\mathbb{N} \cup \{0\}$, where $i \in \mathbb{I} = \{1, \ldots, I\}$ and $j \in \mathbb{J} = \{1, \ldots, J\}$. Let $D = IJ$ be the total number of y-values. For example, y_{ij} could be the number of people living below the poverty line in a survey sample, the indexes i and j may represent province and sex, and D is the total number of domains defined by the crossings of the variables province and sex. In other words, we have a country partitioned into provinces and sexes. We further assume that each province can be grouped into one, and only one, of the K clusters, $\mathbb{I}_1, \ldots, \mathbb{I}_K$, of an income variable. Let $k(i)$ be the number of the category to which province i belongs, so that $k(i) \in \mathbb{K} = \{1, \ldots, K\}$. The number of provinces in the category $\mathbb{I}_K$ is $m_k = \#(\mathbb{I}_k)$, so that $D = J \sum_{k=1}^{K} m_k$.

We are dealing with area-level data for modelling and predicting the target variable y_{ij}. Let us assume that we have p explanatory variables with values $x_{\ell,ij}$, $\ell \in \mathbb{P} = \{1, \ldots, p\}$, $i \in \mathbb{I}$, $j \in \mathbb{J}$. For models with intercept, we take $x_{0,ij} = 1$ for all i and j. In what follows, we present the ARRCP model.

Let u_{ij}, $i \in \mathbb{I}$, $j \in \mathbb{J}$ be i.i.d. $N(0, 1)$ random variables. Let $\phi_\ell > 0$, $\ell \in \mathbb{P}$, be unknown standard deviation parameters. Let $\rho_{rs} \in (-1, 1)$, $r < s, r, s \in \mathbb{P}$, be unknown correlation parameters. Let $\boldsymbol{v}_k = (v_{1,k}, \ldots, v_{p,k})'$, $k \in \mathbb{K}$, be i.i.d. random vectors such that

$$\underset{1\leqslant\ell\leqslant p}{\text{diag}}\,(\phi_\ell)\boldsymbol{v}_k \sim N_p(\boldsymbol{0}, \boldsymbol{V}_{vk}^{\phi\rho}), \quad \boldsymbol{V}_{vk}^{\phi\rho} = \begin{pmatrix} \phi_1^2 & \phi_1\phi_2\rho_{12} & \cdots & \phi_1\phi_p\rho_{1p} \\ \phi_2\phi_1\rho_{12} & \phi_2^2 & \cdots & \phi_2\phi_p\rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_p\phi_1\rho_{1p} & \phi_p\phi_2\rho_{2p} & \cdots & \phi_p^2 \end{pmatrix}.$$

Therefore, the variance of $\boldsymbol{v}_k$ is

$$\boldsymbol{V}_{vk} = \text{var}(\boldsymbol{v}_k) = \underset{1\leqslant \ell\leqslant p}{\text{diag}}\,(\phi_\ell^{-1})\boldsymbol{V}_{vk}^{\phi\rho}\underset{1\leqslant \ell\leqslant p}{\text{diag}}\,(\phi_\ell^{-1}) = \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{12} & 1 & \cdots & \rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1p} & \rho_{2p} & \cdots & 1 \end{pmatrix}.$$

Let us define the vectors

$$\boldsymbol{u} = \underset{1\leqslant i\leqslant I}{\text{col}}(\underset{1\leqslant j\leqslant J}{\text{col}}(u_{ij})) \sim N_{IJ}(\boldsymbol{0}, \boldsymbol{I}_{IJ}), \quad \boldsymbol{v} = \underset{1\leqslant k\leqslant K}{\text{col}}(\boldsymbol{v}_k) \sim N_{pK}(\boldsymbol{0}, \boldsymbol{V}_v),$$

where $\boldsymbol{V}_v = \underset{1\leqslant k\leqslant K}{\text{diag}}\,(\boldsymbol{V}_{vk})$ and where diag and col are the diagonal and the column operator, respectively. We assume that $\boldsymbol{u}$ and $\boldsymbol{v}$ are independent. The distribution of the target variable y_{ij}, conditioned to the random effects u_{ij}, $v_{\ell,k(i)}$, $\ell \in \mathbb{P}$, is

$$y_{ij}|_{u_{ij},v_{1,k(i)},\ldots,v_{p,k(i)}} \sim \text{Poisson}(\nu_{ij}p_{ij}), \quad i \in \mathbb{I},\ j \in \mathbb{J},$$

where the offset (or size) parameters $\nu_{ij} > 0$ are known and correspond to the sample size when the model is applied to real data, and the binomial probability, p_{ij}, is the target parameter with range (0,1). For the natural parameters, we assume

$$\eta_{ij} = \log\mu_{ij} = \log\nu_{ij} + \log p_{ij} = \log\nu_{ij} + \sum_{\ell=1}^{p}\beta_\ell x_{\ell,ij} + \sigma u_{ij} + \sum_{\ell=1}^{p}\phi_\ell v_{\ell,k(i)}x_{\ell,ij}, \quad i \in \mathbb{I},\ j \in \mathbb{J}, \tag{18.1}$$

where $\mu_{ij} = E[y_{ij}|u_{ij}, v_{1,k(i)}, \ldots, v_{p,k(i)}]$. We may write $\boldsymbol{x}_{ij}\boldsymbol{\beta} = \sum_{\ell=1}^{p}\beta_\ell x_{\ell,ij}$, where $\boldsymbol{\beta} = \underset{1\leqslant \ell\leqslant p}{\text{col}}\,(\beta_\ell)$ is the column vector of regression parameters and $\boldsymbol{x}_{ij} = \underset{1\leqslant \ell\leqslant p}{\text{col}'}\,(x_{\ell,ij})$ is the row vector of known auxiliary variables. To finish the definition of the ARRCP model, we assume that the y_{ij}'s are independent conditioned to $\boldsymbol{u}$ and $\boldsymbol{v}$. The variance component parameters are $\sigma > 0$, $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_p)' \in \mathbb{R}_+^p$ and $\boldsymbol{\rho} = (\rho_{12}, \ldots, \rho_{1p}, \ldots, \rho_{p-1p})' \in (-1,1)^{p(p-1)/2}$, where $\mathbb{R}_+ = (0,\infty)$. The vector of model parameters is $\boldsymbol{\theta} = (\boldsymbol{\beta}', \sigma, \boldsymbol{\phi}', \boldsymbol{\rho}')'$. The total number of random effects is $H = IJ + pK$.

With the ARRCP model defined, we proceed to carry out the maximization, deriving the maximum likelihood estimators of the model parameters, $\hat{\boldsymbol{\beta}}$, $\hat{\sigma}$, $\hat{\boldsymbol{\phi}}$, $\hat{\boldsymbol{\rho}}$, and the mode predictors of the random effects, by means of a Laplace approximation algorithm (ML Laplace algorithm), using the R package lme4 1.1-33.

In addition, we define the best predictor (BP), the simplified best predictor (SP), the empirical best predictor (EBP), the empirical simplified best predictor (ESP) and the plug-in predictor (IN), to predict the porverty proportion by province and sex. The best predictor (BP) of p_{ij} is

$$\hat{p}_{ij}^{bp} = \hat{p}_{ij}^{bp}(\boldsymbol{\theta}) = E_\theta[p_{ij}|\boldsymbol{y}] = E_\theta[p_{ij}|\boldsymbol{y}_{jk(i)}]$$

The simplified best predictor (SP) of p_{ij} is

$$\hat{p}_{ij}^{sp} = \hat{p}_{ij}^{sp}(\boldsymbol{\theta}) = E_\theta[p_{ij}|\boldsymbol{y}] = E_\theta[p_{ij}|\boldsymbol{y}_{ij}]$$

The empirical best predictor (EBP) of p_{ij} is $\hat{p}_{ij}^{ebp} = \hat{p}_{ij}^{bp}(\hat{\boldsymbol{\theta}})$, and the empirical simplified best predictor (ESP) of p_{ij} is $\hat{p}_{ij}^{esp} = \hat{p}_{ij}^{sp}(\hat{\boldsymbol{\theta}})$. These predictor requires approximating a multivariate integral which we approximate by a Monte Carlo method.

The behaviour of these predictors is then evaluated by bootstrap simulation studies of at least 1000 iterations, comparing them with other predictors, such as the plug-in predictor (IN) of p_{ij} and μ_{ij} under the ARRCP model which has the form

$$\hat{p}_{ij}^{in} = \exp\{\boldsymbol{x}_{itj}\hat{\boldsymbol{\beta}} + \hat{\sigma}\hat{u}_{ij} + \sum_{\ell=1}^{p} \hat{\phi}_\ell \hat{v}_{\ell,k(i)} x_{\ell,ij}\},$$

where $\hat{u}_{ij}$ and $\hat{v}_{\ell,k(i)}$, $\ell = 1, \ldots, p$, are the mode predictors taken from the output of the ML Laplace algorithm, being $\hat{\mu}_{ij}^{in} = \nu_{ij}\hat{p}_{ij}^{in}$ the IN predictor of $\mu_{ij} = \nu_{ij}p_{ij}$.

As a result of the simulations, the IN predictor shows the best computational efficiency and accuracy trade-off. To understand these results, it should be noted that the estimation of the BP and EBP is a very complex multivariate integral approximation problem. In order to perform this Monte Carlo approximation, we have generated simulations with different configurations in the number of replications, until we have tested, for the moment, the approximation with the generation of 2500 independent random variables. However, as we can see from the results, this number is insufficient. The BP and EBP incorporate the Monte Carlo variance, which is high, and which has a variance underlying the estimation of the multivariate integral, which would require substantially more than 2500 for an optimal approximation. By contrast, these simulations are highly computationally expensive, increasing simulation times to the order of days.

On the other hand, SP and ESP also experience this problem, although to a lesser extent, since the integral to be approximated is considerably less complex than BP and EBP. The results in terms of bias are worse than those of the BP, EBP and IN predictors, although the estimation of the root mean squared error (RMSE) is notably better, being closer to the IN than the rest. This, together with an efficient computational time, makes it a candidate for use when few computational resources are available to simulate the EBP and/or the IN predictor cannot be used. Although, at the expense of obtaining more efficient results in terms of computational performance, and having verified the high performance of the IN predictor, we chose it as the starting predictor for this study.

Consequently, a second simulation study is performed to test the performance of the MSE estimator of this predictor based on the parametric bootstrap with and without bias correction. While the performance of the MSE estimators of both approaches was optimal, due to the substantial improvement in bias and the low computational cost, with virtually unbiased estimation, in the application to real data we use the parametric bootstrap estimation with bias correction.

4 Application to real data

During the model selection phase, we consider several criteria: a) significance of model parameters and socio-economic interpretability; b) convergence of the ML-Laplace approximation algorithm; c) validity of model assumptions; and d) lower conditional AIC.

As a result of the selection process, we define the ARRCP model, introduced in 18.1 in Section 3, with the following variables: y_{ij} is the sample count of people below the poverty threshold in province i and sex j, $\nu_{ij} = n_{ij}$ is the sample size, $x_{0,ij}$ is the intercept and $x_{1,ij}$, $x_{2,ij}$, $x_{3,ij}$ and $x_{4,ij}$ are the values of the auxiliary variables age3, edu1, cit1 and lab2 respectively. The selected model contains two random slopes for $x_{1,ij}$ and $x_{4,ij}$, so that the corresponding model parameters and random effects are $\phi_1, \phi_4, \rho_{14}, u_{ij} \sim N(0,1)$, $(v_{1,k}, v_{4,k})' \sim N_2(\mathbf{0}, \boldsymbol{V}_{14})$, $\mathbf{0} = (0,0)'$ and $\boldsymbol{V}_{14}(\rho_{14}) = \begin{pmatrix} 1 & \rho_{14} \\ \rho_{14} & 1 \end{pmatrix}$. The natural parameter is

$$\eta_{ij} = \log \mu_{ij} = \log n_{ij} + \sum_{\ell=1}^{5} \beta_\ell x_{\ell,ij} + \sigma u_{ij} + \phi_1 v_{1,k(i)} x_{1,ij} + \phi_4 v_{4,k(i)} x_{4,ij}.$$

The estimated model parameters are socioeconomically interpretable, with the auxiliary variables edu1 and lab2 having a protective effect on poverty, and age3 and cit1 helping to reduce it. In addition, we obtained the basic percentile bootstrap confidence intervals observing that, at 95 %, all are significant.

Once the model is selected, it undergoes the diagnostic phase starting with the evaluation of the Pearson residuals.

Having obtained good results, we proceed to further assess the performance in estimating the poverty ratio by province and sex. For this purpose, in Figure 1 (left) we plot the IN predictions and the classic Hájek estimates of the poverty proportions. This figure compares both types of estimators and analyses the effect of using this type of estimator in unplanned domains.

We can see that the direct estimator and the IN predictor diverge noticeably in domains with lower sample sizes, becoming closer as sample size increases. This trend is consistent in the RMSE estimate plotted on the right, where the error magnitude of the direct and IN estimator tend to decrease and equalise as the sample size increases. However, it is notable the smooth and decreasing behaviour of the RMSE of the IN predictor while in the case of the direct estimator the RMSE estimates are characterised by an abrupt trend with characteristic peaks.

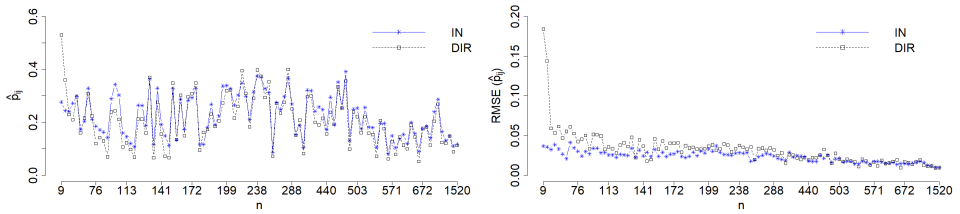


Figure 1: Poverty proportions estimates $\hat{p}_{ii}$ and associated $RMSE_{ii}$, ordered by sample size

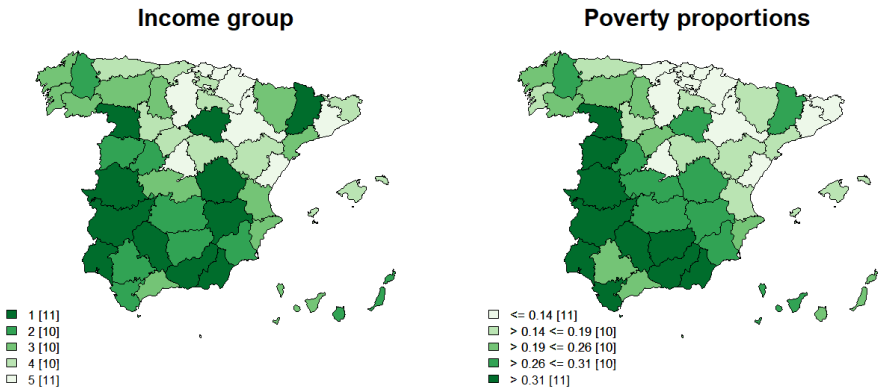


Figure 2: Estimated poverty proportions in Spanish provinces by income group

In order to improve visual comprehension, we represent the estimates of poverty proportions on maps widely used by socio-political powers. In particular, we illustrate in Figure 2 an example of poverty maps at the provincial level and compare it with a poverty map produced taking into account the existing K categories in the data. As can be seen, the intensity of colours is generally the same, while in those that diverge, they do so between contiguous groups, without a substantial difference, as in the provinces of Cádiz or Salamanca.

These maps allow us to evaluate at a glance the socio-economic state of the country, with a substantial percentage of areas with high levels of poverty, not being equally distributed, with a clear north-south, east-west pattern of increasing poverty rates. In addition, we have carried out a study of differences in poverty rates by sex using bootstrap basic percentile confidence interval, and although, at the 95 % and 90 % percentile, significant differences were only detected in 6 provinces, in line with other studies in the field, in all cases the poverty rate was higher among women.

We would like to complement the study with more indicators to overcome limitations such as the fact that the cost of living is not the same in all regions of the country, but the poverty line is the same. This does not detract from the conclusions, but it could help to have a more global vision in decision making.

Acknowledgments

This research is part of the grant PID2020-113578RB-I00, funded by MCIN/AEI/10.13039/501100011033/. It has also been supported by the Spanish grant PID2022-136878NB-I00, the Valencian grant Prometeo/2021/063, by the Xunta de Galicia (Competitive Reference ED431C-2020/14) and by CITIC that is supported by Xunta de Galicia, collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Sistema Universitario de Galicia (CIGUS). The first author was also sponsored by the Spanish Grant for Predoctoral Research Trainees RD 103/2019 being this work part of grant PRE2021-100857, funded by MCIN/AEI/10.13039/501100011033/ and ESF+.

Bibliography

A. P. Dempster, D. B. Rubin, and R. K. Tsutakawa. Estimation in covariance components models. *Journal of the American Statistical Association*, 76(374):341–353, 1981.

N. Diz-Rosales, M. J. Lombardía, and D. Morales. Poverty mapping under area-level random regression coefficient poisson models. *Journal of Survey Statistics and Methodology*, 2023, accepted for publication.

Eurostat. Persons at risk of poverty or social exclusion by age and sex - eu 2020 strategy. *https://ec.europa.eu/eurostat/databrowser/view/ilc_peps01/default/table?lang=en*, 2022. [Online; accessed 4-June-2022].

R. E. Fay III and R. A. Herriot. Estimates of income for small places: an application of james-stein procedures to census data. *Journal of the American Statistical Association*, 74(366a):269–277, 1979.

T. Hobza and D. Morales. Small area estimation under random regression coefficient models. *Journal of Statistical Computation and Simulation*, 83(11):2160–2177, 2013.

I. Molina and J. Rao. Small area estimation of poverty indicators. *Canadian Journal of statistics*, 38(3):369–385, 2010.

D. Morales, M. D. Esteban, A. Pérez, and T. Hobza. *A course on small area estimation and mixed models*. Springer, Switzerland, 2021.

N. N. Prasad and J. N. Rao. The estimation of the mean squared error of small-area estimators. *Journal of the American statistical association*, 85(409):163–171, 1990.

D. A. Tirado, A. Díez-Minguela, and J. Martinez-Galarraga. Regional inequality and economic development in spain, 1860–2010. *Journal of Historical Geography*, 54:87–98, 2016.

Efficient Database Evolution in Digital Library Reengineering

Delfina Ramos-Vidal, and Nieves R. Brisaboa

Database Laboratory, Faculty of Computer Science, Universidade da Coruña, 15071 A Coruña, Spain
Centro de Investigación CITIC, Universidade da Coruña, 15071 A Coruña, Spain
Correspondence: delfina.ramos@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: With the advancement of internet applications, extensive information systems were created to effectively manage and provide easy access to documents, which coincided with a global initiative to convert physical documents into digital format, making them accessible through the internet. After two decades, these databases are well-structured and organized, although the software used to manage them is gradually becoming outdated. Additionally, once the initial digitization and creation of metadata are completed, it is sensible to enhance the metadata further to provide more detailed information about the documents. In this article we propose a tool to facilitate the evolution of large documentary databases.

1 Introduction

The significant effort to make information public has resulted in huge documentary databases accessible worldwide through digital libraries. Currently, many of these digital library systems or documentary database management systems (documentary DBs) have become technologically obsolete. When considering the reengineering of these systems, apart from updating their software and providing them with new functionalities, there is also a need for the migration of documentary databases. Since the original effort to digitize documents has already been carried out, it is common for the migration process to involve an increase in digitized assets and metadata information to expand the database.

The aim of this article is to propose a tool that facilitates the migration of legacy documentary databases to new data models, automating the process as much as possible in order to save time, both in analysis and implementation.

Over the past twenty years, our research group has developed various digital libraries, such as the Galician Virtual Library[1], the Galician Edition during the Franco era[2], or the Hemeroteca of the Royal Galician Academy. All these platforms require their systems to be updated technologically and to take advantage of this update to provide them with new functionalities. The issue arises from the need to preserve current data while modifying the database schema or integrating new data sources.

Currently, this migration process is based on SQL scripts that import complete tables from the original database, Excel spreadsheets where users can verify and complete data extracted from legacy tables, or CSV files obtained from different external data sources that are intended to be linked with existing data. This work is slow and error-prone since it is not a direct process

[1] Galician Virtual Library: *https://bvg.udc.es/*
[2] Galician Edition during the Franco era: *https://ediciongalizafranquista.udc.gal/*

where the script for data importations must allow domain experts to review original data, which often contain errors after decades of use, fulfil tables with new data they want to introduce into the new database, and make decisions regarding which data to migrate in each case. Based on the experience of migrating these use cases and generalising the real problems and needs observed, we have designed a tool to automate the migration process.

2 State of the art

In the field of databases, issues related to data update, integration, and loading have already been addressed through Extraction, Transformation, and Loading (ETL) tools. These tools were originally designed to move data from operational transactional databases to data warehouses but have been adopted in various other contexts, such as database migration due to system updates. Our initial approach was to investigate whether a standard ETL system could be suitable for our purposes.

Currently, there are ETL solutions available in the market, such as Talend, Pentaho, or Google Cloud Dataflow, among others (Sreemathy et al. (2021)). All of these tools automate most of an organization's workflow without the need for human interaction, providing a highly reliable service, and supporting both on-premise databases and cloud databases. However, the transformations supported by these tools are typically quite generic and limited, including operations like row and column transposition, table joining and splitting, data sorting and filtering, but they do not allow for the definition of domain-specific rules.

Additionally, once a transformation is executed, the user does not have decision-making power to handle errors that may arise, necessitating a subsequent data review step for error correction. In previous work by Ramos Vidal et al. (2022), we presented an initial approach to automating the data transformation process for legacy database migration using Domain-Specific Languages.

3 Our proposal

Figure 1 depicts the architecture of the database migration tool we propose. The primary component of our solution is what we call the "Migration and Collation Tool." Within the tool, three phases can be executed independently. On one hand, the "Model Definition Interface" allows for defining the data model for the data originating from both the Source Database and External Data Sources, as well as the data model for the Destination Database. In this step, the "Migration Rules" specific to the business domain are also defined, which will be checked during the conversion process, and the data model that data requiring expert review should adopt. Additionally, the tool includes a "Data Importer" that enables establishing a connection with both the Source Database and External Data Sources outside the system.

The imported data is processed in the "Converter" component, where data from all sources is evaluated, and necessary transformations are performed based on the migration rules defined earlier. Once converted to the new data model, they are stored in the Destination Database. Data for which the transformation fails or is ambiguous according to the migration rules will be placed in the Review Database. Finally, the records in the Review Database will be displayed in the "Debugging Tool," following the data model for review defined in the previous stage, allowing the user to decide which data to retain in case of ambiguity or what corrective action to take in case of errors before moving them to the Destination Database.

The migration for "Edición en la Galicia franquista" will serve as a running example[3]. The original database (A) contains information about authors with literary production during the Franco era (1936-1975) in Galicia. The goal is to expand the library's corpus with data from an external source (B) covering the period from 1936 to 2000. To automate the migration, we

[3] Edición en la Galicia franquista: *https://ediciongalizafranquista.udc.gal/*

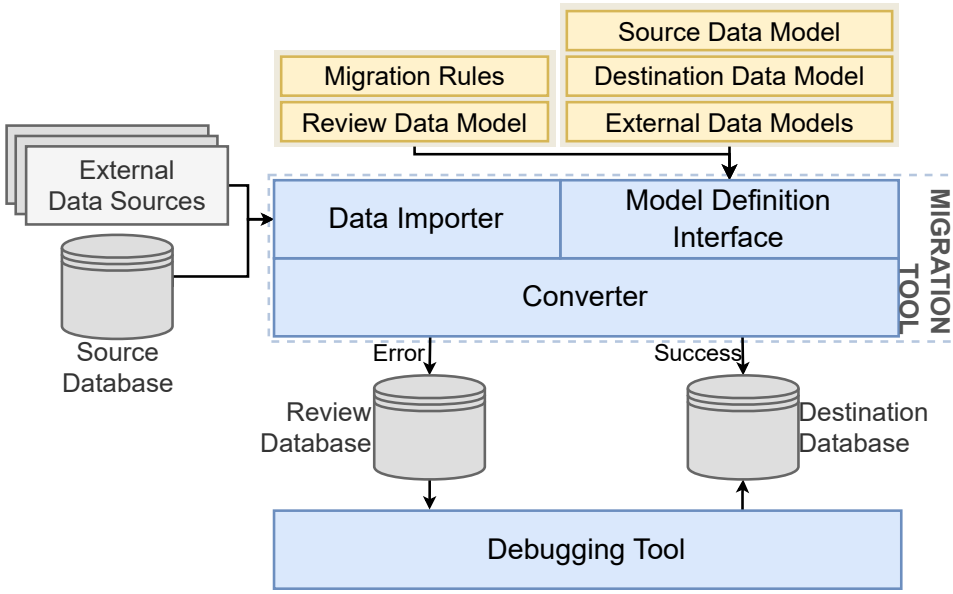


Figure 1: Architecture of the Migration Tool

analyse the possible scenarios that could arise when combining the data sources.

To begin with, if an author is only present in A, the data is automatically migrated to the destination database (C). The same applies when an author has production only after 1975, i.e., they are only present in B. However, in the case of authors with production both before and after 1975, their data will be present in both data sources, so we must evaluate potential cases of inconsistency. We assume that authors in table B are new additions, but a check on table A reveals that there is already another author with the same name, so the row must be sent for review for an expert to verify if they are the same person or not. On the other hand, if the author is present in both tables, we check that all attributes (name, surname, alias, date of birth, date of death, observations) match. It could happen that errors were made when entering authors' names, or an author passed away in recent years, causing a mismatch in the date of death since it doesn't exist in table A. These checks can be defined using rules like:

```
IF (NombreA=NombreB AND AliasA=AliasB)
    THEN (Insert IN C)
ELSE IF (NombreA!=NombreB)
    THEN (Error 1)
ELSE IF (NombreA=NombreB AND AliasA!=AliasB)
    THEN (Error 2)
```

During the conversion process, the attributes defined in the rule are evaluated, and in case of inconsistencies, they are sent to the Review Database for an expert user in the domain to decide which data should be retained or corrected in each case. Once the issues are processed, this data is moved to "C," concluding the migration process.

At the current moment, we are finishing the design phase of the proposal. The next step is to begin the implementation and validation phase, which will include both software-specific verifications and validations, as well as a set of tests with real users in a simulated real-world context.

Acknowledgements

CITIC is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS), and by PRE2021-099351, MCIN/AEI+"FSE+";GRC[ED431C 2021/53]: GAIN/Xunta de Galicia; TED2021-129245B-C21(PLAGEMIS): MCIN/AEI+"NextGenerationEU"/PRTR; PID2020-114635RB-I00(EXTRACompact): MCIN/AEI; PID2021-122554OB-C33 (OASSIS): MCIN/AEI+EU/ERDF A way of making Europe; PDC2021-120917-C21 (SIGTRANS): MCIN/AEI+"NextGenerationEU"/PRTR

Bibliography

D. Ramos Vidal, A. Cortiñas, M. R. Luaces, O. Pedreira, and A. S. Places. Dsl para la migración de bases de datos legacy en el marco de una lps. In *Actas de las XXVI Jornadas de Ingeniería Del Software y Bases de Datos* (*JISBD 2022*), Santiago de Compostela, 2022.

J. Sreemathy, R. Brindha, M. Selva Nagalakshmi, N. Suvekha, N. Karthick Ragul, and M. Praveennandha. Overview of etl tools and talend-data integration. In *2021 7th International Conference on Advanced Computing and Communication Systems* (*ICACCS*), volume 1, pages 1650–1654, 2021. doi: 10.1109/ICACCS51430.2021.9441984.

Analysis of the NS-3 module QKDNetSim for the Simulation of QKD Networks

David Soler, Iván Cillero, Francisco J. Nóvoa, Carlos Dafonte, Ana Fernández Vilas, and Manuel Fernández-Veiga

Laboratorio Interdisciplinar de Aplicaciones de Inteligencia Artificial, Faculty of Computer Science, Universidade da Coruña, 15071 A Coruña, Spain
Laboratorio de Telemática, Faculty of Computer Science, Universidade da Coruña, 15071 A Coruña, Spain
I&CLabs, *AtlanTTic*, University of Vigo, 36310 Vigo, Spain
Correspondence: david.soler@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: Quantum Key Distribution (QKD) is a promising technology that allows two nodes to privately agree on a key through a quantum channel. Unfortunately, QKD is still in experimental phase and researchers must rely on simulators to replicate the behaviour of a quantum network. One of the most widespread is *QKDNetSim*, a module for the C++ network simulator *NS-3*. However, this module is very limited in its behaviour, so it does not faithfully represent a real quantum network. In this work we analyse the structure and components of *QKDNetSim*, as well as its shortcomings and how they affect the quality of the simulation.

1 Introduction

Quantum technologies have experienced a significant advance in recent years Cao et al. (2022); Illiano et al. (2022); Singh et al. (2021). The increasing computing power of quantum computers is endangering the cryptographic algorithms that are employed to distribute keys between users, including protocols as widespread as HTTPS. Nevertheless, quantum technologies also provide an alternative to these algorithms: the Quantum Key Distribution (QKD) protocols allow two nodes to agree on a key through a quantum channel, in such a way that it would be impossible for an eavesdropper to obtain the key without being detected. This key can then be used to encrypt communications between the two nodes.

The number of implemented QKD networks is currently very small due to the high cost of the required material and the lack of maturity in the technology. Thus, researchers must employ simulators that imitate the behavior of a quantum network. There are multiple alternatives depending on the scope of the research Aji et al. (2021): some simulators focus on representing the physical layer of the quantum channel, while others allow users to define entire networks in which nodes can execute QKD between them. The simulators "Qunetsim" and "NetSquid", both written for Python, are among the most popular options.

The network simulator NS-3 is widely used in the scientific and educational communities due to its level of detail and its customizing capabilities. There exists a module implemented for NS-3 for the simulation of quantum networks, named QKDNetSim Mehic et al. (2017), that was developed by researchers of the Technical University of Ostrava.

The advantages of QKDNetSim over other simulators come from the granularity of NS-3: this simulator allows for in-depth configuration of every component, and packets sent over the simulated network are fully defined, including headers for all protocols involved. This level of

detail, when applied to QKDNetSim's simulation of quantum networks, could help newcomers to understand the fundamentals of quantum communications. However, this module contains numerous conceptual and implementation flaws that affect its ability to faithfully represent the behavior of a real quantum network.

In this work we evaluate the quantum network simulation module QKDNetSim: of which elements it is composed, what procedures are followed to simulate the key exchange and the encryption of information, and to which degree it imitates the behavior of a real quantum network. We will also analyze the limitations and errors that prevent QKDNetSim from achieving its purposed objectives.

The rest of this work is organised as follows: in Section 2 we introduce readers to the basics of Quantum Key Distribution, which will be required for understanding the components and behaviour of the simulator. In Section 3 we analyse the module QKDNetSim. First, we introduce the components of a node connected to a quantum channel. Then, we comment on QKDNetSim's errors and how they affect its usefulness. Finally, Section 4 will conclude this document and discuss future improvements.

2 Background

2.1 Quantum key distribution

Quantum channels are communication channels that are analogous to classical channels, but transmit *qubits* instead of bits. Qubits represent a superposition of their base states $|0\rangle$ and $|1\rangle$ in the following way:

$$|\Phi\rangle = \alpha|0\rangle + \beta|1\rangle$$

where α and β represent the relative probability of measuring each of the base states.

Unlike classical bits, qubits possess two properties that make them specially suited for transmitting private information: the **no-cloning theorem** (it is impossible to perfectly clone a quantum state) and the **uncertainty principle** (whenever a state is measured, it collapses into a base state and loses its superposition). Taken together, these two properties ensure that any attacker trying to eavesdrop a quantum channel will inevitably modify the qubits that are being transmitted, such that the legitimate interlocutors will be aware of the intrusion.

Since quantum channels are capable of transmitting information in a private manner, they can be used to establish common cryptographic material between two parties. This is called Quantum Key Distribution (QKD), and it is believed to be information-theoretically secure, that is, it would be impossible for an attacker to obtain the secret no matter how many resources they have access to. The shared secret can then be used in any classical cryptographic algorithm that is considered secure, like AES.

Figure 1 shows the structure of a QKD communication: Peers Alice and Bob are connected through both classical and quantum channels. They must first establish a common secret executing a QKD protocol like BB84 Bennett and Brassard (1984). If they detect any eavesdropping, they abort the protocol. Then, Alice uses this common secret as a symmetric key to encrypt a message which she transmits to Bob. Finally, he decrypts the message.

As mentioned, the main advantage of QKD over other key distribution protocols is that QKD is information-theoretically secure. While classical algorithms like Diffie-Hellman or RSA base their security on the complexity of concrete mathematical problems, QKD's security is guaranteed by the laws of physics.

2.2 QKD network architecture

In experimental settings, the execution of QKD is very slow and thus it would be very inefficient to block the generation of a packet to obtain key material through QKD Mehic et al. (2020). For that reason, it is common to asynchronously execute the QKD protocol at a previous time, and

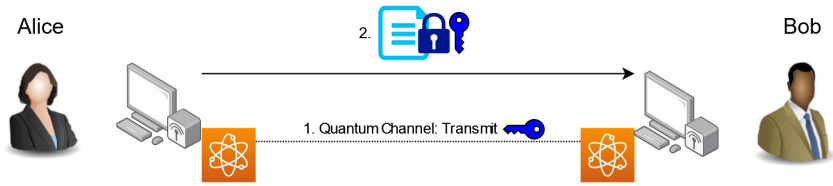


Figure 1: Diagram of a QKD communication.

store the resulting key material in a buffer to be consumed when needed Cao et al. (2019). Thus, the processes of generating and consuming key material are decoupled.

However, this introduces the need to synchronise the buffers of adjacent nodes. If Alice encrypts a packet using material from her buffer, Bob should be able to obtain the exact same key from his. If an error occurs and the buffers are desynchronised, the obtained keys would not match and Bob would not be able to decrypt Alice's messages.

3 Analysis of QKDNetSim

3.1 NS-3

NS-3 is a network simulator written in C++ that is widely used by students and academics. It provides an environment for the simulation of events in a computer network, as well as classes that can generate said events. It is open source, and there are multiple modules developed by the community, including QKDNetSim, the object of study of this work.

NS-3 provides classes that represent certain components of real-world networks. *Nodes* represent the computers inside the network, and they are connected to each other through *NetDevices*. *Applications* simulate the execution of programs inside Nodes, and they generate *Packets*. *Helpers* are classes that simplify the process of modifying Nodes, usually installing a NetDevice or Application.

The user can determine the duration of the simulation, as well as the timing of every event (when the Applications start sending Packets and how often). When it ends, it is possible to generate a PCAP file listing all packets to analyse the simulation.

3.2 Structure of a QKDNetSim node

The module QKDNetSim adds to NS-3 the possibility of creating quantum channels between pairs of nodes. To this end, the following components are added to each pair of nodes:

- **Send/Receive Applications**: they simulate programs that create packets to be encrypted by the rest of QKDNetSim's components. When they are generated or received, they are sent to the Manager to perform the pertinent cryptographic operations.
- **QKD Manager**: the central component of the module, which serves as connection between the others. The Manager processes incoming and outgoing packets, identifies which operations need to be performed and calls the pertinent component to execute them.
- **Cryptography Handler**: receives petitions from the Manager to encrypt or decrypt packets. It implements multiple cryptographic algorithms, and has access to the Key Buffer.
- **Simulated Quantum Channel**: imitates the behaviour of a quantum channel. Each of the nodes possesses a *Charging Application*, which constantly generate new shared key material.
- **Key Buffer**: Stores key material transmitted through the quantum channel for future use.

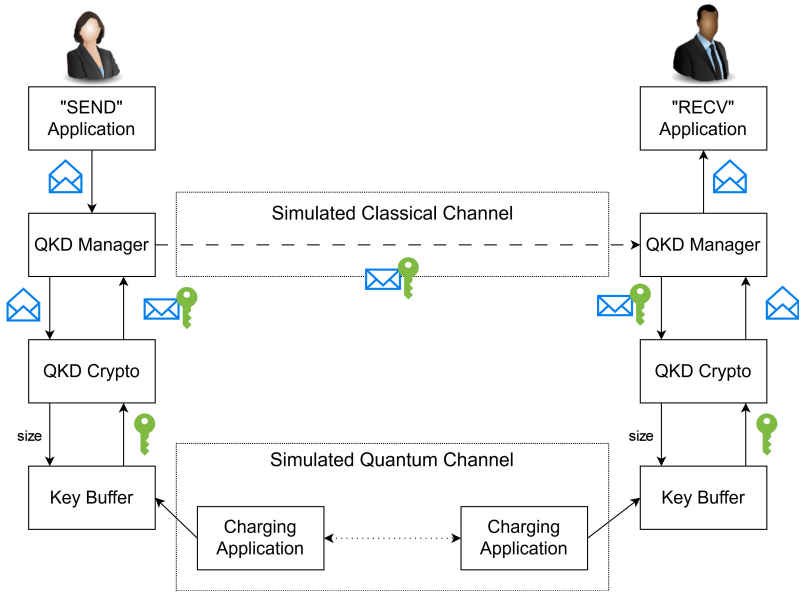


Figure 2: Architecture of QKDNetSim.

```
> Frame 1964: 462 bytes on wire (3696 bits), 462 bytes captured (3696 bits)
  Raw packet data
> Internet Protocol Version 4, Src: 10.1.1.1, Dst: 10.1.1.2
> Transmission Control Protocol, Src Port: 49153, Dst Port: 8000, Seq: 261000, Ack: 1, Len: 410
> Data (410 bytes)

0000  45 00 01 ce 03 c2 00 00  40 06 00 00 0a 01 01 01   E······ @·····
0010  0a 01 01 02 c0 01 1f 40  00 03 fb 88 00 00 00 01   ·····@ ·····
0020  80 10 80 00 00 00 00 00  08 0a 00 00 3c 3c 00 00   ·····  ····<<··
0030  3c 1e 00 00 41 44 44 4b  45 59 3a 33 30 30 30 30   <···ADDK EY:30000
0040  30 30 3b 30 30 30 30 30  30 30 30 30 30 30 30 30   00;00000 00000000
0050  30 30 30 30 30 30 30 30  30 30 30 30 30 30 30 30   00000000 00000000
0060  30 30 30 30 30 30 30 30  30 30 30 30 30 30 30 30   00000000 00000000
```

Figure 3: Packet containing key material sent by Charging Application.

The interaction between elements is shown in Figure 2.

3.3 Simulation of a Quantum Channel

As mentioned, the quantum channel is simulated through the *Charging Applications* that are installed for each pair of connected nodes. Figure 3 shows the contents of a packet exchanged between the Charging Applications of two adjacent nodes, which includes (after the label *ADDKEY*) the key material that will be added to the buffers.

QKDNetSim uses buffers to store key material, as explained in Section 2.2. the Charging Applications are constantly generating new key material and storing it into their respective Key Buffers. Whenever the Send Application generates a new packet, some of the previously generated key material is consumed to encrypt it.

Figure 4 shows the amount of key material stored inside a Key Buffer during a simulation. It increases periodically when the Charging Applications creates new key material, and it decreases when new packets are encrypted.

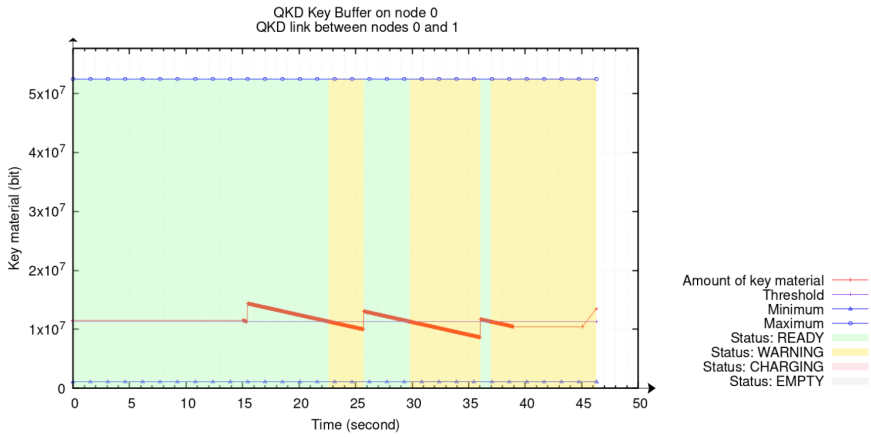


Figure 4: Amount of available key material inside the Key Buffer throughout the simulation.

3.4 Shortcomings

The objective of QKDNetSim is to faithfully represent the behaviour of a real quantum network. However, some of its components have been simplified to a point that they do not longer work as their real counterparts would. Instead of making the simulation easier to understand, these simplifications limit the usefulness of QKDNetSim. Furthermore, there are implementation errors that render some components unusable.

We highlight the following shortcomings of QKDNetSim:

1. The messages sent through the Simulated Quantum Channel are all identical. The Simulated Quantum Channel is tasked with the creation of new key material between two nodes. However, all messages exchanged between the Charging Applications of adjacent nodes contain only a string of '0's, as can be seen in Figure 3. Since all the key material that is created is the same, all packets are encrypted with the same key. Thus, in QKDNetSim the nodes are not actually executing a Key Distribution protocol, since they already know that the key is a string of '0's. This defeats the purpose of QKD and makes QKDNetSim a much less realistic simulator.

2. The Key Buffers do not store keys. The key material generated through the Simulated Quantum Channel is not stored anywhere. This should be the function of the Key Buffer: instead, it only stores a number which represents the *amount* of key material it should have.

Whenever the Simulated Quantum Channel tries to insert new key material in the Key Buffer, it is discarded and the *amount* is increased. A similar process occurs when the Manager requests key material from the Buffer: a new string of '0's is returned and the *amount* is decreased by its length. The plot shown in Figure 4 represents this *amount*, since is the only metric the Buffer can provide.

Since there is no real key material inside the Buffers, there is also no mechanism for ensuring, maintaining or recovering synchronicity between Key Buffers of connected nodes. This makes QKDNetSim less realistic, as it does not represent a problem that needs to be addressed in real quantum networks.

3. Encryption is disabled and does not work. The Cryptography Handler provides implementations for the encryption algorithms AES and One-Time Pad (OTP). However, these

implementations contain errors: they incorrectly assume keys are represented as an array of bytes, when they are an array of bits. Thus, whenever the simulator tries to encrypt a packet, it crashes.

Even if these errors are corrected and packets are adequately encrypted, the PCAP still shows them in plaintext. This is due to the order in which QKDNetSim performs its operations: packets are logged and inserted into the PCAP before being encrypted. The fact that contents of the PCAP do not correspond to the real messages that were exchanged makes the simulation harder to understand.

4 Conclusion

In this work, we have analysed the Quantum Key Distribution module of the simulator NS-3, called QKDNetSim. The module defines a set of components that work together replicate a quantum channel between two nodes. The structure of QKDNetSim is very promising, since its each of its components is clearly defined, performs a very specific action and adequately represents an element that exists in real quantum networks.

However, some of this components are only surface-deep, and they usually do not perform the actions that they are supposed to. Correcting the shortcomings explained in this work would greatly improve the quality of QKDNetSim, such that it could represent a real quantum network in a more faithful way. Furthermore, these modifications do not require substantial alteration to the foundation of the module, thus simplifying the process of improving QKDNetSim.

Acknowledgements

This work is part of the project TED2021-130369B-C31, TED2021-130369BC32, TED2021-130369B-C33 and TED2021-130492B-C21 funded by MCIN/AEI/10.13039/501100011033 and by "ERDF A way of making Europe".

The work is also funded by the Plan Complementario de Comunicaciones Cuánticas, Spanish Ministry of Science and Innovation(MICINN), Plan de Recuperación NextGenerationEU de la Unión Europea (PRTR-C17.I1, CITIC Ref. 305.2022), and Regional Government of Galicia (Agencia Gallega de Innovación, GAIN, CITIC Ref. 306.2022)

Bibliography

A. Aji, K. Jain, and P. Krishnan. A survey of quantum key distribution (qkd) network simulation platforms. In *2021 2nd Global Conference for Advancement in Technology (GCAT)*, pages 1–8, 2021.

C. H. Bennett and G. Brassard. Quantum cryptography: Public key distribution and coin tossing. *Theoretical Computer Science*, 560:7–11, dec 1984.

Y. Cao, Y. Zhao, J. Wang, X. Yu, Z. Ma, and J. Zhang. Kaas: Key as a service over quantum key distribution integrated optical networks. *IEEE Communications Magazine*, 57(5):152–159, 2019.

Y. Cao, Y. Zhao, Q. Wang, J. Zhang, S. X. Ng, and L. Hanzo. The evolution of quantum key distribution networks: On the road to the qinternet. *IEEE Communications Surveys & Tutorials*, 24(2):839–894, 2022.

J. Illiano, M. Caleffi, A. Manzalini, and A. S. Cacciapuoti. Quantum internet protocol stack: A comprehensive survey. *Computer Networks*, 213:109092, aug 2022.

M. Mehic, O. Maurhart, S. Rass, and M. Voznak. Implementation of quantum key distribution network simulation module in the network simulator ns-3. *Quantum Information Processing*, 16(10):1–23, oct 2017.

M. Mehic, M. Niemiec, S. Rass, J. Ma, M. Peev, A. Aguado, V. Martin, S. Schauer, A. Poppe, C. Pacher, and M. Voznak. Quantum key distribution: A networking perspective. *ACM Computing Surveys*, 53(5), sep 2020.

A. Singh, K. Dev, H. Siljak, H. D. Joshi, and M. Magarini. Quantum internet—applications, functionalities, enabling technologies, challenges, and research directions. *IEEE Communications Surveys & Tutorials*, 23(4):2218–2247, 2021.

Virtual Reality for Cognitive Stimulation in the Elderly and Individuals with Disability

Iago Fernández-Garrido and Ángel Gómez

Centro de Investigación CITIC, Faculty of Computer Science, Universidade da Coruña, 15071 A Coruña, Spain
Centro de Investigación CITIC - Laboratorio Interdisciplinar de Aplicaciones de la Inteligencia Artificial (LIA2), Faculty of Computer Science, Universidade da Coruña, 15071 A Coruña, Spain
Correspondence: iago.fgarrido@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: The recent advances in virtual reality enable the creation of highly realistic and immersive environments. The purpose of this project is to develop a virtual reality tool for cognitive stimulation in the elderly and people with disabilities through gamification. This tool will feature a range of activities, with adjustable difficulty levels and accessibility options. These activities can be combined with a 360º video to create a gymkhana or a circuit, providing users with an engaging and customized experience to enhance cognitive abilities and mental agility.

1 Introduction

Today, thanks to improvements in quality of life and advancements in the healthcare field (medicine, technology, assistance, etc.), life expectancy has increased, and people are reaching advanced ages in better health conditions. Therefore, active aging should be one of the fundamental pillars of both current and future society. The concept of "active aging" was promoted by the World Health Organization (WHO) in the 1990s. The WHO defined it as the process of optimizing health, participation, and security opportunities to enhance the quality of life as people age (World Health Assembly, 1999).

As part of this improvement in quality of life, cognitive stimulation is a crucial component (Franco-Gómez, Laura and García-González, Laura, 2022), the more stimuli a person receives, the greater their autonomy will be, and they will better cope with the cognitive decline associated with aging. These techniques can also be applicable to individuals with intellectual disabilities.

Taking all of this into account,this project proposes the implementation of a virtual reality (VR) tool that enables cognitive stimulation for both individuals with intellectual disabilities and older adults experiencing cognitive decline. This tool will feature a visually enjoyable and user-friendly design to encourage its utilization by the application's target users. Aspects such as usability, simplicity, and user-friendly controls within the virtual environment will be prioritized to ensure an intuitive and enjoyable user experience, minimizing any potential resistance.

The tool will offer a variety of activities designed to target different cognitive areas, with the flexibility to adjust their difficulty according to the user's preferences. Additionally, a 360º video viewing mode will be available, allowing for the integration of activities within the video, such as an obstacle course or gymkhana.

2 Technologies

The use of virtual reality technologies for this project enables the creation of immersive and interactive environments. In this context, we have opted for the Meta Quest 2 virtual reality headset (Meta, 2023a) this virtual reality solution is driven by its capacity to operate autonomously, negating the requirement for dedicated room setups or computer connections (though the latter is available if needed), as well as its integrated hand-tracking technology.

For its software development, we have chosen the Unity development engine (Unity, 2023), which offers a wide range of libraries for VR development and is one of the primary engines for 3D environments.

3 Description of the application

The Virtual Reality tool currently in development starts with a main menu, allowing the professional to choose from various options and environments for conducting activities.

As depicted in the application's navigation diagram (Figure 1), starting from the main menu, users will have the option to directly select an activity or access the 'Walk Mode' configuration. This mode will also include a random mode and the possibility to save preconfigured walks.

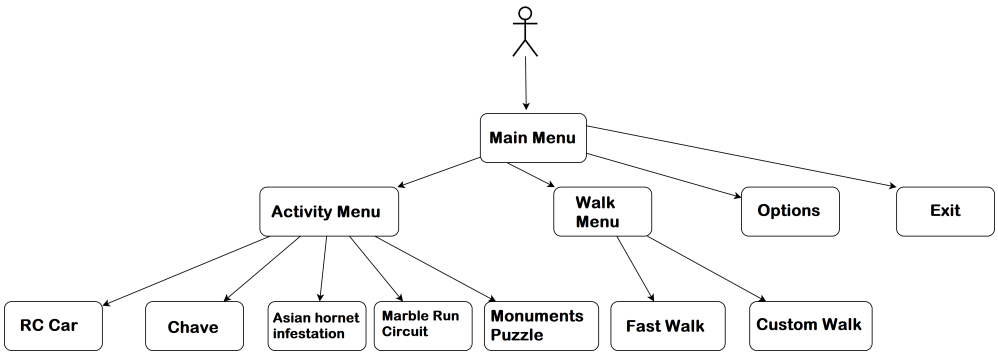


Figure 1: Overall Navigation Diagram

There is a 'Walk Mode,' which involves a 360º video in which various tests or activities are played during its execution, which the user must complete to continue the video playback. Both the video and the displayed activities can be selected in advance before starting the activity. Also, there is the option to perform each activity directly without going through the video process. These activities will be customizable in terms of their difficulty and duration.

We have collaborated with the Adcor Foundation (Adcor, 2023) to develop the following activities:

Asian hornet infestation: It involves a shooting challenge where the user must aim using the controllers at hornets that will appear in the environment.By pressing the primary button, they can shoot a ball, and if it hits the hornet, it will disappear. The difficulty of this challenge can be adjusted by changing the number of hornets to shoot down, their movement speed across the screen, or their stationary presence. Additionally, you can modify the number of hornets present simultaneously. This activity helps the user enhance reflexes, upper body coordination, and stability.

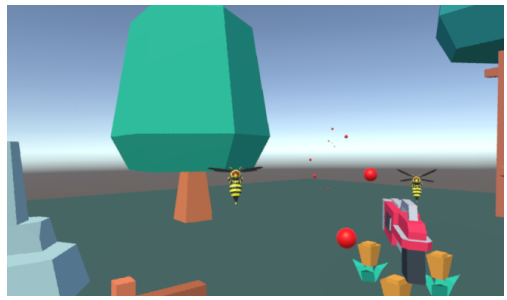


Figure 2: First-Person Screenshot of Asian hornet infestation

RC car: The user will see a radio-controlled car inside a race circuit and will be responsible for controlling it while following a series of waypoints that will appear to complete the course. The time taken to complete the circuit and the time spent off the road will be recorded. The difficulty can be adjusted by increasing or decreasing the length of the circuit. This activity helps improve the user's coordination and reflexes.

Figure 3: First-Person Screenshot of RC car

Chave: An adaptation of the traditional Galician game with the same name (Chave Compostela, 2023), in this activity, the user grabs a 'pello,' either with the controller or using the hand tracking feature, and throws it at the 'chave' with the goal of knocking it down. If needed, the user can approach to the 'chave' as closely as necessary to increase the difficulty. This activity helps improve the user's coordination, balance, and aim.

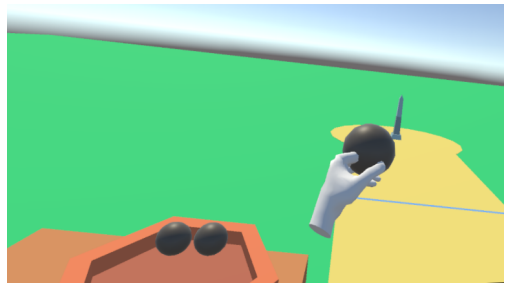


Figure 4: First-Person Screenshot of Chave

Monuments Puzzle: users are tasked with reconstructing various historical monuments that have been broken into pieces within a 3D object. They can use the controllers or their hands as needed. To simplify the task, the pieces don't have to fit perfectly; if the user makes an approximate placement, the program automatically connects them. The difficulty can be adjusted by changing the requirements for when this approximation is needed and the number of puzzles to complete. This activity helps enhance coordination and creativity.

Figure 5: First-Person Screenshot of Monuments Puzzle

Marble Run Circuit: this challenge involves arranging a series of ramp components so that the marble exiting the right square reaches the left square. To assist with this process, guide marks will be placed to indicate where the ramps should be positioned. Similar to the previous activity, if the user brings their position close enough, the ramps will be automatically placed. The ramps can be manipulated using the hands or the controllers. The difficulty can be adjusted by modifying the requirement for approximation or removing it entirely, and the number of puzzles to complete can also be varied. This activity helps improve coordination and creativity.

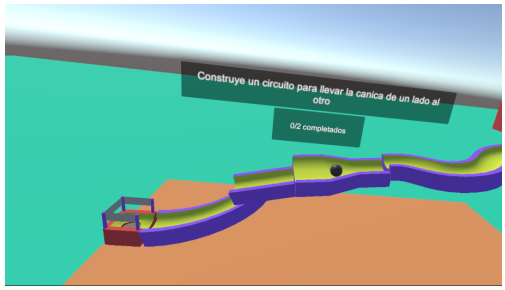


Figure 6: First-Person Screenshot of Marble Run Circuit

4 Discussion

Currently, there are several platforms for selling video games and virtual reality programs, such as Steam (Steam, 2023)] and Meta's own store (Meta, 2023b). The majority of commercial applications are geared towards a more general audience and may not be specialized for individuals with disabilities or the elderly, especially in the context of cognitive stimulation. Additionally, this market often lacks extensive customization options for difficulty levels and controls, as industry standards, while intuitive for those familiar with them, may not be as user-friendly for individuals who are less accustomed to such technologies.

The use of virtual reality for rehabilitation and cognitive stimulation offers numerous advantages, including the high degree of customization for activities, the gamification of physical

tasks that might otherwise be tedious, the safety of performing them in a controlled environment, and the ability to tailor difficulty levels to individual users, among many others.

Therefore, the application being developed in this project can be of significant assistance for cognitive stimulation in elderly individuals and/or those with disabilities. It allows for the adaptation of environments and controls to the user's needs, providing visual and auditory feedback during gameplay.

This tool can be used as a complement to other activities within the center, such as the use of 'Personalized Virtual Reality Environments for Intervention with People with Disability' (Lagos-Rodríguez, Manuel et al., 2022) at the time.

5 Conclusions

The application described in this article is a tool designed for use in the healthcare domain, specifically for cognitive stimulation in the elderly and individuals with disabilities. The utilization of virtual environments tailored to the user's abilities by rehabilitation professionals serves as a motivating factor, encouraging their active participation and commitment to the treatment process.

The environments and activities in this project offer clear advantages. It allows for the adaptation of controls for different user profiles in each of the proposed exercises, enabling the use of controllers and/or hand detection at any time. Additionally, the difficulty level can be adjusted for the various activities in the application.

The proposed tool facilitates gamification of different rehabilitation exercises. Since they take place in a closed and virtual environment, the supervising professional can monitor the user's actions at all times. This application offers a high degree of customization for the various activities and routes, tailoring them to the user's needs.

In the future, it is expected that this application will grow and become a valuable low-cost complement to the activities carried out in various centers involved in the project.

Acknowledgments

ADCOR Foundation
CITIC is financed by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS). The research carried out by TALIONIS group is financed by the Xunta de Galicia (Aid from the Consellería de Cultura, Educación, Formación Profesional e Universidades for the consolidation and structuring of competitive research units) ED431B 2022/39

Bibliography

Adcor. Adcor Website. *https://www.adcor.org/*, 2023. [Online; accessed 10-September-2023].

Chave Compostela. Xogo da chave de Santiago de Compostela e arredores. *http://chavecompostela.com/*, 2023. [Online; accessed 10-September-2023].

Franco-Gómez, Laura and García-González, Laura. *Estimulación Cognitiva Mediante Realidad Virtual: Programa de Estimulación Cognitiva para Adultos Mayores Cognoscitivamente Sanos Mediante Realidad Virtual-NEUROLAU App*. Tesis de grado Maestría en Neuropsicología Clínica, Facultad de Psicología, Universidad CES, Medellín, Colombia, 2022.

Lagos-Rodríguez, Manuel, Gómez-García, Ángel, Pereira-Loureiro, Javier, and Pousada-García, Thais. Personalized Virtual Reality Environments for Intervention with People with Disability. *Electronics*, 11(10):1–14, 2022.

Meta. Meta Quest 2. *https://www.meta.com/es/quest/products/quest-2/*, 2023a. [Online; accessed 10-September-2023].

Meta. Meta Experiences. *https://www.meta.com/es-es/experiences/*, 2023b. [Online; accessed 10-September-2023].

Steam. Steam Store. *https://store.steampowered.com/?l=spanish*, 2023. [Online; accessed 10-September-2023].

Unity. Unity Technologies. *https://unity.com/es*, 2023. [Online; accessed 10-September-2023].

. World Health Assembly. Envejecimiento activo, 1999.

Reinforcement and Homework Control for Children with ASD using a Mobile Application for Apple Watch

Noé Vila, Paula M. Castro, and Adriana Dapena

CITIC Research Center & Department of Computer Engineering, Universidade da Coruña, 15071 A Coruña, Spain
Correspondence: paula.castro@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: The article shows the design and development of a pair of mobile applications for Apple Watch and iPhone devices to reinforce and control tasks for children with autism disorders. This work takes advantage of the possibilities of smartwatches, such as internet connectivity, access to biometric sensors and ease of communication with users, to improve their autonomy in carrying out everyday routines. The application has been designed taking into account the needs and characteristics of this group, collaborating with two non-profit organisations in the area of A Coruña dedicated to the care of children with autism. The obtained final product is functional and has an impact on the quality of life of these children. This development contemplates the possibility of incorporating artificial intelligence to break down a task into smaller sub-tasks, in order to facilitate its understanding and the completion of the activity. Its use could also be extended to adults, if symbols are incorporated instead of pictograms.

1 Introduction

Wearable devices are electronic devices that can be worn as accessories or clothing. These devices are designed to be portable, often incorporating sensors, communication technology and data processing capabilities. Wearable devices include smartwatches, fitness wristbands or headsets and smart glasses, among others (see McCann and Bryson (2022)). These devices allow wearers to monitor and collect information about their health, physical activity or sports performance, for example, as well as provide notifications and additional functionalities (see Sazonov (2020)). They aim to provide a more personalised and accessible technological experience, integrating into the daily lives of users. In particular, in the field of therapies for children with Autism Spectrum Disorder (ASD) (Landa (2011); Organization (2021); Volkmar and Lord (2012)), wearable devices could provide them with a constant and unobtrusive support tool, helping them to organise their daily activities and develop autonomous task completion skills. The watch's display and notifications can serve as visual and auditory reminders, making it easier to follow instructions and complete tasks. In addition, the watch's touch interaction and sensors offer the potential to provide immediate feedback and visual rewards, which encourages task completion and fosters positive reinforcement. Taking into account the difficulties that children with ASD have in planning and organising their activities, it is necessary to develop tools that facilitate the daily lives of children with ASD and their environment, reducing situations of rejection and anxiety. In this sense, an app for Apple Watch, such as the one shown in this article, could be a perfect solution, as this device provides an accessible and portable interface for monitoring and reinforcing their tasks (see Apple Inc. (2022)).

To date, there are no applications specifically designed for wearable devices aimed at people with ASD. Although there are several apps and technological tools for people with ASD, most of them focus on mobile devices such as phones and tablets, offering a wide range of functions such as Augmentative and Alternative Communication (AAC), visual support, routines, and task tracking. The lack of apps specifically designed for wearable devices and people with ASD can be attributed to several factors. Firstly, the development of wearable apps requires special care of the wearable interface and usability, taking into account the specific needs and characteristics of users. This involves adapting the interface to be intuitive, easy to use and accessible, which makes it very difficult to develop. In addition, the design of such applications for people with ASD also needs to consider the particular nature of their needs and characteristics. For example, they may have difficulties with verbal communication and social understanding, requiring specific approaches to the presentation of information and interaction, therefore, existing applications oriented to function as reminders would not work for these individuals.

This paper presents in Section 2 the objectives for the development of a pair of iPhone and Apple Watch applications with the purposes above described; Section 3 shows material and methods used for project development; Section 4 presents the apps developed for iPhone and Apple Watch; in Section 5 the results obtained and expected from these applications are shown and discussed and, finally, the conclusions are presented in Section 6.

2 Objectives

The aim of this work is to design and develop a pair of mobile applications for Apple Watch and iPhone devices that reinforces and monitors tasks in children with autism spectrum disorders. The application leverages the capabilities of smartwatches, such as internet connectivity, access to biometric sensors, and user-friendly communication, to improve the autonomy of children with autism in performing daily tasks. The application is designed focusing on the specific needs and characteristics of this group, with a user-friendly and accessible interface. Therefore, the primary objective is to promote independence and specific skills in performing daily activities, which can have a significant impact on the quality of life of children with autism. The use of an Apple Watch would allow children with ASD to have a constant and discreet support tool for such purposes.

The project must have two types of users who can interact with the two applications: the administrator and the user. The administrator can be any type of parent or educator of the child with ASD, and the application (app in the following) that is intended for him/her is the resulting iPhone app. On the other hand, the user of the Apple Watch app will be the child with ASD. This user will only have access to this app.

3 Material and Methods

For the development of the project we have chosen to use an agile methodology such as SCRUM (Beck et al. (2001); Schwaber (2004)), due to its ability to deal with the challenges of developing a mobile application, such as changing requirements or efficient management of time and resources.

The choice of development tools such as watchOS (Apple Inc., 2023d), Xcode (Apple Inc., 2023e), Swift (Apple Inc., 2023a), SwiftUI (Apple Inc., 2023b) and WatchKit (Apple Inc., 2023c) over alternatives is closely linked to the choice of the Apple Watch as primary device. These tools were designed and developed specifically for Apple Watch development, making them almost exclusive to this task. However, these tools offer a complete and efficient development environment.

In this work we will include pictograms from Portal Aragonés de Comunicación Aumentativa y Alternativa (ARASAAC) (see Palao (2020)), a trademark backed by the Government of Aragon that offers a wide collection of pictograms, named as ARASAAC pictograms.

ARASAAC focuses on providing graphic resources for people with autism, intellectual disabilities, language difficulties or elderly people, among others. The inclusion of this collection of pictograms for communication in our apps is due to their recognition as the most widely used set of pictograms by both associations working with people who have this disorder and their families.

Generative Pre-trained Transformer (GPT)-3 is a language model developed by OpenAI (OpenAI, 2023). This model is based on the neural network architecture known as Transformer and is trained using large volumes of text data from the web. GPT-3 has gained recognition in the field of natural language processing due to its good ability to generate coherent and contextually relevant text in a variety of applications, including language translation, creative textual content creation, question answering and many other tasks. GPT-3.5 Turbo represents an evolution in the GPT-3 series of models, developed by OpenAI in 2022. This version is characterised by significant improvements in text generation. The use of GPT-3.5 Turbo in this project internally provides the possibility of creative textual content creation to allow the decomposition of complex tasks into more manageable sub-tasks, with the purpose of assisting the user in their execution. For example, when the task "Swimming pool" is given, it could provide a simplification into smaller tasks, such as "Hat", "Goggles", "Swimming suit" and "Swimming", which allows to automatically obtain the sub-tasks of a task without the need to set them beforehand.

4 Project Development

We will briefly present in the following two sections the apps developed for iPhone and Apple Watch.

4.1 iPhone App

This app allows users to identify themselves in order to have access to the data they manage under the application. In this way, tasks and data are stored and can be accessed from different devices, such as many models of iPhones and even iPads. The information is kept in the user's account, which ensures availability and continuity of tasks at any time and place. The administrator user can access the child's sensitive data collected from the use of the Apple Watch app. This ensures an adequate level of protection of information, as authorisation is required to access such sensitive data each time it is to be consulted.

The iPhone app also offers a number of account management options for users, so it includes the following: changing the password, logging out, recovering a forgotten password by sending a recovery email and creating a new account. This gives the user control to manage their account in a convenient and tailored way.

This iPhone app also gives users the ability to manage their tasks easily and efficiently, creating, deleting and listing them in an orderly fashion as new tasks are added (see Figure 1). The tasks focus on the ARASAAC pictograms, also allowing customisation to suit the needs of each child, and establishing different parameters to be defined by the user. In addition to the functions related to the tasks, the app also offers a pictogram search engine (see Figure 2). This is particularly useful for users, as they will be able to access a wide variety of pictograms through the ARASAAC database. By using this search engine further facilitates the process of finding and selecting the appropriate pictograms for each specific task or situation. This ensures a personalised experience adapted to the needs of each child.

Some of the sensitive information, to which the iPhone app user will have access, relates to the data collected by the child's Apple Watch app. These collected data are mainly focused on providing relevant information to improve the efficiency and effectiveness of assigned tasks. These include success and failure rates of tasks, time for task completion, time when a task is cancelled, or biometric data provided by the own Apple Watch. Emphasis is placed on the importance of guaranteeing the protection of this sensitive data and ensuring that it can only be

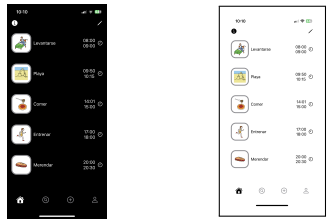


Figure 1: Home page for iPhone app.

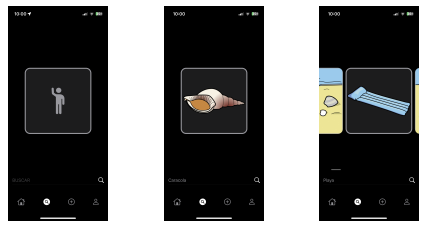


Figure 2: Pictogram search engine.

accessed by the administrator or authorised user. Authentication through password or biometric methods (via facial recognition or fingerprint reader) is an additional security measure that is implemented to maintain the confidentiality and security of the information.

4.2 Apple Watch App

The Apple Watch app provides a complete and up-to-date overview of all the tasks assigned to the child throughout the day. In addition to task displaying, it also provides relevant information on the nature of each one. The current homework is prominently highlighted, allowing the user to interact to successfully complete it, request help or cancel it as needed. This feature gives the user full control over tasks in real time. The app also allows the user to view all tasks, including past and future tasks that are still pending (see Figure 3). This is very useful to have an overview of all scheduled and completed tasks. In this way, the user can progress and make sure that all tasks are completed properly.

Figure 3: Past, current and future tasks.

At any time, a short press on the current task will trigger the display of a spinning coin animation (see Figure 4). This animation will be available for all tasks, whether current, past or future. Through this animation, the task information, including its name, can be accessed in capital letters for clearer understanding, as well as start and end time, represented only with symbols and numbers, which facilitates its use for children with ASD.

Moreover, this app is able to autonomously decompose tasks into sub-tasks using Artificial Intelligence (AI), adapting at all times to a non-verbal language so that the target audience is able to understand it (see Figure 5).

Interaction with the app is constantly guided through intuitive actions, such as swiping the

Figure 4: 3D coin spinning animation.

Figure 5: Example of the use of AI in help: results for two applications.

screen and manipulating the digital crown to navigate between the different tasks predefined by administrators. The inclusion of pictograms and geometric shapes, such as circles and progress bars, is a special consideration to the condition of these children. Careful attention is paid to the presentation colours, which are automatically displayed in a tone that harmonises with the pictogram currently in focus (see Figure 3), and which also play a representative role in the buttons: red colour indicates unsuccessful completion and, otherwise, green is used (see Figure 5). This creates a coherent visual association to facilitate navigation and understanding.

During the use of this app, the user's heart rate per minute is monitored in the background. If it is detected that the heart rate is too high, the user is warned that it is advisable to relax. This situation usually occurs when one of the tasks is causing stress. To provide support, an additional option is offered to terminate the current task, which helps to reduce that level of stress. A request for help is also provided if needed. It is important to note that all user activity within the app is recorded in real time. This includes the time for completing tasks and the total time for application or help use. These logs provide valuable information for detailed tracking and analysis of the user's performance. In addition, this allows to check the effectiveness of both apps and to make improvements based on usage patterns and user needs.

5 Discussion and Results

The data collected through these applications will allow professionals and families of children with ASD to monitor the daily tasks they perform and reinforce behaviours that will enable them to make progress. The statistical data collected will help achieve the improvement in their autonomy and self-esteem that is essential for the quality of life of these people. The project ensures that all this information is recorded, as can be seen in Figure 6, with the administrator being able to consult the data collected by means of interactive graphics. In this figure, as an example, it can be observed that the time spent for the completion of the task is the 99 % of the maximum time, the task is successfully completed 9 of 10 times it has been done and the evolution of the number of times the child has consulted the help during the last 10 days.

On the other hand, although from validation and testing with two entities they indicated that the final product was perfectly adapted to their needs, they pointed as improvement the selection of those tasks to control the child's keystrokes and the possibility of replacing pictograms with real pictures for some children. On the other hand, one of the entities highlights the advantage of using a wearable device such as a smartwatch, as the user can carry with him/her the programming of the tasks he/she has to carry out, knowing it in advance in a standardised way, and not worrying about forgetting the device somewhere. In this way, we reduce the situations of anxiety and stress that this type of planning tends to cause in children with these

disorders.

Nowadays, AI is a technology in constant evolution and increasingly present in our daily lives. In the case of this project, even though it was not in the initial objectives and requirements, it was decided to integrate AI to improve task prediction and decomposition. This is due to the fact that children with autism often present situations of stress and anxiety when faced with complex tasks, and they need clear commands and direct instructions to help them perform the different activities of daily life. Thus, they can have automatically generated support thanks to AI that is perfectly adapted to their individual needs.

As for the viability of this developments, it can be affirmed that it is completely viable and this is evident thanks to the factors that will be explained below.

Firstly, it should be noted that the market for mobile applications aimed at children with ASD is constantly growing. This market is characterised by the scarcity of products available and an increasing demand for technological solutions that can improve the quality of life of these children. In addition, the growing popularity of mobile devices and wearable technology, such as the Apple Watch used in this work, offers a significant market potential for the pair of applications here developed. In fact, to date, there is no other application available for the Apple Watch that focuses on this specific audience, which represents a unique opportunity. Also, the resulting Apple Watch app is the first in the world of existing apps for children with ASD to incorporate AI in this specific audience.

From a technical perspective, it is important to note that the developed project is fully viable. This is based on the fact that a fully functional and ready-to-use pair of apps has been created. Moreover, from an economic point of view, the cost of development and maintenance is reasonable.

Finally, it is relevant to mention that the proposed mobile app complies with all applicable laws and regulations, including laws of European Union. Furthermore, any use of external sources is appropriately referenced within the application to ensure compliance with all applicable laws and regulations, including data protection and privacy laws.

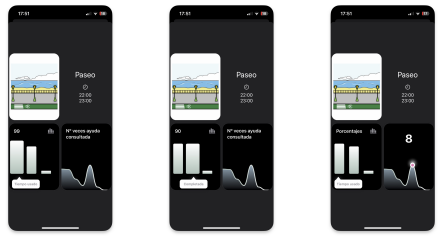


Figure 6: Results with interactive graphics.

6 Conclusions

In the work carried out, it is worth highlighting the achievement of developing two fully functional applications that cover all the objectives proposed at the beginning of the project. The aim of these developments is to meet the real needs of people with diversity, mainly those with ASD, who require comfortable and easy-to-use applications to carry out everyday tasks and thus improve their quality of life. This work responds to these needs, paying meticulous attention to the design of a simple and user-friendly interface, and taking into account the active collaboration of two entities in the area of A Coruña, which participate in the development throughout the whole process. The tests with real users, i.e., children with ASD, will be carried out in the coming months.

The main future direction of this project is its publication in the Apple App Store. Other possible lines are the internationalisation of the app or the use of descriptive icons instead of pictograms, to make them more suitable for adults.

Acknowledgements

This work has been supported by grant ED431C 2020/15 funded by Xunta de Galicia and ERDF Galicia 2014-2020; by grant PID2019-104958RB-C42 (ADELE) funded by MCIN/AEI/10.13039/501100011033; and by project TED2021-130240B-I00 (IVRY) funded by MCIN/AEI/10.13039/501100011033 and by the European Union NextGenerationEU/PRTR.

Bibliography

Apple Inc. Apple watch, 2022. URL *https://www.apple.com/watch/*.

Apple Inc. Swift programming language. Apple Developer, 2023a. *https://developer.apple.com/swift/*.

Apple Inc. SwiftUI framework. Apple Developer, 2023b. *https://developer.apple.com/xcode/swiftui/*.

Apple Inc. Watchkit framework. Apple Developer, 2023c. *https://developer.apple.com/watchkit/*.

Apple Inc. WatchOS. Apple Developer, 2023d. *https://developer.apple.com/watchos/*.

Apple Inc. Xcode. Apple Developer, 2023e. *https://developer.apple.com/xcode/*.

K. Beck, M. Beedle, A. Van Bennekum, A. Cockburn, W. Cunningham, M. Fowler, ..., and J. Kern. Manifiesto por el desarrollo ágil de software, 2001. URL *https://agilemanifesto.org/*.

R. J. Landa. Early identification of infants at risk for autism spectrum disorder. *The Journal of Pediatrics*, 159(2):191–192, 2011.

J. McCann and D. Bryson. *Smart clothes and wearable technology*. Woodhead Publishing, 2022.

OpenAI. Openai, 2023. URL *https://www.openai.com/*.

W. H. Organization. Autism spectrum disorders, 2021. URL *https://www.who.int/es/news-room/fact-sheets/detail/autism-spectrum-disorders*.

S. Palao. Pictogramas ARASAAC. ARASAAC, 2020. *http://www.arasaac.org*.

E. Sazonov. *Wearable Sensors: Fundamentals, implementation and applications*. Academic Press, 2020.

K. Schwaber. *Agile Project Management with Scrum*. Microsoft Press, enero 2004. ISBN 0-7356-1993-X.

F. R. Volkmar and C. Lord. Autism spectrum disorders and other developmental disorders. In *The American Psychiatric Publishing Textbook of Psychiatry*, pages 918–959. American Psychiatric Publishing, 6th edition, 2012.

Enhancing Pathological Detection and Monitoring in OCT Volumes with Limited Slices using Convolutional Neural Networks and 3D Visualization Techniques

Emilio López-Varela, Noelia Barreira, Nuria Olivier Pascual, and Manuel G. Penedo

Grupo VARPA, Instituto de Investigación Biomédica de A Coruña (INIBIC), Universidade da Coruña, Xubias de Arriba, 84, 15006 A Coruña, Spain
Centro de Investigación CITIC, Universidade da Coruña, 15071 A Coruña, Spain
Servizo de Oftalmoloxía, Complejo Hospitalario Universitario de Ferrol (CHUF), Av. da Residencia S/N, Ferrol 15405, Spain
Correspondence: e.lopezv@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: Optical Coherence Tomography (OCT) is a non-invasive imaging technique with a crucial role in the monitoring of a wide range of diseases. In order to make a good diagnosis it is essential that clinicians can observe any subtle changes that appear in the multiple ocular structures, so it is imperative that the 3D OCT volumes have good resolution in each axis. Unfortunately, there is a trade-off between image quality and the number of volume slices. In this work, we use a convolutional neural network to generate the intermediate synthetic slices of the OTC volumes and we propose a few variants of a 3D reconstruction algorithm to create visualizations that emphasize the changes present in multiple retinal structures to aid clinicians in the diagnostic process.

1 Introduction

Optical Coherence Tomography (OCT) is a non-invasive imaging modality of great importance for the diagnosis and ongoing monitoring of a wide range of ocular and systemic diseases. By providing an accurate examination of the three-dimensional anatomy of the eye, OCT facilitates the early detection and monitoring of diseases such as central serous chorioretinopathy López-Varela et al. (2023b), diabetic macular edema Vidal et al. (2020), age-related macular degeneration López-Varela et al. (2022), hypertension, and even multiple sclerosis Garcia-Martin et al. (2021); López-Varela et al. (2022). These diseases have a profound impact on the lives of patients, ranging from irreversible visual impairment to gradual loss of mobility. Hence, the timely and accurate diagnosis of these conditions becomes essential Medeiros et al. (2005).

The effectiveness of OCT in diagnosing disease depends on the ability of clinicians to quickly detect even the most subtle changes in ocular structures. The task of visualizing these alterations in the discrete 2D slices that constitute each OCT volume is labor-intensive and time-consuming. Moreover, this method is prone to subjectivity and relies heavily on the experience of the clinician, who must extrapolate three-dimensional structural changes from observations of individual slices Apostolopoulos et al. (2017). In contrast, 3D reconstructions provide a complete and accurate visualization of all structures in the OCT volume at a glance. This mode

of visualization substantially streamlines the workflow of the clinician, increasing the robustness and accuracy of diagnoses and significantly reducing the workload and, consequently, the direct and indirect costs incurred by healthcare services.

In order to use 3D visualization effectively, it is imperative that 3D OCT volumes have good resolution in each axis. Unfortunately, due to the acquisition mechanism of OCT, there is a trade-off between the image quality and the number of volume slices, which limits the resolution in some of the axes and forces to choose different settings of these parameters depending on the objective of the study Huang et al. (1991). Many routine clinical studies opt for OCT volumes composed of a limited number of high-resolution images, which decreases the effectiveness of 3D visualizations in the diagnostic process by profoundly reducing the resolution of the cube in the slice dimension as shown in Figure 1.

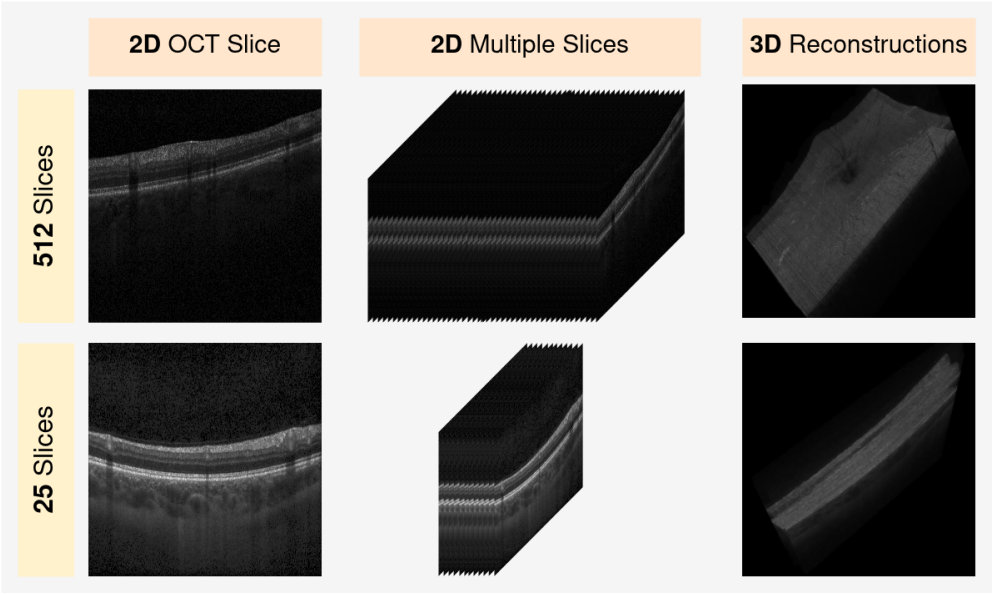


Figure 1: Two OCT cubes with different number of slices. The 2D representation of a single image and a 3D reconstruction of the volume are shown.

It is therefore of great interest to create a system capable of synthesising the intermediate images between adjacent slices in order to increase the cross-sectional resolution of the volume. As a consequence, some papers have begun to propose different initial approaches based on neural networks to solve similar problems in OCT López-Varela et al. (2023a) or in other medical imaging modalities such as magnetic resonance imaging Peng et al. (2020). In addition to having an OCT volume with good resolution, it is essential that the 3D reconstruction algorithm produces a visualization where each of the analyzed diagnostic features is accurately identified. Structural alterations, such as the presence of fluid, can be discerned in a standard reconstruction, but it is challenging for a clinician to estimate subtle changes in a more general feature, such as layer thickness, at a glance. These types of changes are much better visualized using alternative visualization systems, such as depth-based heatmaps, in combination with complementary techniques like structure segmentation. Therefore, there is a paramount interest in designing and creating a system that employs a combination of multiple visualizations that complement each other, emphasizing different characteristics of the analyzed structure.

Given all these problems, in this work, we use a convolutional neural network trained in an unsupervised manner, which leverages information from adjacent slices to generate the intermediate synthetic slices of the volume. To fully exploit this enhanced resolution cube and

accurately monitor changes in multiple diagnostic biomarkers, several variants of a 3D reconstruction algorithm are proposed. These algorithms generate diverse visualizations with complementary functions, highlighting changes in various retinal structures, including layer thickness. These visualizations hold the potential to significantly aid clinicians in the diagnostic process, emphasizing the paramount importance of a comprehensive visual approach in medical diagnostics.

2 Materials and Methods

A complete diagram showing the different processes of the proposed diagnostic visualization system can be observed in Figure 2. To assess the effectiveness of the proposed visualization system, in this work, we use the dataset and the network architecture presented in López-Varela et al. (2023a). This dataset consists of 42 OCT cubes obtained using two identical Spectralis R OCT capture devices from Heidelberg Engineering. These cubes encompass three distinct categories of OCT volumes captured by different configurations commonly used in clinical practice and have a limited number of slices. This characteristic significantly degrades 3D visualizations, highlighting the indispensability of the generative network for enhancing the visualization quality. In addition to the generative network, and as a complement to our visualization system, we employ a segmentation network trained to segment the multiple layers of the retina. This segmentation network is a variation of the UNet Ronneberger et al. (2015) with the classic encoder-decoder structure. The segmentation masks generated by this network enable us to individually fine-tune the appearance of each layer within the 3D visualization, with the purpose of emphasizing the most significant diagnostic features.

Finally, to establish a 3D reconstruction algorithm to exploit the OCT volumes with enhanced resolution, we have adapted and improved the reconstruction algorithm proposed in López-Varela et al. (2022), creating four complementary variants that highlight different aspects of the analyzed layers. These variants are obtained by altering three parts of the reconstruction methodology: the voxel values of each OCT volume, the color-opacity transfer function, and the composition system of the ray casting algorithm.

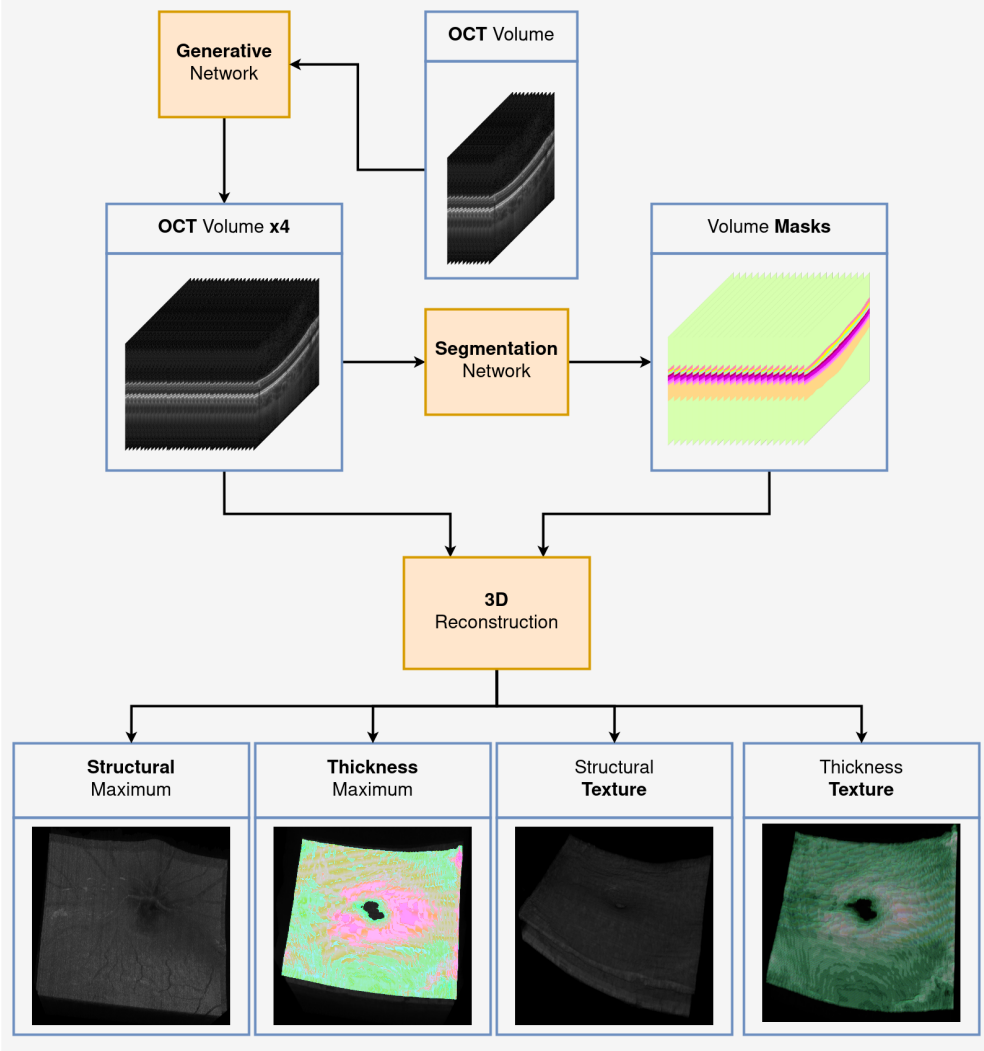


Figure 2: Diagram of the diagnostic visualization system.

- **Structural Maximum**: To obtain this variant, the voxel values are adjusted using the segmentation mask of the analyzed structure. This way, the graylevel intensities of voxels belonging to the layer are compressed into the range $[20, 256]$, while values that do not belong to it are compressed into the range $[0, 19]$. The opacity transfer function is adjusted to increase transparency for values outside the layer in the range $[0, 19]$, and the maximum intensity is used as the composition system.
- **Thickness Maximum**: In addition to the processing applied in 'Structural Maximum', the intensities of voxels belonging to the layer are adjusted to obtain a heat map indicating the thickness. To achieve this, a distance transform is applied to the mask cube. The result of applying the distance transform is a volume where the graylevel intensities of points inside the layer are changed to show the distance to the closest boundary from each point. The voxels of the original cube are modified based on the distance-transformed cube to obtain an estimation of the layer thickness. Lastly, the transfer function is adjusted to apply a Red-Green color scale based on the intensity value of

each voxel.

- **Structural Texture**: In addition to the processing applied in 'Structural Maximum', a front-to-back composition scheme using alpha blending is used as the composition system.
- **Thickness Texture**: In addition to the processing applied in 'Thickness Maximum', the opacity gradient transfer function is modified according to the distance-transformed cube, and the composition scheme is applied as the composition system.

3 Results and Discussion

In this section, we present the results obtained by employing our visualization methodology on OCT volumes with a low number of slices. To assess the utility of each component comprising our system, we display the four proposed visualization variants before and after applying the generative network to a 25-slice OCT volume. Figure 3 illustrates this effect in two planes of two distinct layers of the retina, the retinal nerve fibre layer (RNFL) and the ganglion cell layer (GCL). To focus the visualization on these layers, we utilize the masks obtained through the proposed segmentation network. The use of this network enables the individual extraction and emphasis of each clinically significant structure.

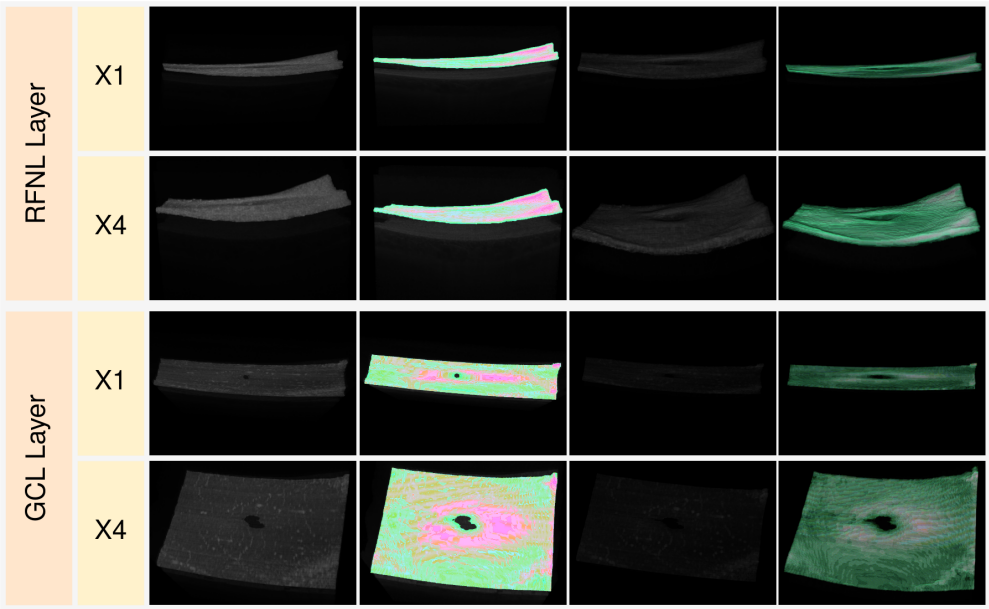


Figure 3: Sample figure caption.

In general, it can be observed that the use of the generative network is essential for achieving accurate reconstruction that displays the details of each layer. When using the original OCT volume, the lack of resolution prevents us from discerning any kind of detail in the layer, such as blood vessels or the macular hole. It even hinders our ability to distinguish the shape and thickness of the structure. Conversely, these characteristics are clearly distinguishable when using the enhanced OCT volumes. These results reinforce the conclusions obtained in López-Varela et al. (2023a), where the realism of enhanced OCT volumes and the utility of synthetic layer generation were validated by clinical experts.

Each variant of the reconstruction algorithm emphasizes a feature with diagnostic significance. The first variant allows us to observe the internal details of the layer clearly, such as blood vessels, and provides good contrast to distinguish the multiple structures that compose the layer at different depths. In contrast to this, the third variant shows the surface details of the layer, allowing for precise observation of its exact shape, which helps identify any malformations. The second and fourth variants display the thickness of the layer on a green-red color scale, enabling the observation of pathological thickening or thinning. Detecting these thickness changes in the layers is crucial for diagnosing and monitoring many diseases, such as multiple sclerosis, making the contribution of these two visualizations to the system significant. These types of visualizations also assist in visualizing other pathological structures, such as fluid accumulation, making their use versatile and of general nature. The diagnostic system as a whole enables precise visualization and differentiation of the multiple pathological features present in the ocular structure at a glance. Therefore, we can conclude that this system has great potential to assist clinicians and significantly lighten the diagnostic workload.

Acknowledgements

This research was funded by Government of Spain, Ministerio de Ciencia e Innovación y Universidades, Government of Spain, RTI2018-095894-B-I00 research project; Ministerio de Ciencia e Innovación, Government of Spain through the research projects with reference PID2019-108435RB-I00, PDC2022-133132-I00 and TED2021-131201B-I00; Consellería de Cultura, Educación e Universidade, Xunta de Galicia through the Grupos de Referencia Competitiva, grant ref. ED431C 2020/24; CITIC, as Research Center accredited by Galician University System, is funded by "Consellería de Cultura, Educación e Universidade from Xunta de Galicia", supported in an 80% through ERDF Funds, ERDF Operational Programme Galicia 2014-2020, and the remaining 20% by "Secretaría Xeral de Universidades", grant ref. ED431G 2019/01. Emilio López Varela acknowledges its support under FPI Grant Program through PID2019-108435RB-I00 project.

Bibliography

S. Apostolopoulos, S. De Zanet, C. Ciller, S. Wolf, and R. Sznitman. Pathological oct retinal layer segmentation using branch residual u-shape networks. In *Medical Image Computing and Computer Assisted Intervention- MICCAI 2017: 20th International Conference, Quebec City, QC, Canada, September 11-13, 2017, Proceedings, Part III 20*, pages 294–301. Springer, 2017.

E. Garcia-Martin, M. Ortiz, L. Boquete, E. M. Sánchez-Morla, R. Barea, C. Cavaliere, E. Vilades, E. Orduna, and M. J. Rodrigo. Early diagnosis of multiple sclerosis by oct analysis using cohen's d method and a neural network as classifier. *Computers in Biology and Medicine*, 129: 104165, 2021.

D. Huang, E. A. Swanson, C. P. Lin, J. S. Schuman, W. G. Stinson, W. Chang, M. R. Hee, T. Flotte, K. Gregory, C. A. Puliafito, et al. Optical coherence tomography. *Science*, 254(5035):1178–1181, 1991.

E. López-Varela, P. L. Vidal, N. O. Pascual, J. Novo, and M. Ortega. Fully-automatic 3d intuitive visualization of age-related macular degeneration fluid accumulations in oct cubes. *Journal of Digital Imaging*, pages 1–12, 2022.

E. López-Varela, N. Barreira, N. O. Pascual, M. R. A. Castillo, and M. G. Penedo. Generation of synthetic intermediate slices in 3d oct cubes for improving pathology detection and monitoring. *Computers in Biology and Medicine*, 163:107214, 2023a.

E. López-Varela, J. de Moura, J. Novo, J. I. Fernández-Vigo, F. J. Moreno-Morillo, and M. Ortega. Fully automatic segmentation and monitoring of choriocapillaris flow voids in octa images. *Computerized Medical Imaging and Graphics*, 104:102172, 2023b.

E. López-Varela, N. Barreira, N. O. Pascual, E. G. Ben, S. R. Cid, and M. G. Penedo. Fully automatic segmentation of the choroid in non-edi oct images of patients with multiple sclerosis. *Procedia Computer Science*, 207:726–735, 2022. Knowledge-Based and Intelligent Information and Engineering Systems: Proceedings of the 26th International Conference KES2022.

F. A. Medeiros, L. M. Zangwill, C. Bowd, R. M. Vessani, R. Susanna Jr, and R. N. Weinreb. Evaluation of retinal nerve fiber layer, optic nerve head, and macular thickness measurements for glaucoma detection using optical coherence tomography. *American journal of ophthalmology*, 139(1):44–55, 2005.

C. Peng, W.-A. Lin, H. Liao, R. Chellappa, and S. K. Zhou. Saint: spatially aware interpolation network for medical slice synthesis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 7750–7759, 2020.

O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*, pages 234–241. Springer, 2015.

P. L. Vidal, J. de Moura, M. Díaz, J. Novo, and M. Ortega. Diabetic macular edema characterization and visualization using optical coherence tomography images. *Applied Sciences*, 10(21): 7718, 2020.

Valuation and Risk Modeling of Renewable PPAs

Julio Pombo-Romero, Oliver Rúas-Barrosa, and Carlos Vázquez

Instituto de Energía Solar - Universidad Politécnica de Madrid, 28031, Madrid, Spain
Department of Mathematics, Faculty of Informatics and Centro de Investigación CITIC, Campus Elviña s/n, 15071-A Coruña (Spain)
Correspondence: oliver.ruas@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract:
Renewable energy (RE) projects aim to sell electricity to the consumers, which may be carried out by means of a Power Purchase Agreement (PPA). These PPAs will yield payments over a long period of time so as to refinance the project. Nevertheless, the present literature which addresses the economic appraisal of renewable energy projects focuses on the project developer and whether an onsite PPA is deemed, the actual cost of energy is not taken into consideration and nor the embedded options of the PPA. We propose such a model and carry out a comparison with the most common approach, a swap between fixed and market price.

1 Introduction

The development of renewable energy (RE) projects has enabled the decarbonization of some regions. These projects involve a high capital expenditure, which is repaid over long periods of time. One way of ensuring this repayment and thus making the renewable deployment possible is by offering long term agreements, for instance, a Power Purchase Agreement (PPA), which is a bilateral contract between 2 parties: an energy producer and an energy consumer (the offtaker) which settle on a price for the electricity supply during a long period of time Peña et al. (2022).

The vast majority of the renewable projects appraisals take into consideration the viability of a project by computing a single value, the Levelized Cost of Energy (LCOE), which encompasses the capital expenditure, maintenance expenditure, the interest rate linked to the general market conditions and the riskiness of the project, and technical features of the project. This value is the unit cost of energy and yields a value which allows for a comparison with other projects Carrêlo et al. (2020). For the completion of the calculation, several features can be added such as a probability distribution for the energy production Freeman et al. (2018) and how the placement affects the projects' viability Gabrielli et al. (2022). Another common approach in projects assessments is that of real options, which is focused on the development stage of the project or on how it should be run.

It is clear that these approaches pay attention to the party that develops the project, while the offtakers' point of view is dismissed. Regarding this latter point, the valuations reviewed consider the value of a PPA as a swap between two parties. One of them pays a fixed amount and the other one the market price. Nevertheless, this is only true when the producer must purchase the electricity, that is, acting as a distributor company, while for the case the producer owns a RE project, the unit cost of energy (LCOE) is the real price of the supplied electricity.

Therefore, we aim at developing a model for the valuation of a PPA taking into consideration the embedded options of a PPA.

2 Mathematical model

In this section we briefly describe the main ideas to pose the mathematical modelling for RE PPA pricing, also including the specific stochastic model that has been considered for the electricity price evolution. Note that this price model is a building block and can be replaced by alternative models.

2.1 PPA model

The optional component of a RE PPA is assessed with respect to its alternative (that supplied amount of electricity at the market price). For that purpose, we will assume that a RE PPA is signed at t_0 and expires at t_M, having $M+1$ times when the electricity supply takes place. For each unit of supplied electricity $U(t_k)$ for $k = 0, \dots, M$, which we will consider constant and equal to $1\, MW \cdot h$ for simplification, the offtaker will pay an amount P_{PPA}, and if at $t_k > t_0$ the offtaker decides not to pay, then s/he will quit the RE PPA. Therefore, by paying at t_0, the offtaker will obtain savings equal to the difference between the market price (P_t) and P_{PPA} and s/he also buys a call option on the underlying asset which matures at t_1, which we denote as $H^1(P_{t_1})$. By paying at t_0 and t_1, the offtaker will obtain savings at t_0 and t_1 and s/he buys a call option on the underlying asset which matures at t_2, that is $H^2(P_{t_2})$, and so on. In this way we have the following set of possible scenarios

$$
\begin{aligned}
&(P_{t_0} - P_{PPA}) + e^{-r(t_1-t_0)} H^1(P_{t_1}) \\
&(P_{t_0} - P_{PPA}) + e^{-r(t_1-t_0)} \left((P_{t_1} - P_{PPA}) + e^{-r(t_2-t_1)} H^2(P_{t_2}) \right) \\
&(P_{t_0} - P_{PPA}) + e^{-r(t_1-t_0)} (P_{t_1} - P_{PPA}) + e^{-r(t_2-t_0)} \left((P_{t_2} - P_{PPA}) + e^{-r(t_3-t_2)} H^3(P_{t_3}) \right) \\
&\quad \vdots \\
&(P_{t_0} - P_{PPA}) + \cdots + e^{-r(t_{M-2}-t_0)} \left((P_{t_{M-2}} - P_{PPA}) + e^{-r(t_{M-1}-t_{M-2})} H^{M-1}(P_{t_{M-1}}) \right) \\
&(P_{t_0} - P_{PPA}) + \cdots + e^{-r(t_{M-1}-t_0)} \left((P_{t_{M-1}} - P_{PPA}) + e^{-r(t_M-t_{M-1})} H^M(P_{t_M}) \right).
\end{aligned}
$$

Due to the possibility of quitting the PPA at any time $t_k > t_0$, the offtaker will pursue the scenario with the maximum payoff. For this reason, the PPA value for the offtaker, which we will denote by V_{PPA}, will be given by the maximum of the possible values of the PPA

$$
V_{PPA}(t_k) = \max_{k+1 \leqslant m \leqslant M} \left(\sum_{j=k}^{m-1} e^{-r(t_j - t_k)} \left[P_{t_j} - P_{PPA} \right] + e^{-r(t_{m-1}-t_k)} e^{-r(t_m - t_{m-1})} H^m(P_{t_m}) \right), \tag{24.1}
$$

with $k = 0, \dots, M-1$.

The financial appraisal of the PPA is not over, it remains some measurements which may be computed, more precisely, the default probability defined as

$$
PD(t_0) = \mathbb{P}\left[V_{PPA}(t_0) \leqslant 0 \right] \tag{24.2}
$$

$$
PD(t_k) = \mathbb{P}\left[V_{PPA}(t_k) \leqslant 0 \mid V_{PPA}(t_{k-1}) > 0 \right] \quad \text{for } k = 1, \dots, M-1. \tag{24.3}
$$

This enables us to compute the collateral requirements the producer might demand. For that, it is supposed that at each time t_k the deemed loss will be equivalent to the investment not hedged, that is, the capital expenditure not amortized. The expression for the expected loss is

$$\begin{aligned} EL(t_k) &= PD(t_k) \cdot CAPEX^{not}_{amortized}(t_k) \\ &= PD(t_k) \cdot (CAPEX(t_k) - AMORTIZATION(t_k)), \end{aligned} \tag{24.4}$$

where $CAPEX$ refers to the capital expenditure, and the Total Expected Loss (TEL) will be

$$TEL(t_0) = \sum_{k=0}^{M-1} e^{-r(t_k - t_0)} EL(t_k). \tag{24.5}$$

2.2 Electricity Price model

Given that the electricity market price is the underlying variable of the model, it is necessary to set a model which describes the main features of the electricity market price such as the seasonality and the mean reversion. We have chosen the following model

$$P_t = e^{f(t) + Y_t}, \tag{24.6}$$

where $f(t)$ is the seasonality function and Y_t is the stochastic part of the price process, described respectively by a sinusoidal function and the Ornstein-Uhlenbeck process:

$$f(t) = a_0 + \sum_{i=1}^{4} a_i \cos(2\pi\gamma_i (t - \tau_i)) \tag{24.7}$$

$$dY_t = \alpha(\mu - Y_t)\, dt + \sigma dW_t, \quad Y_{t_0} = \log P_{t_0} - f(t_0). \tag{24.8}$$

3 Results and discussion

In order to shed some light into the valuation, a numerical simulation was carried out. Furthermore, we noticed that in Spain a toll is added up to the electricity market price, that is

$$P_t = \widehat{P}_t + \xi,$$

where ξ denotes this toll.

3.1 Numerical settings

The variables of the model were chosen as the actual values used in a RE PPA designed for a photovoltaic irrigation project in Spain, for which the offtaker was determined that s/he will need $300\, MW \cdot h$ in a yearly basis. For that purpose, the producer establishes that a $200\, kWp$ system will be required, which implies a capital expenditure of 244960 €. The duration of the agreement is 20 years and payments will take place annually, $t_k = 0, 1, \ldots, 19$. The LCOE of that system has been estimated as $76.69 \frac{€}{MW \cdot h}$ and an interest rate of 2% was considered. The $AMORTIZATION$ variable of the Equation (24.4) is equal to 11484 € for years $1 - 10$ and 13012 € for years $11 - 20$ Carrêlo et al. (2020). On the other hand, the parameters of the seasonality part of the electricity price model were estimated by a least squares method (Nelder-Mead optimization algorithm) and the stochastic part by means of the maximum likelihood method using daily data of the electricity price from 2010 to 2022 in Spain. γ_i, for $i = 1, \ldots, 4$, was obtained analysing the periodogram of the raw data. Below is shown a summary of the parameters

Table 1: Estimated parameters of the seasonality function and stochastic process.

	i				
	0	1	2	3	4
a_i	4.001137	-0.187542	8.264600	-0.079046	0.637311
τ_i	–	-0.046877	0.100468	0.024929	0.647611
γ_i	–	1	$\frac{365}{7}$	$\frac{365}{7/2}$	$\frac{365}{7/3}$
	α	μ	σ	$\hat{P}_{t_0}$	ξ
	39.105053	0.004586	2.180066	143.17	20.00

3.2 Results

According to the methodology proposed, and given the numerical settings presented, for a $P_{PPA} = LCOE = 76.69\ \frac{€}{MW \cdot h}$ the expected value of the PPA at t_0 is $\mathbb{E}\left[V_{PPA}(t_0)\right] = 66630.84\ €$, with a 95% CI equal to $(65721.46, 67540.23)$. A comparison using a swap-based approach was made so as to clarify the performance of our model. A swap-based PPA value may be computed as Edge (2015)

$$V_{PPA}\left(t_k\right) = \sum_{i=k}^{M} U(t_i)\left(P_{t_i} - P_{PPA}\right) e^{-r(t_i - t_k)}, \quad \text{for } k = 0, \ldots, M. \tag{24.9}$$

As a result, our proposed model provides PPA values greater than the first difference, that is $U\left(P_{t_0} - P_{PPA}\right)$, even for P_{PPA} values considerably higher than the $LCOE$ level. On the other hand, the swap approach provides negative values for a P_{PPA} value greater than $89\ \frac{€}{MW \cdot h}$.

For the other financial measurement, which determines the amount, in case of early termination, of the CAPEX that will not be repaid, the $TEL(t_0)$ amounts to $7448.69\,€$ with a 95% CI equal to $(6103.72, 9258.43)$. A comparison with the swap-based approach has also been conducted and is plotted in Fig. 1, where we can observe that by applying our model, the amount required to hedge against default is lower than for the swap approach, indeed for a $P_{PPA} \in [70, 95]$, the collateral requirements are from 76% to 9% lower. Furthermore, for a P_{PPA} value equal to the $LCOE$ level in a swap-based approach the collateral requirements amount to $18362.33\,€$ with a 95% CI equal to $(16091.37, 21093.69)$.

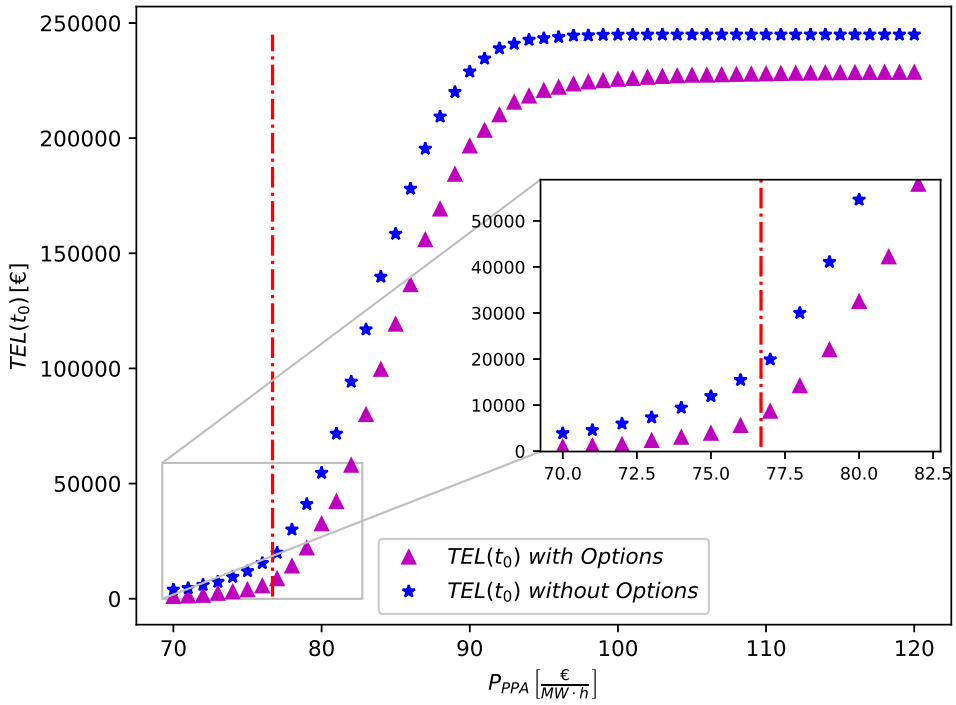


Figure 1: $TEL\,(t_0)$ assuming a RE PPA without optional value (blue stars) and $TEL\,(t_0)$ including the options (purple triangles) for different levels of P_{PPA}. Red-dotted line represents the LCOE value.

4 Conclusions and future research/work

We have found that the proposed model for the PPA valuation describes some features which were usually neglected in the modelling, such as the embedded options. Furthermore, it also tackles the specific terms of a PPA including the price and amount of energy to be supplied and the alternative source of electricity supply, the public grid, which in our case is the underlying variable. Once this model is given, the probability of default can be computed, which turns out to be the probability that the value of the PPA becomes negative.

Albeit the results depend on the electricity price model, we think that by plugging an alternative electricity price model the same methodology can be applied and the same kind of results may be obtained.

Future research related to this work should be focused on building a model which considers the agreed price as non-fixed, with payments taking place in a monthly basis and even a floating consumption.

Acknowledgements

CITIC is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS).

Bibliography

I. B. Carrêlo, R. H. Almeida, L. Narvarte, F. Martinez-Moreno, and L. M. Carrasco. Comparative analysis of the economic feasibility of five large-power photovoltaic irrigation systems in the mediterranean region. *Renewable Energy*, 145:2671–2682, 2020.

P. Edge. An approximation of counterparty credit risk in long term power purchase agreements (PPAs). In *2015 12th International Conference on the European Energy Market (EEM)*, pages 1–4, Lisbon, Portugal, 2015. IEEE.

J. M. Freeman, N. A. DiOrio, N. J. Blair, T. W. Neises, M. J. Wagner, P. Gilman, and S. Janzou. System Advisor Model (SAM) general description (version 2017.9. 5). Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States), 2018.

P. Gabrielli, R. Aboutalebi, and G. Sansavini. Mitigating financial risk of corporate power purchase agreements via portfolio optimization. *Energy Economics*, 109:105980, 2022.

J. I. Peña, R. Rodriguez, and S. Mayoral. Modeling Renewable Power Purchase Agreements Prices. *Available at SSRN*, 2022.

Embedded System for Monitoring and Prevention of Hardware/Software Incidents in Data Center Devices

Alberto Pérez-Aboli, Carlos Dafonte, and Ángel Gómez

Faculty of Computer Science, Universidade da Coruña, 15071 A Coruña, Spain
Centro de Investigación CITIC - Laboratorio Interdisciplinar de Aplicaciones de la Inteligencia Artificial (LIA2), Faculty of Computer Science, Universidade da Coruña, 15071 A Coruña, Spain
Centro de Investigación CITIC - Laboratorio Interdisciplinar de Aplicaciones de la Inteligencia Artificial (LIA2), Faculty of Computer Science, Universidade da Coruña, 15071 A Coruña, Spain
Correspondence: alberto.aboli@udc.es; carlos.dafonte@udc.es; angel.gomez@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: Real-time monitoring of the status of the devices available in Data Centers (DC) is a critical task to try to provide quality services to customers, but also to detect possible hardware/-software incidents or to act against cyber-attacks. To perform these functions, there are specific system and network monitoring tools on the market that supervise different device functionalities (CPU, memory, storage, network, processes, services, etc.). However, these systems do not include the monitoring of basic elements for the correct operation of these devices, such as environmental parameters (power supply, temperature, humidity, noise, etc.). In this project we have designed and built a prototype embedded system, both hardware and software, to monitor in real time each device inside a rack and capable of integrating a wide variety of sensors. The hardware device is composed of: motherboard, power supply, infrared temperature sensor, noise level sensor, current consumption meter, 220VAC real input voltage meter, electronic module to switch on/off the 220VAC output and WiFi module to send the information.

1 Introduction

A data center or DC is the facility that centralizes the IT infrastructures of an organization. Within these infrastructures are the most important and critical devices for providing services to customers (Web, storage, computation, networks, etc.). The DCs are large facilities with high energy consumption due to the fact that they gather a large number of equipment in the smallest possible space, where temperature and cooling become important in order to maintain the DC in optimal working conditions. Having a more efficient and exhaustive control of the most critical parts of a data center is vital for any corporation.

For the software monitoring of devices and services there are many widely used monitoring tools, among them we can highlight open source solutions such as Nagios (Nagios, 2023) or Zabbix (Zabbix, 2023). This type of tool collects information from each piece of equipment by means of software agents installed in them. The information received is stored and processed in real time for proper display on the application's web interface. In this way, it is possible to see graphically any anomaly or incident detected in the operation of the monitored devices.

For hardware monitoring of the environment there are specific commercial products such as: humidity and temperature sensors (NetBotz APC, 2023a), vibration sensors (NetBotz APC, 2023b), integrated PDU (Power Distribution Unit) systems for voltage control in racks (NetShelter APC, 2023). These types of devices provide certain advantages but also have important limitations: they are closed products, their cost is high, and multiple devices are required to cover different types of measurements.

This project proposes the creation of an embedded device that can monitor the components of a DC individually by means of various types of sensors. In addition, this device will be designed following a modular architecture to facilitate future extensions, should be as small as possible and with a reduced manufacturing cost. The control of each of the equipment will be done in real time and in a non-invasive way: it will be connected directly using the base of the power input socket of the equipment to be monitored.

The system will provide relevant data for data center administrators such as measurements of the noise generated by the equipment's fans, changes in the temperature of the enclosure compared to that of its immediate surroundings, changes in consumption and voltage changes occurring in the 220VAC network. The monitoring of this information, collected in real time, will help in the management and prevention of incidents. The detection of anomalous operation such as changes in power consumption, unusual temperatures, strange noises or vibrations could reveal information about what is happening with the status of the device being monitored.

This device would also help to study the profitability of the equipment, when a device starts to consume too much energy and overheat due to wear and tear, it stops performing as expected at the computational level. This information would help in making decisions for equipment renewal.

2 Project description

Creating a device with the indicated characteristics is not a simple task; it requires both hardware and software design, involving the use of numerous professional tools and equipment to design and manufacture the hardware itself, which will be developed at each stage.

The hardware objective of this project is to create an electronic device that is easy to install, non-invasive and as simple as possible to set up. This device will send critical measurement data in real time on the status of each of the devices/servers within a data center. Among the functionalities that will be available, we can highlight:

- External infrared temperature sensor, which will provide the temperature of the case of the device to be monitored and the temperature of its immediate surroundings.
- Voltage sensor, which will measure the actual voltage at all times at the 220VAC input.
- Current consumption sensor, which will control the actual consumption in amperes (A) of the device to be monitored, including the consumption of the monitoring device itself.
- Microphone of very small size, which will measure the noise/vibration of the case of the device to be monitored, with the idea of capturing the noises of the fan bearings.
- Microcontroller with WiFi, which will be the core of this embedded system, will be in charge of the computational part and of sending telemetry data via WiFi.

The software objective of this project is to obtain all the monitoring data from the aforementioned sensors and send them via WiFi, so that each user can decide how to integrate them into their control system. Specifically, we can cite the following objectives:

- Use a free software focused on IoT to manage this device.
- Develop custom firmware for the sensors of each device module.
- Access the device via IP through the Web browser, using a Web interface; the device can be configured and the monitoring data can be viewed at the same time as it is sent via

WiFi. In the Web interface, the user can configure certain parameters such as the data of his MQTT or HTTP server to receive the data.

3 Development

The main device of the embedded system is the processing unit, a device with built-in WiFi has been selected, all in a single component called SoC. The SoC model selected for our system was the Espressif ESP32-S3-WROOM-1U, 32-bit MCU 2.4 GHz Wi-Fi Bluetooth 5 (LE) dual-core (Espressif, 2023). It offers high performance at a very competitive price. In addition, there are open source solutions for IoT applications that can be used with this SoC type device. The Expressif ESP-S3-DevKitC-1U development board (Figure 1) contains the chosen SoC and will be used to support design development.

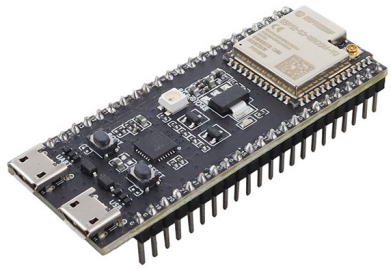


Figure 1: Expressif ESP-S3-DevKitC-1U development board.

In this project we need to implement a series of modules with the corresponding sensors for each functionality. Some of them are the following (Figure 2): a sensor to measure the current consumption in amperes (model ACS712-20A [1]), a sensor to measure the actual value of the AC voltage at the input (model ZMPT101B [2]), a sensor to measure the noise level (model MAX9814ETD+T [3]), an infrared temperature sensor (model MLX90614 [4]).

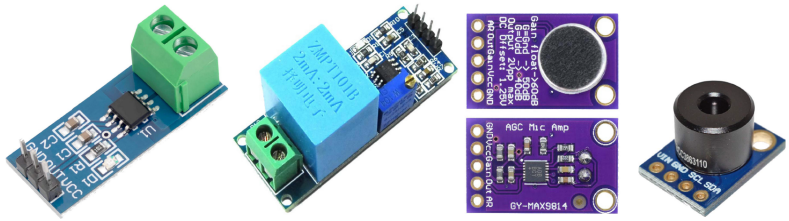


Figure 2: Selected sensor development boards and models.

In summary, the final hardware device will be composed of: motherboard, power supply, infrared temperature sensor, noise level sensor, current consumption meter, actual 220VAC input

[1] Allegro ACS712-20A, datasheet available at: *https://www.farnell.com/datasheets/1759100.pdf*

[2] InnovatorsGuru Módulo-ZMPT101B, datasheet available at: *https://www.circuits-diy.com/zmpt101b-ac-single-phase-voltage-sensor-module*

[3] ANALOGDEVICES MAX9814ETD+T, datasheet available at: *https://www.mouser.es/datasheet/2/609/MAX9814-3131693.pdf*

[4] MELEXIS MLX90614, datasheet available at: *https://www.farnell.com/datasheets/3959635.pdf*

voltage meter, 220VAC output on/off electronic module and WiFi module to send the information. During the development of the device it is very important to carry out an adequate design in order to place the different components in such a way that they occupy the smallest possible size without affecting its correct operation. Figure 3 shows the final design of the device and an image of the prototype in the firmware programming phase (Figure 4).

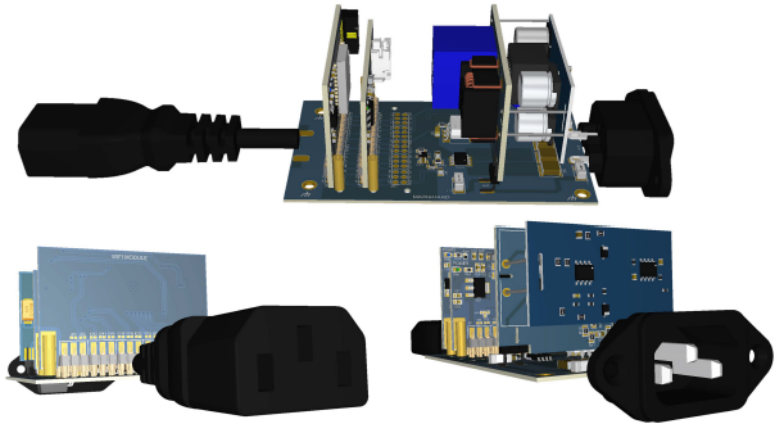


Figure 3: Hardware device design.

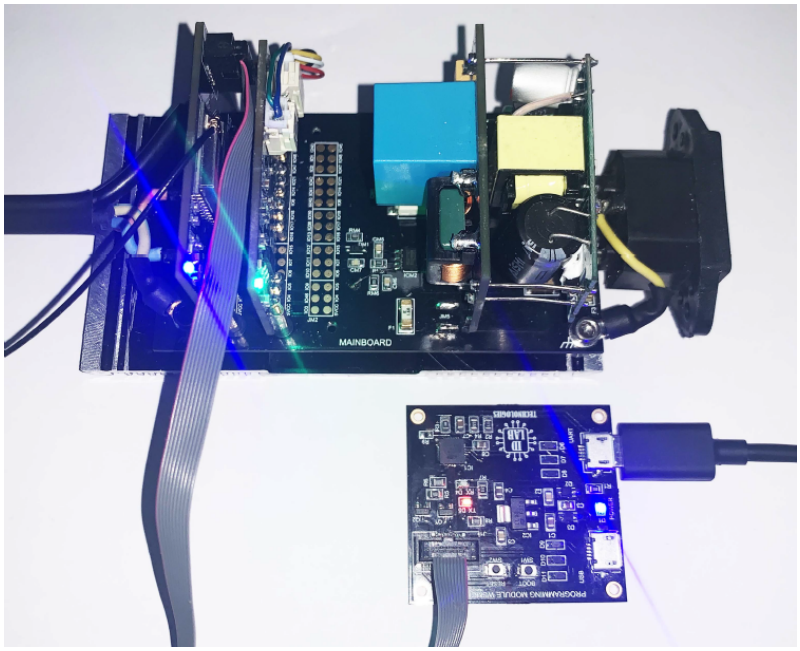


Figure 4: Final prototype in firmware programming phase.

For the development of the system software, we chose to use the open source Tasmota firmware (Tasmota, 2023), specially designed for microcontrollers manufactured by Espressif. Tasmota offers many possibilities of customized configuration, starting from the default

configuration you can make modifications in the code to adapt it to your needs and generate the compiled firmware for your project. Among the functionalities it provides is the Web interface (WebUI), which allows access to the device via IP from the Web browser. Using Tasmota's WebUI you can customize some parameters, configure the WiFi or send monitoring data for centralization in environments such as Home Assistant or Zabbix. To control Tasmota devices the MQTT protocol (MQTT, 2023) is used, it is a standard for IoT messaging that works with publishings and subscriptions to a topic through a broker or agent that manages those publishings and subscriptions.

4 Conclusions

This paper presents an embedded system capable of monitoring the environmental parameters of the equipment in a data center individually and assisting technical personnel in the prevention of incidents. It is designed as a low-cost hardware device, small size and modular architecture, allowing future expansion and scalability. The modules that compose it are of expansion type, being able to remove, add or modify modules by configuring the GPIOs to be used. To install the developed device it is only necessary to insert the device between the two input and output sockets of the equipment to be monitored. The information collected by each device will be sent via WiFi to the DC monitoring server. Additionally, through a Web interface included in the firmware of the device, it is possible to access the configuration or visualize the monitoring information of each equipment in real time.

The preliminary results obtained have been very positive, but the implementation and testing in a real controlled environment is proposed as a line of future work, which would allow the results to be studied and compared with other measurement tools.

Acknowledgements

CITIC is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS).

Bibliography

Espressif. Espressif Systems: ESP32-S3-DevKitC1U. *https://docs.espressif.com/projects/esp-idf/en/latest/esp32s3/hw-reference/esp32s3/user-guide-devkitc-1.html*, 2023. [Online; accessed 1-September-2023].

MQTT. MQTT: The Standard for IoT Messaging. *https://mqtt.org*, 2023. [Online; accessed 1-September-2023].

Nagios. Nagios Open Source. *https://www.nagios.org/*, 2023. [Online; accessed 1-September-2023].

NetBotz APC. NetBotz Sensors: APC Temperature & Humidity Sensor. *https://www.apc.com/us/en/product/AP9335TH/apc-temperature-humidity-sensor*, 2023a. [Online; accessed 1-September-2023].

NetBotz APC. NetBotz Sensors: NetBotz Vibration Sensor - 12 ft. *https://www.apc.com/us/en/product/NBES0306/netbotz-vibration-sensor-12-ft-/*, 2023b. [Online; accessed 1-September-2023].

NetShelter APC. NetShelter: APC Rack PDU, 2G, metered, 0U, 30A, 200V and 208V, 36 C13 and 6 C19 sockets. *https://www.apc.com/us/en/product/AP8841/apc-rack-pdu-2g-metered-0u-30a-200v-and-208v-36-c13-and-6-c19-sockets/*, 2023. [Online; accessed 1-September-2023].

Tasmota. Tasmota: Open source firmware for ESP devices. *https://tasmota.github.io/docs/*, 2023. [Online; accessed 1-September-2023].

Zabbix. Zabbix - The Enterprise-Class Open Source Network Monitoring Solution. *https://www.zabbix.com/*, 2023. [Online; accessed 1-September-2023].

Integration of Captive Movement Systems in Virtual Environments for Rehabilitation in People with Disabilities

Manuel Lagos Rodríguez, Nereida Canosa, and Thais Pousada

TALIONIS Research Group, CITIC, Universidade da Coruña, 15071 A Coruña, Spain
Correspondence: m.lagos@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: Generally, virtual reality (VR) systems used to improve physical and cognitive skills in people with disabilities generate a limited play area, allowing only a few meters of movement. This work proposes the integration of the Virtuix Omni Pro device, a captive movement system that enables unlimited displacement. This represents an opportunity in VR research for gait training in people with disabilities and adds value in terms of realism and immersion for the practice of activities in which movement is essential, such as shopping in a supermarket or moving freely through an urban environment.

1 Introduction

Virtual Reality (VR) is a technology capable of generating highly realistic virtual environments. Depending on the technology used, a great immersion of the person in these environments can be achieved, causing the sensation of being present in them. Most VR equipment consists of helmets or glasses with stereoscopic vision, controls, and a set of sensors capable of tracking the person's movements. It should be noted that the tracking is limited by the gaming area generated by the device and by the actual space of the room in which it is used.

VR is a motivating, safe, and entertaining means in the process of improving physical and cognitive skills in people with disabilities (Lagos Rodríguez et al., 2022). The existing literature has studied the use of this technology in people with different health situations that limit their autonomy (Chau et al., 2021), (Domínguez-Téllez et al., 2019), (Lee and Jin, 2023). The results point to further research in this line and the development of new VR applications that contribute to improving the daily lives of people with disabilities.

This work is intended to continue this line of research. As mentioned above, VR equipment has a limited play area, which means that the person can only move a few meters. In addition, the actual space itself is also a barrier in this regard. This can be a limitation when working on skills that require complete freedom of movement. Furthermore, being able to move freely through a virtual scenario significantly increases the realism and immersion of the environment. As a solution, a captive movement system has been integrated to allow unlimited movement in virtual environments.

A virtual reality application that simulates a supermarket is also being developed. In it, different daily activities can be carried out that will allow both physical and cognitive work. The fact of working with day-to-day activities can be more enjoyable and motivating for the person compared to other more mechanical or repetitive exercises. In addition, the exercises performed have a direct application in your life.

Therefore, the objective of this work is to integrate a captive movement system in a virtual environment that simulates a supermarket. The final application will consist of three very common activities in a real supermarket and will allow the work of physical and cognitive skills in people with disabilities.

2 Materials and Methods

The project has started with several meetings with organizations in the field of disability. The researchers have presented the VR equipment, as well as the captive movement system, and different tests have been carried out with healthcare professionals and people who need to work on some physical or cognitive ability. These meetings have made it possible to confirm the security of the platform and to define the virtual environment, as well as the tasks to be performed in it.

Subsequently, the development of the tool was started. The first step was the search for a 3D model of a supermarket and its products and equipment. The selection process considered the completeness of the product catalog, the realism of the 3D objects, and their configuration options.

Once the 3D models have been acquired, they have been integrated into a virtual scene. The placement of the objects in each scenario has been done considering the tasks to be programmed in the scenario. The aim is to obtain a realistic, intuitive, and motivating scenario for the user.

After obtaining the first version of the virtual supermarket, the interaction system has been implemented and the captive movement platform has been integrated. It is worth mentioning that the interaction is done through a hand-tracking system, which avoids the use of remotes.

Finally, the programming of the set of activities defined in the meetings held at the beginning of the project was started. Thus, the objective is to develop three activities that allow the work of both physical and cognitive skills.

In addition to the tests carried out by the research team, the organization of test sessions with professionals and users of different entities in the field of disability is proposed. The purpose will be to detect possible errors and implement improvements that will contribute to increasing the usability of the tool.

Below, the hardware and software elements that have been selected for the development of the project are briefly defined:

2.1 Hardware

- **HTC VIVE Pro 2:** It is a virtual reality equipment created by HTC. It consists of a kit made up of glasses, controllers, and position sensors. The glasses have a screen that allows the visualization of 3D virtual environments and a high-quality audio system that increases the immersion capacity (HTC VIVE, n.d.).

- **Leap Motion:** It is an optical hand and finger tracking system. It allows interaction with digital content with the hands, avoiding the use of a keyboard, mouse, or controller. It is possible to integrate it into a virtual reality system, providing greater realism by being able to touch and grasp virtual objects (Leap Motion, n.d.).

- **Virtuix Omni Pro:** It is a captive movement system. It consists of a sliding surface and a harness. It transforms the movement of the feet into movement in the virtual environment (Virtuix Omni Pro, n.d.) (see Fig. 1).

Figure 1: Virtuix Omni Pro

2.2 Software

- **Unity:** It is a real-time 3D content development platform. It allows the creation of interactive environments for a wide variety of devices such as computers, consoles, or VR equipment (Unity, n.d.).

- **OpenXR:** It is a free and open API standard developed by Khronos. It facilitates the development of VR applications compatible with VR devices from different manufacturers (OpenXR, n.d.).

- **Unity Leap Motion Modules:** It is a set of libraries provided by the Leap Motion manufacturer to access the device API and to facilitate the development of functionalities in the VR application (Leap Motion Modules, n.d.).

3 VR Application

The application consists of a virtual supermarket with a high degree of immersion. It has a wide catalog of products, which allows it to carry out a variety of exercises related to cognitive training. Likewise, the environment has the usual equipment of real supermarkets, with cash registers, shelves at different heights, fridges, and freezers with different types of openings, and baskets and trolleys to store the shopping. This is noteworthy because it facilitates the implementation of tasks for physical work while contributing to greater realism in the training process. Also noteworthy is the inclusion of audio and image elements to motivate the person and guide him/her in the completion of each test.

The development of three activities with tasks that allow training of physical and cognitive skills in people with disabilities is proposed. Each of them is described below:

- **Activity 1:** The person must pick up products at different heights and place them in a basket. The objective is to work on the mobility of the upper and lower limbs when reaching each product and placing it inside the basket (see Fig. 2).

Figure 2: Activity 1

- **Activity 2:** The person will have a shopping list and must select only the correct products and place them in a trolley. This task is intended to work on memory and concentration (see Fig. 3).

Figure 3: Activity 2

- **Activity 3:** The person must make the payment for the purchase made. To do this, the person will have bills and coins of different values. Thus, he/she must check the amount on a screen and select the appropriate bills and coins until the correct amount is added up. The objective is to train mental agility (see Fig. 4).

Figure 4: Activity 3

4 Conclusions

The integration of a captive motion system provides a more realistic and natural way of carrying out certain activities that require moving through a virtual environment. In addition, it is an added value in VR research for gait training in people with disabilities.

The programming of tasks based on daily life favors the transfer of the results obtained to the person's life. In addition, it could also imply a better adherence to the prescribed exercises.

Acknowledgments

This research had the support for its publication by the project TED2021-130127A-I00, funded by MCIN/AEI/10.13039/501100011033 and by the European Union "NextGenerationEU". CITIC is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS). The research carried out by TALIONIS group is financed by the Xunta de Galicia (Aid from the Consellería de Cultura, Educación, Formación Profesional e Universidades for the consolidation and structuring of competitive research units) ED431B 2022/39.

Bibliography

P. H. Chau, Y. Y. J. Kwok, M. K. M. Chan, K. Y. D. Kwan, K. L. Wong, Y. H. Tang, K. L. P. Chau, S. W. M. Lau, Y. Y. Y. Yiu, M. Y. F. Kwong, W. T. T. Lai, and M. K. Leung. Feasibility, acceptability, and efficacy of virtual reality training for older adults and people with disabilities: Single-arm pre-post study. *Journal of Medical Internet Research*, 23(5), 2021.

P. Domínguez-Téllez, J. Moral-Muñoz, E. Casado-Fernández, A. Salazar, and D. Lucena-Antón. Efectos de la realidad virtual sobre el equilibrio y la marcha en el ictus: revisión sistemática y metaanálisis. *Revista de Neurología*, 06(69):223–234, 2019.

HTC VIVE, n.d. URL *https://www.vive.com/eu/product/vive-pro2-full-kit/overview/*. [Online; accessed september 2023].

M. Lagos Rodríguez, A. Gómez García, J. Pereira Loureiro, and T. Pousada García. Personalized virtual reality environments for intervention with people with disability. *Electronics*, 11(10), 2022.

Leap Motion, n.d. URL *https://www.ultraleap.com*. [Online; accessed september 2023].

Leap Motion Modules, n.d. URL *https://developer.leapmotion.com/unity*. [Online; accessed september 2023].

H. K. Lee and J. Jin. The effect of a virtual reality exergame on motor skills and physical activity levels of children with a developmental disability. *Research in Developmental Disabilities*, 132, 2023.

OpenXR, n.d. URL *https://www.khronos.org/openxr*. [Online; accessed september 2023].

Unity, n.d. URL *https://unity.com*. [Online; accessed september 2023].

Virtuix Omni Pro, n.d. URL *https://www.virtuix.com/developer*. [Online; accessed september 2023].

A System for the Management, Visualization and Annotation of Digital Collections of Handwritten Historical Documents

Carlos Andrés Correa-Guillén, Ángel Gómez, and Estefanía López-Salas

Faculty of Computer Science, Universidade da Coruña, 15071 A Coruña, Spain
Centro de Investigación CITIC - Laboratorio Interdisciplinar de Aplicaciones de la Inteligencia Artificial (LIA2), Faculty of Computer Science, Universidade da Coruña, 15071 A Coruña, Spain
Grupo de Estudios Territoriales (GET), Faculty of Architecture, Universidade da Coruña, 15071 A Coruña, Spain
Correspondence: carlos.guillen@udc.es; angel.gomez@udc.es; estefania.lsalas@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: Today, there are huge amounts of handwritten documents accessible in digital collections. However, working with these documents for research purposes is limited due to the lack of appropriate technological tools to read, transcribe and analyze them in the same digital environment. This paper presents a web-application that allows a dynamic interaction with a particular collection, the so-called General Answers of the Cadaster of Ensenada. The app was designed to give access and support the analysis of these historical handwritten documents that are crucial to study the natural and human environment of the 18^{th}-century Spain. The app allows parameterized searches, the visualization and annotation of digitized documents, and its integration with other applications through an API. In addition, it includes a functionality to segment and classify handwritten numbers automatically.

1 Introduction

Spanish digital archives and libraries, such as the Portal of Spanish Archives (PARES),[1] the Virtual Library of Defense (BVD),[2] the Virtual Library of Documentary Heritage (BVPB),[3] or the Hispanic Digital Library (BDH), [4] contains a huge amount of digital documents, from several time periods, that anyone interested can freely access through the corresponding online platforms. The amount of those sources is, in fact, growing rapidly in the last decades. However, its exploitation is still highly limited in the web-environment due to several technical challenges. We can read and explore each document, but it is still not possible to add information over them

[1] Ministry of Culture and Sport - Spain. Portal of Spanish Archives (PARES), accessible at: *https://pares.culturaydeporte.gob.es/inicio.html*.

[2] Ministry of Defense - Spain. Virtual Library of Defense (BVD), accessible at *https://bibliotecavirtual.defensa.gob.es/BVMDefensa/es/inicio/inicio.do*.

[3] Ministry of Culture and Sport - Spain. Virtual Library of Documentary Heritage (BVPB), accessible at: *https://bvpb.mcu.es/es/inicio/inicio.do*.

[4] National Library of Spain. Hispanic Digital Library (BDH), accessible at: *https://www.bne.es/en/catalogues/hispanic-digital-library*.

or to share it with others to favor, for instance, collaborative research based on data coming from historical documentation.

One rich collection accessible in digital archives is the General Answers of the Cadaster of Ensenada conducted in the mid-18^{th} century in the old Crown of Castile. This digital collection, which is accessible at *https://pares.mcu.es/Catastro/*, contains the complete copy of the General Answers that is preserved in the General Archive of Simancas currently, with the records gathered from 13.000 places (Ministry of Culture and Sport – Spain, 2023).

The General Answers were the first part of the works carried out within the Cadaster of Ensenada. This was a large-scale census and statistical investigation that aimed to implement a single tax in the Crown of Castile, that is, a fiscal reform to force each person living into the Castilian territory to pay according to their possessions (Alimento, 2002). In order to achieve that goal, it was necessary to make a register of all the properties as well as a deep on-site study of each of them, but firstly, the cadaster officers carried out a survey with 40 common general questions at each locality. The findings or answers to these questions were gathered in books, called General Answers. One book was generated per each locality of the old Crown of Castile. As a result, it was produced a huge volume of documentation that allow us to approach many different issues related to population, society, economy, or geography, to cite a few, from mid-18^{th}-century Spain (Camarero Bullón, 2002). In this paper we present the web-application Archive Lens developed to facilitate the analysis of that historical documentation by researchers through an effective management, dynamic interaction, advanced search and annotation capabilities that, on the whole, aims to contribute to the enrichment of shared knowledge.[5]

2 Conceptualization

The proposed system was developed to ease the work when a particular theme is investigated, but through the whole collection of the General Answers. In fact, its conceptualization was born in direct relation with the in-progress work within the project Mapping Hospitals.[6] The aim of this project is to investigate the network of hospitals that existed in Galicia in the mid-18^{th} century through the Cadaster of Ensenada.

The 30^{th} question of the cadaster survey was focused on gathering data about the existing hospitals in each locality. Therefore, if we pay attention to this particular question, we have the opportunity to approach the past network of hospitals, hospices, inns or rest houses that were used mainly for pilgrims on travel to Santiago de Compostela, but also by the poor and the ill.

However, the project has some limitations that derived mainly from the huge amount of records to be reviewed. So far, we have only approached the network of hospitals corresponding to Galicia and this work implied the reading of the 30^{th} question from a total of 2912 records (López Salas, 2021). Therefore, to have a tool to ease the management, visualization, annotation and search through these historical documents will reduce the cost and time needed to broaden the scope of the study to other territories of the old Crown of Castile, but also to proposed new fields of research that crosscut large amounts of cadaster records.

3 Description of the application and its functionalities

The application Archive Lens is organized into three main levels: (1) Map; (2) Selector of localities, and (3) Visualization and annotation of documents. The first level, the Map, presents a graphic representation of the geographic area that was affected by the cadastral process in the mid-18^{th} century. In this level, the user is able to interact with the map and to select a province in which a number of historical documents were created (Figure 1).

---

[5] This paper is the result of the work done as a Final Degree Project by Carlos Andrés Correa-Guillén, under the supervision of Ángel Gómez García and Estefanía López Salas, and presented at the Faculty of Computer Science, University of A Coruña, in July 2023 (Correa-Guillén, 2023).

[6] Mapping Hospitals Project, accessible at: *https://mappinghospitals.udc.es/index_es.html*.

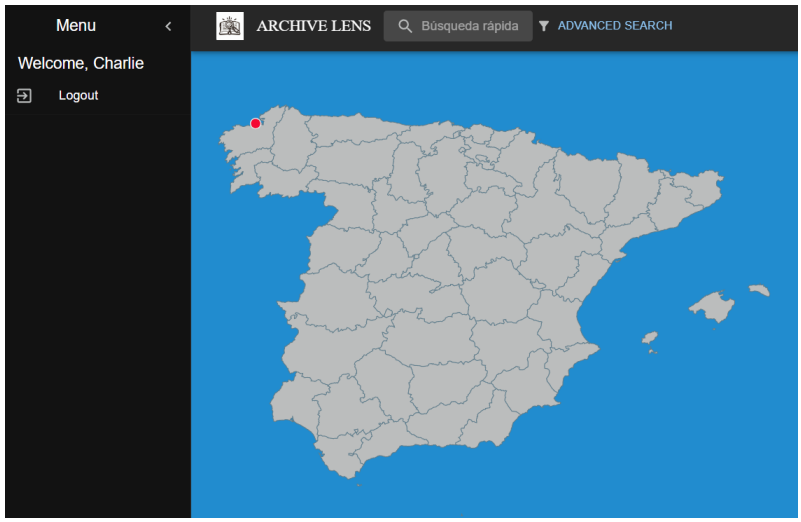


Figure 1: The Application Archive Lens: the map.

When a province is selected, the second level of the application is opened for the selection of a specific locality within the whole province (Figure 2). Each locality is presented as an independent card with information about: the present name of the locality, the name of the locality in the mid-18th century, a short description of the place, a representative image, and the total number of documents that it comprises. These documents are the digitized pages of the General Answers corresponding to each locality of the province. When a specific locality is selected by the user, he/she can access the third and last level of the application, which is devoted to a tool for the visualization and annotation of historical documents (Figure 3).

Figure 2: The Application Archive Lens: the selector of localities.

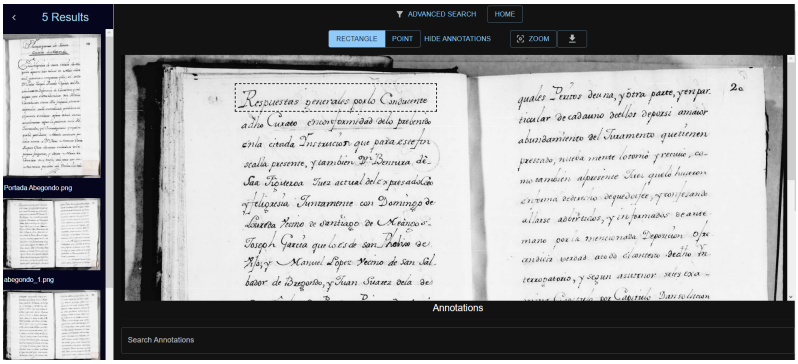


Figure 3: The Application Archive Lens: the visualization and annotation tools.

In the third and last level, the user is able to interact with each digitized page of the General Answers generated for a specific locality. We included traditional tools such as zoom in and zoom out, but also new ones to facilitate the interaction of the user with the handwritten text. In the upper part of the Visualization tool, three buttons are displayed to allow annotations over the historical documentation. The first two buttons, called Rectangle and Point, indicate the two types of annotations the user is able to create. The type called Rectangle allows the user to select a rectangular area of the document to be annotated, for instance, with the corresponding transcription of the historical text. The second type of annotations is thought to add Points to specific parts of the documents where the user may need, for instance, to include a comment for another researcher working with the same locality. Therefore, the annotation tools were designed to facilitate an effective interaction between the reader and the archival documentation while enriching the information that it already contains in a new digital layer.

Another important functionality that was explored during the development of the present work is directly related to improving the search and filter of specific data within this cadaster. Based on the fact that each book of the General Answers is organized with the same structure of 40 questions, we started with the implementation of a functionality to segment and classify handwritten numbers automatically. We applied the TrOCR model that consists of "an image Transformer encoder and an autoregressive text Transformer decoder to perform optical character recognition (OCR)" (Li et al., 2022). The results are shown in Figure 4.

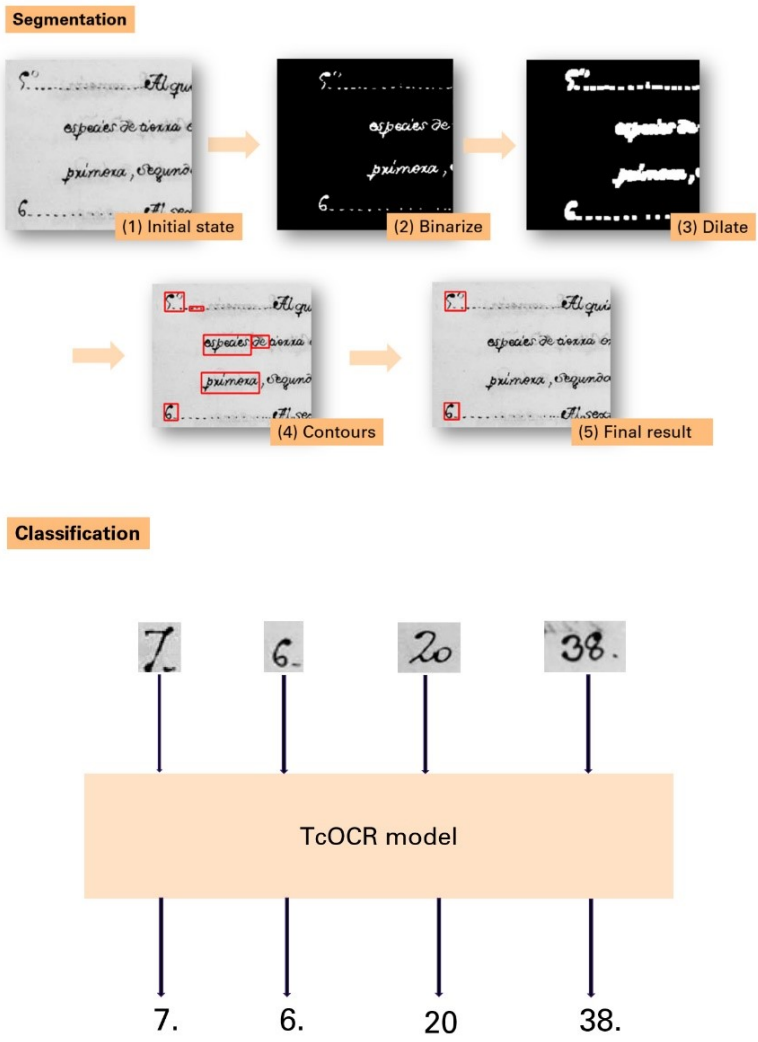


Figure 4: Summary of process for segmentation and classification of handwritten numbers in the Cadaster of Ensenada.

4 Conclusions

The application presented in the paper offers a set of advantages to work with historical handwritten documents as it integrates visualization and annotation capabilities that make easier the reading, non-automatic transcription, and collaborative research in a web-environment. Although it was designed for the Cadaster of Ensenada, it could be easily adapted to other historical documentation. We have opened new opportunities, but the word done also poses some challenges to be faced in future developments. For instance, we still need to improve the management system of users in order to be able to assign roles and verify permissions. This way, we will offer the possibility to promote discussions through annotations and also customization of each user type. In addition, the handwritten text recognition system is only focused on question numbers, but it will be really useful for historical research to explore how to apply it

to the whole text in order to allow more specific and insightful searches in the rich content of this cadaster of the mid-18^{th} century Spain.

Acknowledgements

CITIC is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS).

Bibliography

A. Alimento. Los catastros del siglo XVIII, entre tradición y modernidad. *CT: Catastro*, 46:17–26, 2002.

C. Camarero Bullón. El Catastro de Ensenada, 1745-1756 diez años de intenso trabajo y 80.000 volúmenes manuscritos. *CT: Catastro*, 46:61–68, 2002.

C. A. Correa-Guillén. *Web tool for structured viewing and online annotation of historical manuscripts*. Final Degree Project. Faculty of Computer Science, University of A Coruña, A Coruña, 2023.

M. Li, Lv, Tengchao, Chen, Jingye, Cui, Lei, Lu, Yijuan, Florencio, Dinei, Zhang, Cha, Li, Zhoujun, and Wei, Furu. TrOCR: Transformer-based Optical Character Recognition with Pretrained Models. *Arxiv. Computer and Language*, pages 1–10, 2022.

E. López Salas. Cartografía de la hospitalidad en los caminos de peregrinación: el proyecto Mapping Hospitals. *Anales de Geografía Complutense*, 2(41):389–407, 2021.

Ministry of Culture and Sport – Spain. General Answer of the Cadaster of Ensenada. *https://pares.mcu.es/Catastro/*, 2023. [Online; accessed 1-September-2023].

Understanding Machine Learning Explainability Models in the context of Pancreatic Cancer Treatment

José Bobes-Bascarán, Ángel Fernández Leal, Eduardo Mosqueira-Rey, David Alonso-Ríos, Elena Hernández-Pereira, Vicente Moret-Bonillo

Centro de Investigación CITIC, University of Coruña, Spain
Correspondence: jose.bobes@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: The increasing adoption of artificial intelligent systems at sensitive domains where humans are particularly, such as medicine, has provided the context to deeply explore ways of making machine learning models (ML) understandable for their final users. The success of such systems require the trust of their users, and thus there is a need to design and provide methods to understand the decisions made by such systems. We start from a public Pancreatic Cancer dataset and experiment with different ML models on a diagnosis scenario with the goal to decide whether a patient should be prescribed with a chemotherapy treatment. To validate the diagnosis results we explore different explainability approaches: Decision Tree, Random Forest, and model agnostic ad-hoc models, and compare them against a standard Pancreatic Cancer treatment set of rules. The increasing adoption of artificial intelligent systems at sensitive domains where humans are particularly, such as medicine, has provided the context to deeply explore ways of making machine learning models (ML) understandable for their final users. The success of such systems require the trust of their users, and thus there is a need to design and provide methods to understand the decisions made by such systems. We start from a public Pancreatic Cancer dataset and experiment with different ML models. To validate the diagnostic results we explore different explainability approaches: Decision Tree based approach, Random Forest based approach, and different model agnostic ad-hoc approaches, and we compare them against a standard Pancreatic Cancer treatment set of rules.

1 Introduction

When creating Machine Learning (ML) models there is normally a trade of between interpretability and accuracy. While the former aims to create models that can be understand by their end-users, the search for accuracy often require complex models that are not easy to interpret.

In recent years ML models are being deployed covering a wide range of scenarios were it is of crucial importance the ability to understand how those models behave and reach a certain conclusion. It is equally important to reach a certain level of accuracy so that the models could be trust in real scenarios.

In this research, we continued from our previous work on AL Bobes-Bascarán et al. (2021) Bobes-Bascarán et al. (2023), where we first experimented with generated synthetic data and an Active Learning approach, and then followed with a real dataset introducing medical doctors in the loop of a therapy selection model for pancreatic cancer. On our previous work the goal was to overcame the scarcity of data available on the Pancreatic Cancer context by incorporating humans into the ML loop Mosqueira-Rey et al. (2022b). The focus is now on applying several explainability techniques to get useful insight about the underlining models.

We found two different concepts in the literature that refer to the quality of a system to be understood by its end-users: Explainability and Interpretability.

Explainability is the ability to describe how a model could reach a certain prediction or classification result so that it can be understand by its end-users.

Interpretability is intrinsic to the model itself and refers to the fact that end-users could interpret the relationship between the model inputs and its outputs.

On the one hand, we found that better interpretability and understandability leads to better trust and eases the adoption of AI systems. On the other hand, better accuracy require more complex models. This trade-of is a key factor when dealing with models deployed on domains such as healthcare.

To overcome the understandability issue eXplainable AI (XAI) Gunning (2017): "XAI will create a suite of machine learning techniques that enables human users to understand, appropriately trust, and effectively manage the emerging generation of artificially intelligent partners".

Some authors implicitly assumes that the ease of understanding and clarity targeted by XAI techniques for the model at hand, results in different application purposes, such as a better trustworthiness of the model's output by the audience. They do synthesize and enumerate definitions for these XAI goals, in example: Trustworthiness, Causality, Informativeness and Confidence Arrieta et al. (2020).

Nowadays, there can be found several methods that help to enhance ML models with easier to understand ad hoc exlainability models. Among others features, they provide several charts for both local and global explanations which is a very interesting characteristic for non-experts users. A set of tools helping in the different processes has been described by Mosqueira-Rey et al. (2022a).

2 Dataset

The dataset used in this work was obtained from The Cancer Genome Atlas Program Tomczak et al. (2015) published by the USA National Cancer Institute (NCI) and the National Human Genome Research Institute that contains a complete database of pancreatic cancer cases. It is composed of several research projects, among them, the TCGA-PAAD, currently with 185 diagnosed cases with all the necessary details.

There are 185 cancer positive patients (83 female and 102 male) and for each one we were interested in determining whether chemotherapy treatment was indicated or not based on the diagnostic information available. From the available cases 117 (60%)of them correspond to a chemotherapy treatment and 64 correspond to "Other".

It includes patient demographic information, family history, diagnosis, treatments, and genomic, epigenomic, transcriptomic and proteomic data. It issues information such as: the "stage event" which describes the pathological state of the tumor, the "clinical data" which describes the characteristics of the tumor and the occurrence of "new tumor events" which describes the patient follow-up.

3 Data Preparation

The aim of the experiment we have designed is to classify whether to prescribe a Chemotherapy treatment or not, based on some of the most relevant patient features available in the dataset.

A data curation process has been performed including (1) handling Missing Values/Null Values, (2) removing redundancy, (3) simplifying the target values with only *Chemotherapy* or *Other*, (4) converting some of the categorical variables into numerical, (5) using One-hot encoding for the rest of categorical features, (6) and dropping unneeded features.

The selected features are: pathologic_stage, pathologic_t, pathologic_m, pathologic_n, primary_diagnosis, tissue_or_organ_of_origin, age_at_index, and gender. We refer the reader to the

official NIH GDC documentation were a detailed description is provided.

The target variable is the therapy type to be prescribed to the patient. It can take several values, but we are interested only in whether or not it is chemotherapy. The possible values are 'Chemotherapy', or 'Other', as we have recoded all the alternative values (Hormone Therapy, Vaccine, and Ancillary) as 'Other'.

After preparing the database we performed a split between the training and the test data of 70 and 30% respectively.

4 Generating the models

With the goal of illustrating the balance between the accuracy and the explainability capabilities of the models, we have created several models and discuss their advantages and disadvantages.

4.1 Decision tree

A decision tree is a decision support hierarchical model that uses a tree-like model of decisions and their possible consequences.

We do create a decision tree with only 6 levels, and a minimum of 5 samples per leaf, using the training data.

	precision	recall	f1-score	support
Chemotherapy	0.70	0.88	0.78	34
Other	0.69	0.41	0.51	22
accuracy			0.70	56
macro avg	0.69	0.65	0.65	56
weighted avg	0.70	0.70	0.68	56

Text representation

```
|--- primary_diagnosis_Neuroendocrine carcinoma, NOS <= 0.50
|   |--- age_at_index <= 79.50
|   |   |--- age_at_index <= 62.50
|   |   |   |--- age_at_index <= 59.50
|   |   |   |   |--- age_at_index <= 56.50
|   |   |   |   |   |--- gender_female <= 0.50
|   |   |   |   |   |   |--- class: Chemotherapy
|   |   |   |   |   |--- gender_female >  0.50
|   |   |   |   |   |   |--- class: Other
...
```

Decision Tree Graphical representation (partial chart)

For the explainability aspect, we can count on the graph representation itself that clearly identifies which is the decision path of an instance. For each of the decision nodes, the condition is evaluated and we descent through the appropriate branch, until we get to a leaf level. For each of the nodes the *Gini index*, the number of samples and the specific class is represented.

Furthermore, we enhance the explanation using **Permutation Importance** that is an algorithm that computes importance scores for each of the feature variables of a dataset. These scores are determined by computing the sensitivity of a model to random permutations of feature values. After permuting a certain feature, t-he increase in the prediction error of the model determines the importance score of the feature (see 2).

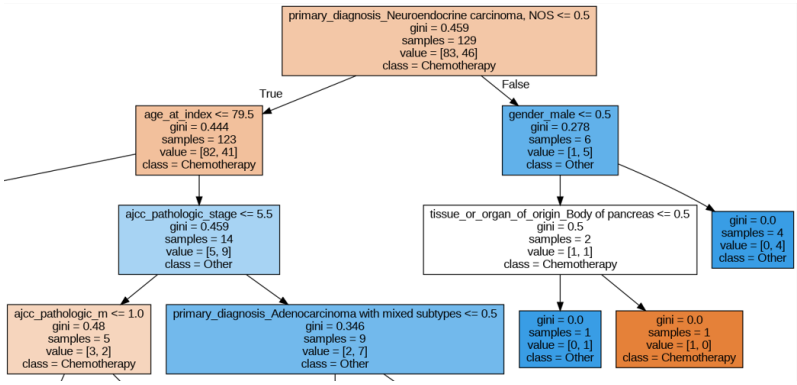


Figure 1: Decision Tree: graphical representation (partial chart).

Weight	Feature
0.4251	age_at_index
0.1077	primary_diagnosis_Neuroendocrine carcinoma, NOS
0.0990	ajcc_pathologic_stage
0.0922	tissue_or_organ_of_origin_Head of pancreas
0.0653	gender_female
0.0565	tissue_or_organ_of_origin_Pancreas, NOS
0.0512	primary_diagnosis_Adenocarcinoma with mixed subtypes
0.0402	ajcc_pathologic_m
0.0376	tissue_or_organ_of_origin_Body of pancreas
0.0251	gender_male
0	ajcc_pathologic_n
0	ajcc_pathologic_t
0	tissue_or_organ_of_origin_Tail of pancreas
0	primary_diagnosis_Carcinoma, undifferentiated, NOS
0	primary_diagnosis_Infiltrating duct carcinoma, NOS
0	primary_diagnosis_Mucinous adenocarcinoma
0	tissue_or_organ_of_origin_Overlapping lesion of pancreas
0	primary_diagnosis_Adenocarcinoma, NOS

Figure 2: Decision Tree: permutation importance.

4.2 Random forest

A Random Forest is an ensemble method which consist of a set of decision trees that combined, normally provide better accuracy than a single tree. Even if Random Forests are more accurate than Decision Trees, they are way more complex in terms of understandability.

	precision	recall	f1-score	support
Chemotherapy	0.64	1.00	0.78	34
Other	1.00	0.14	0.24	22
accuracy			0.66	56
macro avg	0.82	0.57	0.51	56
weighted avg	0.78	0.66	0.57	56

5 Understanding the models

Firstly, we would like to emphasize that the decision tree is already a white-box model and offers a good explainable capabilities by nature. Both the textual and the graphical representations suit many scenarios as it is easy to understand and interpret on single instances.

Nevertheless, using Permutation importance we can synthesize the rules produced by the tree data structure in an aggregated manner. Observing figure 2 we can see that *age at diagnosis*, the *primary diagnosis* being *Neuroendocrine carcinoma*, the *pathologic stage* of the cancer, and the fact of having the *tumor localized at the head of the pancreas* are the more relevant features of the model.

Secondly, for the Random Forest (RF) created using bootstrap aggregation by combining several decision trees, even if the first results deliver a lower accuracy than the DT, the ensemble models in general provide better generalization capabilities. As the RF is more complex than a simple DT we enhanced the RF by means of creating several ad-hoc explainability models on top or it. We have chosen widespread methods such SHAP and LIME.

SHAP is a method based on the cooperative game theory proposed by Shapley Shapley (1953).

It handles explainability by attributing a numerical value to each of the features of the model. This number represents the contribution of the feature to the model prediction or classification result.

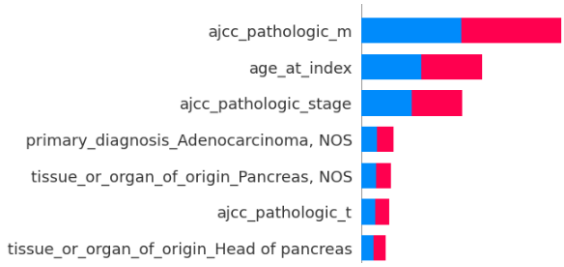


Figure 3: SHAP summary plot for variable importance.

The summary plot 3 give us the variable importance. Features with high predicting power are shown at the top and the ones with low predicting powers are shown at the bottom.

We could also see the contribution of each feature into the prediction probability. The redder the color, the higher the value and vice versa. Also, when the value is on the positive side, it contributes to the class 'Chemotherapy' prediction result probability and vice versa.

LIME, proposed by Ribeiro et. al Ribeiro et al. (2016), is the acronym for Locally Interpretable

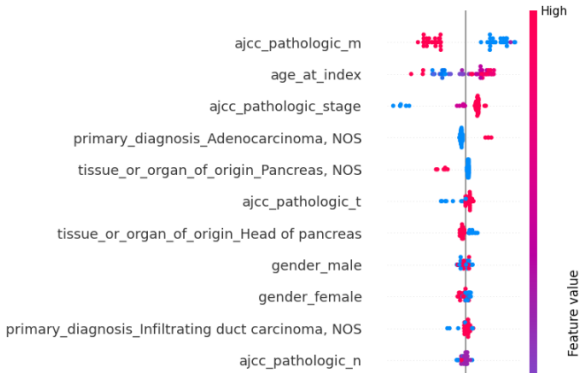


Figure 4: SHAP summary plot for impact on output.

Model-agnostic Explanations. It tries to understand the features that influence a prediction on a single instance focusing on a local level where a linear model is enough to explain the behavior.

We present local explanations for two of the patients available on the dataset.

The first patient (index 3) is a woman who was diagnosed an infiltrating duct carcinoma at the tail of the pancreas when she was 58 years old. The TNM is T3, N0, M1 and the pathologic stage is IV. The model selects a Chemotherapy treatment 6.

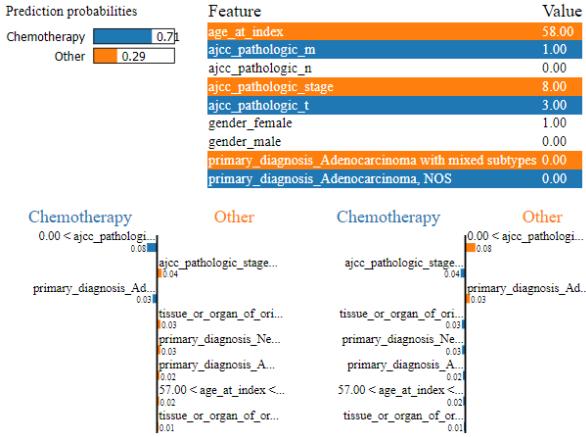


Figure 5: LIME patient 3 resume.

The second patient (index 4) is a woman who was diagnosed and infiltrating duct carcinoma in a none specific area of the pancreas when she was 66 years old. The TNM is T3, N0, M0 and the pathologic stage is IIA. The model selects a Chemotherapy treatment. The model selects a Chemotherapy treatment 5.

6 Conclusions

We have described the importance of interpreting and understanding a ML models, especially in a sensible context, as it is the case of the pancreatic cancer domain.

Through the construction of two different ML models we have illustrated how a simple model could be easily understood with a visual representation, and how difficult it will be to compre-

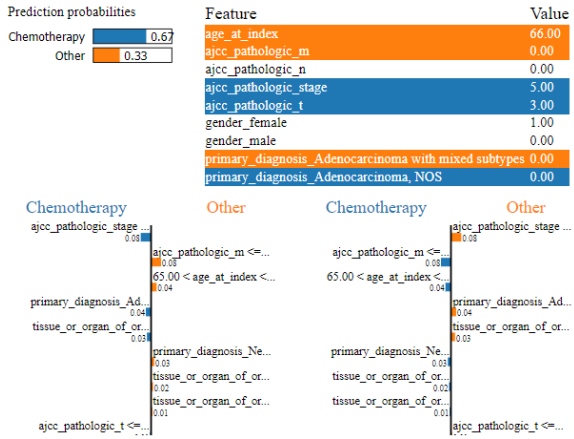


Figure 6: LIME patient 4 resume.

hend a more sophisticated one, as the ensemble random forest.

Several ad-hoc models have been created using the SHAP and LIME methods to provide further explainability features over the original models. Those models provide an easy to understand variable importance for both global and local explanations, that can be combined to fully interpret and rely on a ML model.

We plan to compare the results obtained, to a set of guidelines widely adopted among pancreatic cancer professionals. The idea will be to determine if the explainable model is able to reach the same expert criteria.

In our study, the pathologic M, and the age at the first diagnosis are the most relevant characteristics, but perhaps the medical literature do not agree on those attributes.

7 Acknowledgments

This work has been supported by the State Research Agency of the Spanish Government (grant PID2019-107194GB-I00/AEI/10.13039/501100011033) and by the Xunta de Galicia (grant ED431C2022/44), supported in turn by the EU European Regional Development Fund. We wish to acknowledge support received from the Centro de Investigación de Galicia CITIC, funded by the Xunta de Galicia and the European Regional Development Fund (Galicia 2014-2020 Program; grant ED431G 2019/01).

Bibliography

A. B. Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. García, S. Gil-López, D. Molina, R. Benjamins, et al. Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai. *Information fusion*, 58:82–115, 2020.

J. Bobes-Bascarán, E. Mosqueira-Rey, and D. Alonso-Ríos. Improving medical data annotation including humans in the machine learning loop. *Engineering Proceedings*, 7(1):39, 2021.

J. Bobes-Bascarán, A. Pérez-Sánchez, E. Mosqueira-Rey, D. Alonso-Rıos, and E. Hernández-Pereira. Using active learning to improve the treatment selection on pancreatic cancer patients. In *Proceedings of V XoveTIC Conference. XoveTIC*, volume 14, pages 70–72, 2023.

D. Gunning. Explainable artificial intelligence (xai). *Defense advanced research projects agency (DARPA), nd Web*, 2(2):1, 2017.

E. Mosqueira-Rey, E. Hernández-Pereira, D. Alonso-Ríos, and J. Bobes-Bascarán. A classification and review of tools for developing and interacting with machine learning systems. In *Proceedings of the 37th Annual ACM Symposium on Applied Computing*, pages 1083–1092. Association for Computing Machinery, New York, NY, USA, 2022a. ISBN 9781450387132.

E. Mosqueira-Rey, E. Hernández-Pereira, D. Alonso-Ríos, J. Bobes-Bascarán, and A. Fernández-Leal. Human-in-the-loop machine learning: A state of the art. *Artificial Intelligence Review*, 2022b.

M. T. Ribeiro, S. Singh, and C. Guestrin. "why should i trust you?": Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, page 1135–1144, New York, NY, USA, 2016. Association for Computing Machinery. ISBN 9781450342322.

L. S. Shapley. A value for n-person games. In H. W. Kuhn and A. W. Tucker, editors, *Contributions to the Theory of Games II*, pages 307–317. Princeton University Press, Princeton, 1953.

K. Tomczak, P. Czerwińska, and M. Wiznerowicz. The cancer genome atlas (TCGA): an immeasurable source of knowledge. *Contemporary Oncology*, 19(1A):68–77, 2015.

Estimation of Distance Correlation: a Simulation-based Comparative Study

Blanca E. Monroy-Castillo, M.A. Jácome, and Ricardo Cao

Centro de Investigación CITIC, Universidade da Coruña, 15071 A Coruña, Spain
Research Group MODES, Faculty of Sciences, Universidade da Coruña, 15071 A Coruña, Spain
Research Group MODES, Faculty of Computer Science, Universidade da Coruña, 15071 A Coruña, Spain
Correspondence: b.mcastillo@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: The notion of distance correlation was introduced to measure the dependence between two random vectors, not necessarily of equal dimensions, in a multivariate setting. In their work, Székely et al. (2007) proposed an estimator for the squared distance covariance, and they also proved that this estimator is a V-statistic. On the other hand, Székely and Rizzo (2014) introduced an unbiased version of the squared sample distance covariance, which was subsequently identified as a U-statistic in Huo and Székely (2016). In this study, a simulation is conducted to compare both distance correlation estimators: the U-estimator and the V-estimator. The analysis assesses their efficiency (mean squared error) and contrasts the computational times of both approaches across various dependence structures.

1 Introduction

Distance correlation is a novel measure of dependence between random vectors. The concept of distance correlation was introduced by Székely et al. (2007). They emphasize that distance covariance and distance correlation draw a parallel to product-moment covariance and correlation. Nevertheless, unlike the classical definition of correlation, distance correlation is zero only when the random vectors are independent. Essentially, for all distributions with finite first moments, distance correlation ($\mathcal{R}$) extends the concept of correlation in two fundamental ways:

(i) $\mathcal{R}(X, Y)$ is defined for X and Y in arbitrary dimensions;

(ii) $\mathcal{R}(X, Y) = 0$ characterizes independence of X and Y.

Distance correlation satisfies $0 \leqslant \mathcal{R} \leqslant 1$, and $\mathcal{R} = 0$ if and only if X and Y are independent.

Székely et al. (2007) introduced a sample distance covariance estimator, demonstrating that this estimator functions as a V-statistic. Furthermore, in the work of Székely and Rizzo (2014), intermediate findings are outlined, culminating in an unbiased estimator for squared distance covariance. The unbiased estimator is established as a U-statistic in Huo and Székely (2016), along with the introduction of a novel algorithm. This algorithm shows a computational complexity of $\mathcal{O}(n \log n)$, a noteworthy enhancement compared to the $\mathcal{O}\left(n^2\right)$ complexity associated with the direct implementation of the V-estimator put forward by Székely et al. (2007).

It is important to highlight that, to the best of our knowledge, the merits and drawbacks linked to each distance correlation estimator (U-estimator and V-estimator) have not been thoroughly investigated in the current body of literature. In this study, both estimators are comprehensively examined and compared.

2 Preliminaries

Consider two random vectors, $X \in \mathbb{R}^p$ and $Y \in \mathbb{R}^q$, where both p and q are positive integers. Let ϕ_X and ϕ_Y represent the characteristic functions of X and Y respectively, and denote the joint characteristic function of X and Y as $\phi_{X,Y}$. In this context, the distance covariance between these random vectors X and Y, assuming they possess finite first moments, is defined as a nonnegative scalar denoted as $\mathcal{V}(X,Y)$, given by the square root of:

$$\begin{aligned}\mathcal{V}^2(X,Y) &= ||\phi_{X,Y}(t,s) - \phi_X(t)\phi_Y(s)||^2 \\ &= \frac{1}{c_p c_q}\int_{\mathbb{R}^{p+q}} \frac{||\phi_{X,Y}(t,s) - \phi_X(t)\phi_Y(s)||^2}{|t|_p^{1+p}|s|_q^{1+q}}dt\,ds,\end{aligned}$$

where $c_d = \frac{\pi^{(1+d)/2}}{\Gamma((1+d)/2)}$. Similarly, distance variance ($\mathcal{V}(X)$) is defined as the square root of

$$\mathcal{V}^2(\mathrm{X}) = \mathcal{V}^2(\mathrm{X},\mathrm{X}) = ||\phi_{X,X}(t,s) - \phi_X(t)\phi_X(s)||^2.$$

And the distance correlation ($\mathcal{R}$) between random vectors X and Y, assuming they possess finite first moments, is the positive square root of a nonnegative quantity denoted as $\mathcal{R}^2(X,Y)$. This quantity is defined as follows:

$$\mathcal{R}^2(X,Y) = \begin{cases} \frac{\mathcal{V}^2(X,Y)}{\sqrt{\mathcal{V}^2(X)\mathcal{V}^2(Y)}}, & \mathcal{V}^2(X)\mathcal{V}^2(Y) > 0 \\ 0, & \mathcal{V}^2(X)\mathcal{V}^2(Y) = 0. \end{cases} \tag{29.1}$$

Alternatively, Székely et al. (2007) proposed an equivalent method for computing distance covariance through expectations. This is, if $E|X|_p^2 < \infty$ and $E|Y|_q^2 < \infty$, then $E[|X|_p|Y|_q] < \infty$, and

$$\begin{aligned}\mathcal{V}^2(X,Y) = {} & E[|X_1 - X_2|_p|Y_1 - Y_2|_q] + E[|X_1 - X_2|_p]E[|Y_1 - Y_2|_q] \\ & -2E[|X_1 - X_2|_p|Y_1 - Y_3|_q],\end{aligned} \tag{29.2}$$

where (X_1, Y_1), (X_2, Y_2) and (X_3, Y_3) are independent and identically distributed as (X, Y).

When dealing with an observed random sample $(\mathbf{X}, \mathbf{Y}) = \{(X_k, Y_k) : k = 1, \ldots, n\}$ drawn from the joint distribution of random vectors $X \in \mathbb{R}^p$ and $Y \in \mathbb{R}^q$, Székely et al. (2007) introduced the empirical distance covariance ($\mathcal{V}_n(\mathbf{X}, \mathbf{Y})$) as follows. The empirical distance covariance $\mathcal{V}_n(\mathbf{X}, \mathbf{Y})$ is a nonnegative value defined by the square root of:

$$\mathcal{V}_n^2(\mathbf{X}, \mathbf{Y}) = \frac{1}{n^2}\sum_{k,l=1}^{n} A_{kl}B_{kl}, \tag{29.3}$$

where A_{kl} and B_{kl} denote the corresponding double-centered distance matrices defined as:

$$A_{kl} = \begin{cases} a_{kl} - \frac{1}{n}\sum_{j=1}^{n} a_{kj} - \frac{1}{n}\sum_{i=1}^{n} a_{il} + \frac{1}{n^2}\sum_{i,j=1}^{n} a_{ij}, & k \neq l \\ 0, & k = l, \end{cases}$$

where $a_{kl} = |X_k - X_l|$ the pairwise distances of the X observations, similarly for $b_{kl} = |Y_k - Y_l|$. In the same way

$$\mathcal{V}_n^2(\mathbf{X}) = \mathcal{V}_n^2(\mathbf{X}, \mathbf{X}) = \frac{1}{n^2}\sum_{k,l=1}^{n} A_{kl}^2. \tag{29.4}$$

Theorem 1 in Székely et al. (2007) establishes the nonnegativity of $\mathcal{V}_n^2(\mathbf{X}, \mathbf{Y})$. Furthermore, it demonstrates that under independence, $\mathcal{V}_n^2$ behaves as a degenerate kernel V-statistic. The

computational complexity of this estimator is $\mathcal{O}(n^2)$. Now, the empirical distance correlation $\mathcal{R}_n(\mathbf{X}, \mathbf{Y})$ is defined as the square root of:

$$\mathrm{dCorV}^2(\mathbf{X}, \mathbf{Y}) \quad = \quad \mathcal{R}_n^2(\mathbf{X}, \mathbf{Y}) = \begin{cases} \frac{\mathcal{V}_n^2(\mathbf{X},\mathbf{Y})}{\sqrt{\mathcal{V}_n^2(\mathbf{X})\mathcal{V}_n^2(\mathbf{Y})}}, & \mathcal{V}_n^2(\mathbf{X})\mathcal{V}_n^2(\mathbf{Y}) > 0 \\ 0, & \mathcal{V}_n^2(\mathbf{X})\mathcal{V}_n^2(\mathbf{Y}) = 0, \end{cases} \tag{29.5}$$

which is always non-negative.

Similarly, in Székely and Rizzo (2014), the $\mathcal{U}$-centered matrix is introduced as follows. Consider a symmetric, real-valued $n \times n$ matrix $A = (a_{kl})$ with a zero diagonal, where $n > 2$. The entry in the (k, l)th position of the $\mathcal{U}$-centered matrix $\tilde{A}$ is defined as:

$$\tilde{A}_{kl} = \begin{cases} a_{kl} - \frac{1}{n-2}\sum_{j=1}^{n} a_{kj} - \frac{1}{n-2}\sum_{i=1}^{n} a_{il} + \frac{1}{(n-1)(n-2)}\sum_{i,j=1}^{n} a_{ij}, & k \neq l; \\ 0, & k = l. \end{cases}$$

Here "$\mathcal{U}-$centered" is so named because the inner product,

$$\mathcal{U}_n^2(\mathbf{X}, \mathbf{Y}) \quad = \quad \left(\tilde{A} \cdot \tilde{B}\right) = \frac{1}{n-3}\sum_{i \neq j} \tilde{A}_{kl}\tilde{B}_{kl}, \tag{29.6}$$

defines an unbiased estimator of the squared distance covariance. The work by Huo and Székely (2016) established that the estimator in Equation (29.6) is a U-statistic. This reevaluation paved the way for the creation of an efficient algorithm, which can be executed with a computational complexity of $\mathcal{O}(n \log n)$.
Thus, it is possible to define the empirical distance correlation through U-statistics (dCorU) which is the square root of

$$\mathrm{dCorU}^2(\mathbf{X}, \mathbf{Y}) = \begin{cases} \frac{\mathcal{U}_n^2(\mathbf{X},\mathbf{Y})}{\sqrt{\mathcal{U}_n^2(\mathbf{X})\mathcal{U}_n^2(\mathbf{Y})}}, & \mathcal{U}_n^2(\mathbf{X})\mathcal{U}_n^2(\mathbf{Y}) > 0 \\ 0, & \mathcal{U}_n^2(\mathbf{X})\mathcal{U}_n^2(\mathbf{Y}) = 0, \end{cases} \tag{29.7}$$

where $\mathcal{U}_n^2(\mathbf{X})$ represents the distance variance of $\mathbf{X}$, similarly $\mathcal{U}_n^2(\mathbf{Y})$ for $\mathbf{Y}$.

These results have spurred the development and advancement of numerous software packages, accessible for use in both the R software environment (R Core Team, 2022) and Python (Van Rossum and Drake Jr, 1995). In the realm of Python, libraries such as statsmodels (Seabold and Perktold, 2010), hyppo (Panda et al., 2021), dcor (Ramos-Carreño, 2022), and pingouin (Vallat, 2018) are available. In the R environment, notable packages include energy (Rizzo and Székely, 2022), dcortools (Edelmann and Fiedler, 2022), and the Rfast package (Papadakis et al., 2022).

3 Simulation study

A Monte Carlo simulation study was conducted in order to compare the efficiency of the dCorU and dCorV estimators across diverse dependence structures. For the simulation study, the dcortools package was utilized, specifically employing the distcor function. To calculate the distance correlation through dCorU, the code used is distcor(X, Y, bias.corr = TRUE). While, to compute dCorV, the code is distcor(X, Y, bias.corr = FALSE) or simply distcor(X, Y).

To facilitate a comprehensive comparison of each estimator's efficiency (MSE) and computational time, two different models are utilized. The first one, the bivariate normal model, which

is used because Székely et al. (2007) proved that the distance correlation, in terms of Pearson's correlation coefficient, is given by the square root of:

$$\mathcal{R}^2(X,Y) = \frac{\rho \arcsin \rho + \sqrt{1-\rho^2} - \rho \arcsin \rho/2 - \sqrt{4-\rho^2} + 1}{1 + \pi/3 - \sqrt{3}}.$$

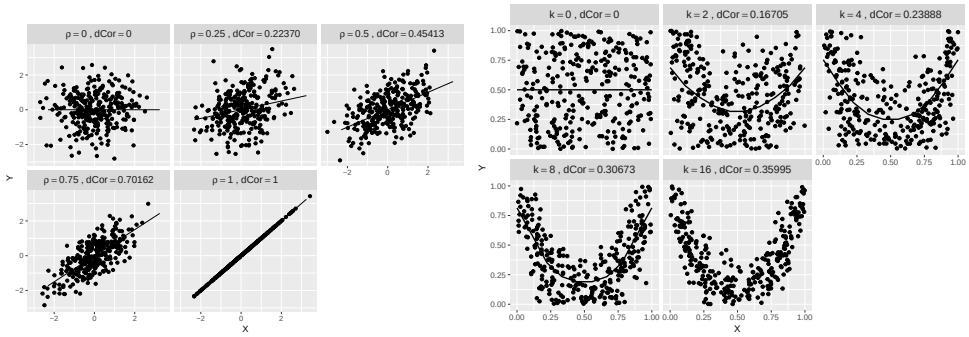
The second model corresponds to a nonlinear model defined as:

$$f_{X,Y}(x,y) = c\left[1 - \left(y - 4\left(x - \frac{1}{2}\right)^2\right)^2\right]^k I_{[0,1]}(x) I_{[0,1]}(y),$$

where $k \in \mathbb{N}$ and c is a constant that depends on the value of k.

These models encompass varying levels of dependence and are evaluated across two sample sizes, 100 and 10000. Each simulation is performed with 1000 Monte Carlo repetitions. The precise computation of the distance covariance for the second model is achieved using Equation (29.2). Similarly, the calculations for $\mathcal{V}(X)$ and $\mathcal{V}(Y)$ were performed. Finally, the value of $\mathcal{R}$ is obtained from Equation (29.1).

Five samples drawn from different values of the parameters of each model are depicted in Figure 1, accompanied by the respective parameter value and the corresponding distance correlation. The lines represent the conditional mean $E[Y|X = x]$ in each case.



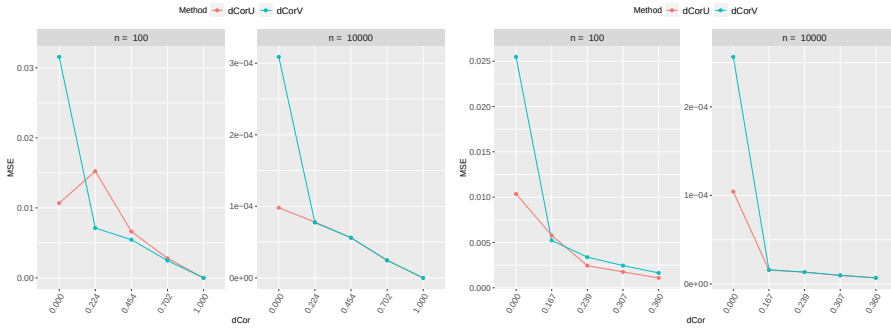
(a) Bivariate normal model samples for different values of ρ.

(b) Nonlinear model samples for different values of k.

Figure 1: Samples ($n = 300$) for different values of ρ, k and their corresponding distance correlation for the respective model.

In the bivariate normal model, complete dependence between X and Y results in $\mathcal{R} = 1$. However, the maximum value of the distance correlation for the nonlinear model, reached when X and Y are totally dependent (i.e. $k \to \infty$), is $\mathcal{R} = 0.41$.

The MSE results for each model are illustrated in Figure 2. Several observations can be made: under independence, dCorU outperforms dCorV in the two cases for both sample sizes. However, when it comes to cases of dependence, whether weak or strong, dCorV shows better performance than dCorU. These disparities are more noticeable with weak dependence and smaller sample sizes. Interestingly, this pattern deviates for the nonlinear model (depicted in Figure 2b), where, even with weak dependence, dCorU surpasses dCorV, although the differences are not substantial. Finally, for a larger sample size ($n = 10000$), both estimators exhibit remarkable similarity across both models.



(a) MSE under the bivariate normal model from $\mathcal{R} = 0$ ($\rho = 0$) to $\mathcal{R} = 1$ ($\rho = 1$).

(b) MSE under the nonlinear model from $\mathcal{R} = 0$ ($k = 0$) to $\mathcal{R} = 0.36$ ($k = 16$).

Figure 2: MSE of dCorU and dCorV under the two models for different distance correlation values.

Finally, the computational time for each estimator are in Table 1. The characteristics of the computer equipment used are the following ones: CPU 12th Gen Intel(R) Core(TM) i7-1280P 2.00 GHz and RAM 16 GB.

Note that there are no substantial differences in the computation times among each of the estimators. However, it is worth noting that these times were acquired using the dcortools package. If a different package, such as energy package, were employed, the timings would likely vary, potentially resulting in an increase.

Table 1: Computational time in secs for 1000 samples.

	$n = 100$		$n = 1000$		$n = 10000$	
	dCorU	dCorV	dCorU	dCorV	dCorU	dCorV
Time	0.36	0.33	0.63	0.58	2.22	2.51

4 Conclusions

This study focused on examining the performance of the dCorU and dCorV estimators for distance correlation through Monte Carlo simulations. The results presented here underscore the significance of the specific scenario. In cases of independence, dCorU demonstrates superiority over dCorV across all considered scenarios. However, when dependence is present, the outcomes diverge. The dCorV estimator aligns with superior results in terms of Mean Squared Error (MSE) for linear model (bivariate normal), as well as for the nonlinear model under weak dependence. Moreover, in the realm of computational efficiency, both estimators, dCorU and dCorV, stand on competitive ground.

Acknowledgements

CITIC is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS).

Bibliography

D. Edelmann and J. Fiedler. *dcortools: Providing Fast and Flexible Functions for Distance Correlation Analysis*, 2022. URL *https://CRAN.R-project.org/package=dcortools*. R package version 0.1.6.

X. Huo and G. J. Székely. Fast Computing for Distance Covariance. *Technometrics*, 58(4):435–447, 2016.

S. Panda, S. Palaniappan, J. Xiong, E. Bridgeford, R. Mehta, and C. Shen. *hyppo: A multivariate hypothesis testing Python package*, 2021. URL *https://github.com/neurodata/hyppo.*

M. Papadakis, M. Tsagris, M. Dimitriadis, S. Fafalios, I. Tsamardinos, M. Fasiolo, G. Borboudakis, J. Burkardt, C. Zou, K. Lakiotaki, and C. Chatzipantsiou. *Rfast: A Collection of Efficient and Extremely Fast R Functions*, 2022. URL *https://CRAN.R-project.org/package=Rfast*. R package version 2.0.6.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2022. URL *https://www.R-project.org/*.

C. Ramos-Carreño. *dcor: distance correlation and energy statistics in Python*, 2022. URL *https://pypi.org/project/dcor/*.

M. L. Rizzo and G. J. Székely. *energy: E-Statistics: Multivariate Inference via the Energy of Data*, 2022. URL *https://CRAN.R-project.org/package=energy*. R package version 1.7-11.

S. Seabold and J. Perktold. *Statsmodels: Econometric and statistical modeling with Python*, 2010. URL *https://github.com/statsmodels/statsmodels/*.

G. J. Székely and M. L. Rizzo. Partial Distance Correlation with Methods for Dissimilarities. *Annals of Statistics*, 42(6):2382–2412, 2014.

G. J. Székely, M. L. Rizzo, and N. K. Bakirov. Measuring and testing dependence by correlation of distances. *Annals of Statistics*, 35(6):2769–2794, 2007.

R. Vallat. Pingouin: statistics in Python. *Journal of Open Source Software*, 3(31):1026, 2018.

G. Van Rossum and F. L. Drake Jr. *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam, 1995.

Towards a FAIR Dataset for Spanish Non-Functional Requirements

María Isabel Limaylla-Lunarejo, Nelly Condori-Fernandez, and Miguel R. Luaces

Database Lab, Fac. Informática, Universidade da Coruña, 15071 A Coruña, Spain
Citius, Universidad de Santiago de Compostela, 15071 A Coruña, Spain
Centro de Investigación CITIC, Universidade da Coruña, 15071 A Coruña, Spain
Correspondence:
maria.limaylla@@udc.es,n.condori.fernandez@usc.es,miguel.luaces@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: Supervised Machine Learning algorithms (ML) have enhanced the performance of the automatic non-functional requirements (NFR) classification in the Requirements Engineering domain. However, the lack of public datasets, dealing with imbalanced datasets and reproducibility are current concerns in ML experiments. We conducted a quasi-experiment to generate a dataset of NFR in the Spanish Language, following the FAIR Principles. We collected 109 requirements from an open access repository of the University of A Coruña, and performed a labeling process based in the categories and subcategories of the ISO/IEC 25010 quality model. Using a Fleiss' Kappa test we obtained a substantial agreement (0.78) at the category level and a moderate agreement (0.48) when the classification is per subcategory.

1 Introduction

Supervised Machine Learning (ML) algorithms and Natural Language Processing techniques have been used to improve the performance of the automatic non-functional requirements classification. However, the lack of publicly datasets for requirements categorized in sub-classes of non-functional classification is still one concern when conducting ML experimentsAhmad et al. (2020); Binkhonain and Zhao (2019). Moreover, the lack of diverse datasets in languages other than English is also currently challenges. Reproducibility is another concern when speaking of ML experiments, in part due to several barriers like accessibility and availability, capability of reuse, among others. To avoid these barriers, Wilkinson el al.(2016) propose the FAIR Guiding Principles, that consist of four principles: Findable, Accessible, Interoperable, and Reusable, guiding how data should be managed to be more easily accessible, understood, exchangeable, and reusable Wilkinson et al. (2016).

In this research we present a summary of the quasi-experiment conducted in a previous work Limaylla-Lunarejo et al. (2023). The main objective of the quasi-experiment was generate a dataset of non-functional requirements written in the Spanish Language, following the FAIR Guiding Principles for facilitating reuse. The Fleiss' Kappa test was used to assess the inter-rater reliability with multiple annotators.

2 Adapted FAIRification process

Some research have explored the application of the FAIR principles to datasets. One example is the FAIRification process propose by GO FAIR[1], an initiative to coordinate and collaborate on the global Internet of FAIR Data & Services (IFDS). Figure 1 presents the FAIRification process adopted by GO FAIR, consisting on seven steps: 1.Retrieve non-FAIR data, 2.Analyse the retrieved data, 3.Define the semantic model, 4.Make data linkable, 5.Assign license, 6.Define metadata for the dataset, and 7.Deploy FAIR data resource. This process have been used as a base on several studies Kochev et al. (2020); Sinaci et al. (2020).

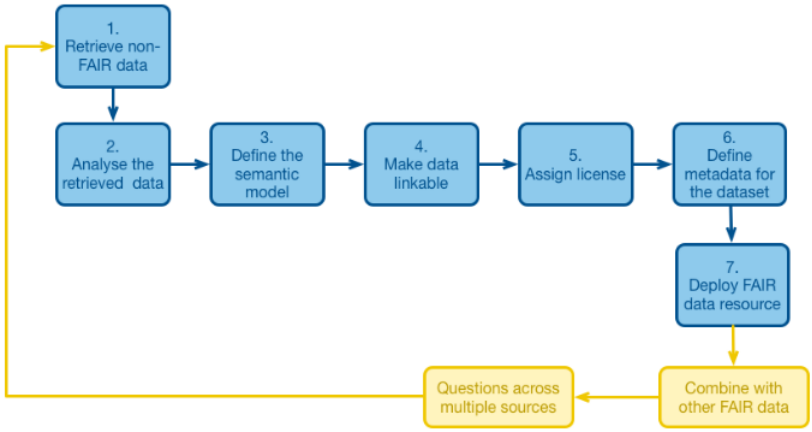


Figure 1: FAIRification Process GO FAIR International Support and Coordination Office. (2022)

While the original process primarily focused on making FAIR existing data (and metadata), our adaptation allowed the generation of FAIR datasets. In order to utilize the GO FAIR Process (FAIRification) for dataset generation, we made adaptations by incorporating, modifying, and excluding certain tasks. Figure 2 shows the adapted FAIRification process, that consist also in seven tasks: 1.Define Semantic Model, 2.Data Definition, 3.Data Collection, 4.Data Labeling, 5.Define Metadata, 6.FAIR Validation, and 7.Data Publishing.

The first two steps perform several definitions before the data collection and labeling. A semantic model involves defining the meaning of entities, their relationships, vocabulary, and ontologies/taxonomies. The Certified Professional for Requirements Engineering (CPRE) glossary Glinz (2011) was used for concepts like 'requirement', 'non-functional requirement', and 'stakeholders'. The ISO/IEC 25010 quality model ISO (2011) was chosen as taxonomy for the NFRs, and an entity-relationship model was selected to represent data with two entities: 'requirements' and 'projects'. We also established the data's structure for each entity, aligning it with the previously established semantic model. The next two tasks are focused on data collection and labeling. We review several Bachelor projects and selected 19 projects with has at least three NFR and collected the information in two Excel files. The labeling process will be present in more detail in the following section. The data collection and labeling could be an iterative process. Once the data is collected and labeled, the metadata is describe, such as the authors, the description, the license, the language used, the process, etc. Finally, the last two tasks is about carrying out an evaluation of the FAIR principles before, during and/or after publishing, the last task. A preliminary version of Metadata and datasets was released on Zenodo CERN and OpenAIRE (2013), an open publishing repository aligned with the FAIR principles. Four

[1] https://www.go-fair.org/fair-principles/fairification-process/

files were published[2], including the data structure, projects list and the requirements with the final label for the categories and subcategories, each one in a CSV file; and an RDF ontology model.

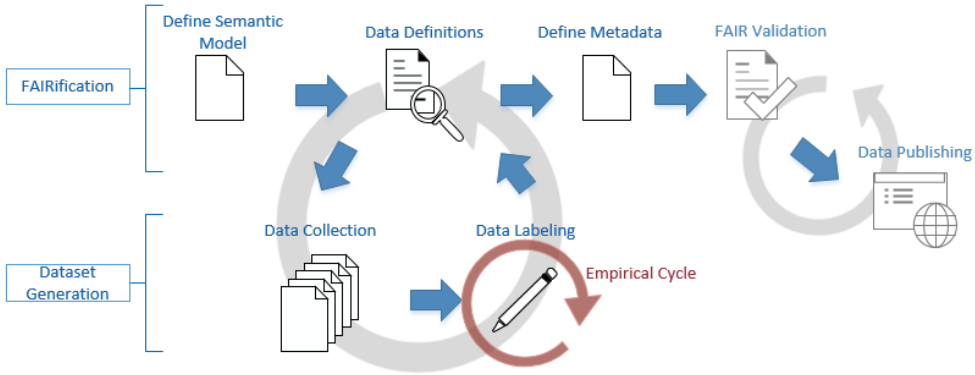


Figure 2: Adapted FAIRification Process for Dataset Generation

3 Experiment and Results

This section introduces the research questions, the instrumentation and process of the quasi-experiment, and the data analysis and results.

3.1 Research Questions

- **RQ1**: Does the adapted FAIRification process contribute to improve the reliability of agreement between multiple annotators?
- **RQ2**: Which NFR categories/subcategories are less reliable due to the high number of disagreements in several requirements?

3.2 Instrumentation and process

We collected 109 requirements from the free repository of the University of A Coruña[3], and performed a labeling process based in the categories and subcategories of the ISO/IEC 25010 quality model[4]. The labeling process was accomplished by seven annotators (four PhD students and three professionals), all native Spanish speakers. The requirements were divided in four groups and assigned one or two groups to each annotators. We performed a meeting with the annotators to explain the process and also prepared a document with the Spanish definitions of software requirements: functional and non-functional, and the definitions of all categories (characteristics) and subcategories (sub-characteristics) of the ISO/IEC 25010. All this definitions were based on the semantic model, the vocabulary, and the taxonomy defined in the first two task of the adapted FAIRification process.

[2] https://doi.org/10.5281/zenodo.7124407
[3] https://ruc.udc.es
[4] https://iso25000.com/index.php/en/iso-25000-standards/iso-25010

3.3 Data analysis and Results

Each requirement obtained between three and five labels from the annotators. For define the final label we based on unanimity or at least a prevailing vote among all the labels (majority). The outcomes of the label consensus for categories are displayed in Table 1. For the 109 labeled requirements a 63% obtained a label by unanimity, a 35% by majority and a 2% didn't agreed. From this two requirements, one was identified as Constraint (not a quality concern) and the other mentions a compact device and it could lead to confusion.

Table 1: Numbers of requirements belonging to categories

Categories	Unanimity	Majority	No agreement	Total
Usability	26	6	-	32
Security	19	6	-	25
Performance	11	8	-	19
Reliability	10	6	-	16
Functional Suitability	1	4	-	5
Maintainability	2	3	-	5
Portability	0	5	-	5
No agreement	0	0	2	2
Total	69	38	2	109

The identical analysis was carried out for the 107 requirements regarding the subcategories label. Around 37% of the requirements were labeled with unanimity in the subcategory level and 42% were labeled with majority agreement. The annotators disagreed on 21% of the classification of the subcategory of requirements. Table 2 presents the numbers of requirements belonging to these three subcategories (Unanimity, Majority, No agreement). Annotators encountered challenges and achieved limited consensus when categorizing requirements under subcategories associated with Usability (Operability, Learnability, User interface aesthetics, Accessibility, User error protection, and Appropriateness recognizability), Security (Confidentiality, Authenticity, and Integrity), and Performance efficiency (Time behavior, Capacity, and Resource utilization). Recurring disagreements were observed for subcategory groups like Operability and Learnability, Time behavior and Capacity, and Confidentiality, Authenticity, and Integrity.

Using a Fleiss' Kappa test Fleiss (1971) we obtained a substantial agreement in the category level (0.78) and a moderate agreement (0.48) when the classification is per subcategory. This indicate that it was possible to label almost all requirements based on unanimity or majority agreement. Regarding subcategories, there was a 20% of requirements without consensus, probably due to numerous subcategories Gut and Bayerl (2004), intersection of some meaning (such as Confidentiality and Integrity subcategories Samonas and Coss (2014)) and a lack of detail in some requirements.

4 Conclusions

In summary, this experiment has made two main contributions. Firstly, it has addressed the prevalent lack of requirement datasets in the Spanish language by successfully generating a new Spanish dataset. Secondly, it has provided valuable insights into the FAIRification process, offering an adapted framework that incorporates dataset creation from its inception. The reliability assessment of agreement among multiple annotators, using Fleiss' Kappa, has demonstrated substantial agreement when the classification is conducted at the category level (0.78) and moderate agreement (0.48) when the classification is done for subcategories. Future work will include additional experiments in NFR labeling and a FAIR validation process.

Table 2: Numbers of requirements belonging to subcategories

Categories	SubCategories	U	M	N	Total
Usability	Accessibility	0	1	-	1
	User error protection	1	1	-	2
	Learnability	1	2	-	3
	Operability	5	6	-	11
	User interface aesthetics	4	4	-	8
	No Agreement	0	0	7	**7**
Portability	Adaptability	0	5	-	5
Security	Authenticity	1	4	-	5
	Confidentiality	1	6	-	7
	Integrity	5	3	-	8
	No Agreement	0	0	5	**5**
Performance efficiency	Capacity	0	1	-	1
	Time behaviour	9	0	-	9
	Resource utilization	1	2	-	3
	No Agreement	0	0	6	**6**
Reliability	Recoverability	2	0	-	2
	Fault tolerance	2	2	-	4
	Availability	6	1	-	7
	Maturity	0	1	-	1
	No Agreement	0	0	2	**2**
Maintainability	Modifiability	1	3	-	4
	Modularity	1	0	-	1
Functional Suitability	Functional completeness	0	1	-	1
	Functional correctness	0	1	-	1
	Functional appropriateness	0	1	-	1
	No Agreement	0	0	2	**2**
	Total	40	45	22	107

Legend: U: Unanimity, M: Majority, N: No agreement

Acknowledgements

CITIC is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS).

Bibliography

A. Ahmad, C. Feng, M. Khan, A. Khan, A. Ullah, S. Nazir, and A. Tahir. A systematic literature review on using machine learning algorithms for software requirements identification on stack overflow. *Security and Communication Networks*, 2020, 2020.

M. Binkhonain and L. Zhao. A review of machine learning algorithms for identification and classification of non-functional requirements. *Expert Systems with Applications: X*, 1:100001, 2019.

CERN and OpenAIRE. Zenodo, 2013. URL *https://www.zenodo.org/*.

J. L. Fleiss. Measuring nominal scale agreement among many raters. *Psychological bulletin*, 76 (5):378, 1971.

M. Glinz. A glossary of requirements engineering terminology. *Standard Glossary of the Certified Professional for Requirements Engineering (CPRE) Studies and Exam, Version*, 1:56, 2011.

GO FAIR International Support and Coordination Office. Fairification process, 2022. URL *https://www.go-fair.org/fair-principles/*. [Online; accessed 02-September-2023].

U. Gut and P. S. Bayerl. Measuring the reliability of manual annotations of speech corpora. In *Speech prosody 2004, international conference*, 2004.

ISO. Iso/iec 25010:2011 systems and software engineering - systems and software quality requirements and evaluation (square) - system and software quality models, 2011.

N. Kochev, N. Jeliazkova, V. Paskaleva, G. Tancheva, L. Iliev, P. Ritchie, and V. Jeliazkov. Your spreadsheets can be fair: A tool and fairification workflow for the enanomapper database. *Nanomaterials*, 10(10):1908, 2020.

M.-I. Limaylla-Lunarejo, N. Condori-Fernandez, and M. R. Luaces. Towards a fair dataset for non-functional requirements. In *Proceedings of the 38th ACM/SIGAPP Symposium on Applied Computing*, SAC '23, page 1414–1421, New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9781450395175.

S. Samonas and D. Coss. The cia strikes back: Redefining confidentiality, integrity and availability in security. *Journal of Information System Security*, 10(3), 2014.

A. A. Sinaci, F. J. Núñez-Benjumea, M. Gencturk, M.-L. Jauer, T. Deserno, C. Chronaki, G. Cangioli, C. Cavero-Barca, J. M. Rodríguez-Pérez, M. M. Pérez-Pérez, et al. From raw data to fair data: the fairification workflow for health research. *Methods of information in medicine*, 59(S 01):e21–e32, 2020.

M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne, et al. The fair guiding principles for scientific data management and stewardship. *Scientific data*, 3(1):1–9, 2016.

Hardening of a Continuous Behavior-based Authentication Distributed System

Julián González-Muñoz, Mario Casado, Daniel Garabato, Francisco J. Nóvoa, and Carlos Dafonte

Telematic Group, Faculty of Computer Science, Universidade da Coruña, 15071 A Coruña, Spain
CITIC Research Center, Computer Science and Information Technologies Department, Campus de Elviña, 15071 A Coruña, Spain
Correspondence: {j.gonzalezm, mario.diez, daniel.garabato, fjnovoa, carlos.dafonte}@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: Password-based traditional authentication systems are increasingly insufficient when it comes to providing security and checking the identity of the authenticated user. What happens when the password of an user has been stolen or an active user is not the same user who authenticated firstly?
A distributed system using AI (Artificial Intelligence) acting as a second factor authentication method by analyzing user's mouse events has to provide confidentiality and integrity in order to protect against different attacks such as Man-In-The-Middle that allow sniffing or data tampering, resulting in an identity spoof.
In order to grant integrity and confidentiality, encryption and authentication must be implemented. Authentication is used to allow one node to produce or consume data from an existent message stream and encryption in order to avoid exposing these data to external agents.
PKI (Public Key Infrastructure) system is widely used over the internet, so it is a trusty authentication and encryption framework. By using PKI in this project, hardening is performed by creating with OpenSSL a trusted Certificate Authority that issues and signs the certificates used by each node in the distributed system. Trust in this Certificate Authority is implemented by creating keystores and truststores for each node with keytool. This project resulted in a secure communication system preventing data from being sniffed or tampered.

1 Introduction

Password-based traditional authentication systems are increasingly insufficient when it comes to providing security and checking the identity of the authenticated user. Biometric identification, physical keys or password policies came up as different solutions to manage these problems, but, even so, these solutions act as a first security barrier. Incident handling techniques or security controls try to increase trustness in traditional authentication systems by using security policies in order to force re-authentication periodically, dificulting identity spoofing of an authenticated user but falling again in the problem of using one isolated event as an identity checking system. At the same time, forcing an user to re-authenticate periodically can worsen the user experience or the user performance during a job.

The need to apply a new authentication system based on a different method of identity demonstration comes up after analyzing the weaknesses of the traditional authentication methods. A distributed system has been built in order to provide a continuous behavior-based

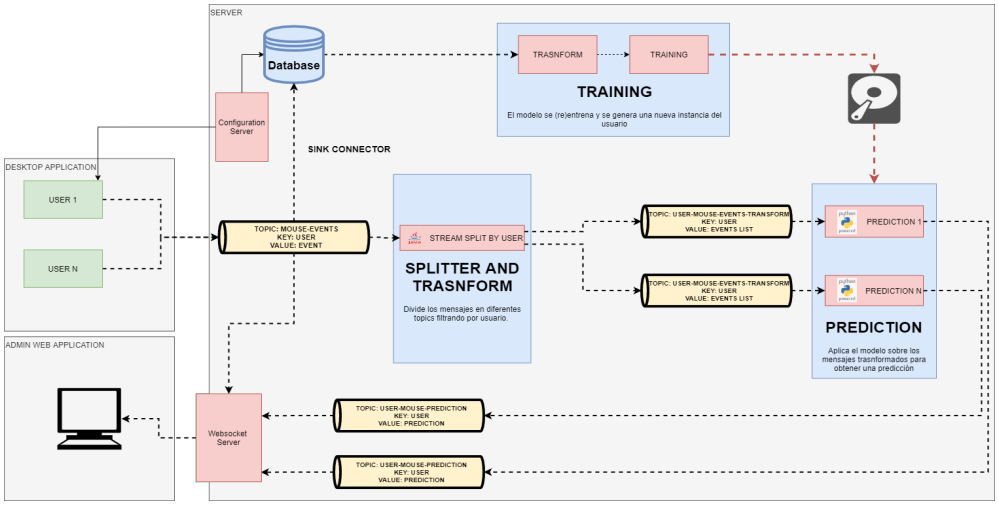


Figure 1: Distributed system deployed in this project

authentication method based on AI (Artificial Intelligence) acting as a second factor authentication method.

Once the system is working and providing authentication service, there is a need to apply security to its infrastructure and communications in order to avoid identity spoofing.

In Section 2 the architecture of the distributed system is explainend in order to contextualize the starting point. Next, in Section 3 the hardening methodology and PKI will be explained as well as the tools used to implement PKI. Finally, in Section 4 conclusions are discussed.

2 System Architecture

The distributed system used in this project uses Kafka messaging system to process all the information gathered from the user mouse movements [Narkhede et al. (2017)].

User mouse movements are sent via Kafka queues to the central node (broker) of the system. Once there, mouse events are splitted and transformed through Kafka Streams in order to use these transformed events as characteristics. The created characteristics are returned to the Kafka system via different queues splitted by user, these queues will be used by the final inference AI system to determine whether an active user is the one previously authenticated or not (Figure 1) [Silvelo et al. (2020)].

By using Confluent Kafka, we can encrypt communications with SSL and authenticate them with the use of different authentication protocols such as SASL, mTLS or HTTP Basic Auth.

3 Hardening of the system

A lot of advantages exist when having a distributed system like the one used in the project (Figure 1). One of them is that it allows the possibility to split each node of the system in different places or servers. In addition, the client-side application will be always communicating through the Internet.

Considering these scenarios and the possibility of a malicious user trying to tamper the data, the communications among nodes of the system must be protected through different methods. The first method of protection is to encrypt communications in order to avoid sniffing and tam-

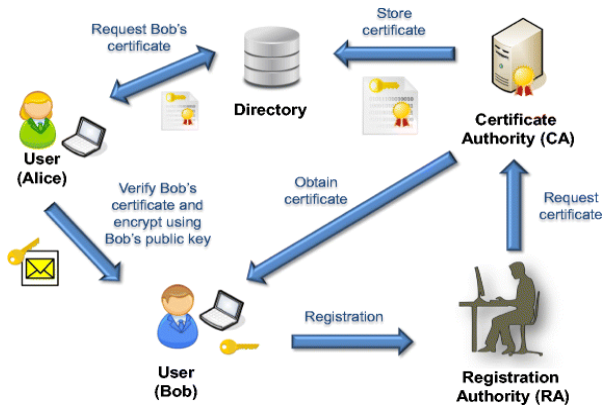


Figure 2: PKI simple schema [Tan et al. (2015)].

pering techniques. The other protection method consists of the authentication among nodes when consuming or producing data, this authentication would avoid malicious agents getting potentially confidential data from the system or producing and sending data to the system.

As mentioned in the previous chapter (Chapter 2) mTLS can be implemented in order to encrypt communications and perform authentication.

3.1 PKI

PKI is the framework that enables entities to securely exchange information using digital certificates. The components that form PKI include the hardware, software, policies, procedures and entities needed to safely distribute, verify and revoke certificates [Ellison (1999) Wood (2002)] (Figure 2).

Key elements of the PKI are:

- **Certificate Authority (CA):** Trusted party who provides certificates and authenticates their identity.
- **Registration Authority:** Party allowed by CA to issue certificates
- **Certificate store:** Enables programs running on the system to access stored certificates, certificate revocation lists (CRLs) and certificate trust lists (CTLs).
- **Certificate database:** This database stores information about issued certificates. In addition to the certificate itself, the database includes the validity period and status of each PKI certificate.

PKI has been used in this project by creating our own Root CA in order to issue all the needed digital certificates to encrypt and authenticate communications among nodes.

OpenSSL

OpenSSL is the world's most widely used implementation of the Transport Layer Security (TLS) protocol. It provides a set of command-line tools that serve a variety of purposes, including support for common PKI operations and TLS testing [Ristić (2023)].

The creation of a private CA is performed by using the OpenSSL command-line tool, creating a CA root certificate which is used to sign all the CSR (Certificate Signing Request) of the distributed system. Signed certificates of each node will be then stored in Keytool's keystores and the private CA will be trusted by all of the nodes by using Keytool's truststores.

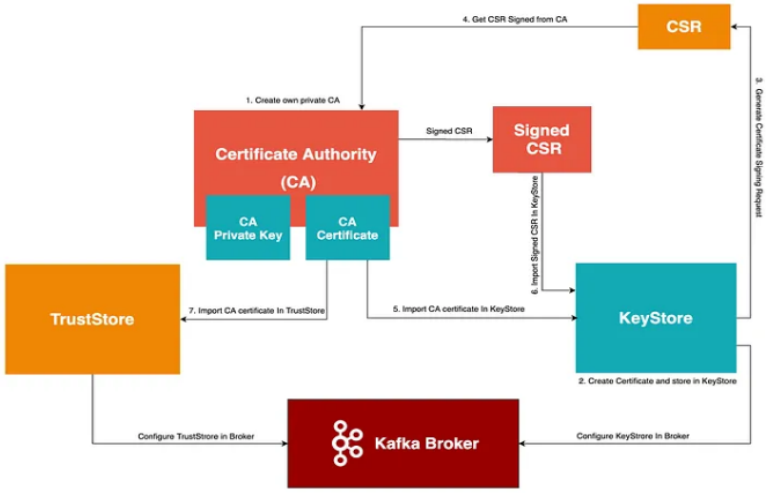


Figure 3: TLS configuration example in the Kafka Broker [Pandey (2020)]

Keytool

Keytool is a Java command-line tool which allows us to manage a keystore (database) of cryptographic keys, X.509 certificate chains, and trusted certificates [Oracle (2020)]. In this project, PKCS12 databases were created with Keytool to store signed certificates signed by the created CA (keystore) as well as the CA root certificate (truststore). Both keystores and truststores can contain more than one certificate.

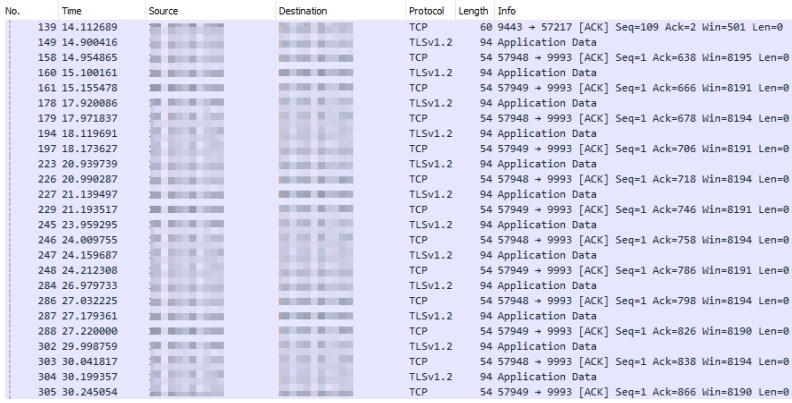
4 Conclusions

The creation of a private CA with OpenSSL (Section 3.1) allows to sign the CSR of each node thus creating each node's signed certificate and storing it in their Keytool's keystores. These signed certificates will be trusted by using Keytool's truststores where the CA root certificate is stored. This implementation can be shown in the Figure 3 where the structure of the mTLS communication configuration is represented.

The implementation of continuous authentication as a second factor method is a need nowadays, it provides an extra layer of security in the systems but it is not enough if it is not properly hardened. By using PKI, data flows in the distributed system are encrypted and authenticated, preventing identity spoofing from malicious actors (Figure 4).

Acknowledgements

This work was funded by the Spanish MCIN/AEI/10.13039/501100011033 and European Union Next Generation EU/PRTR through grant TED2021-130492B-C21 and the Galician Regional Government, Xunta de Galicia, through grants ED431B 2021/36 and ED431G 2019/01



No.	Time	Source	Destination	Protocol	Length	Info
139	14.112689	[redacted]	[redacted]	TCP	60	9443 → 57217 [ACK] Seq=109 Ack=2 Win=501 Len=0
149	14.900416	[redacted]	[redacted]	TLSv1.2	94	Application Data
158	14.954865	[redacted]	[redacted]	TCP	54	57948 → 9993 [ACK] Seq=1 Ack=638 Win=8195 Len=0
160	15.100161	[redacted]	[redacted]	TLSv1.2	94	Application Data
161	15.155478	[redacted]	[redacted]	TCP	54	57949 → 9993 [ACK] Seq=1 Ack=666 Win=8191 Len=0
178	17.920086	[redacted]	[redacted]	TLSv1.2	94	Application Data
179	17.971837	[redacted]	[redacted]	TCP	54	57948 → 9993 [ACK] Seq=1 Ack=678 Win=8194 Len=0
194	18.119691	[redacted]	[redacted]	TLSv1.2	94	Application Data
197	18.173627	[redacted]	[redacted]	TCP	54	57949 → 9993 [ACK] Seq=1 Ack=706 Win=8191 Len=0
223	20.939739	[redacted]	[redacted]	TLSv1.2	94	Application Data
226	20.990287	[redacted]	[redacted]	TCP	54	57948 → 9993 [ACK] Seq=1 Ack=718 Win=8194 Len=0
227	21.139497	[redacted]	[redacted]	TLSv1.2	94	Application Data
229	21.193517	[redacted]	[redacted]	TCP	54	57949 → 9993 [ACK] Seq=1 Ack=746 Win=8191 Len=0
245	23.959295	[redacted]	[redacted]	TLSv1.2	94	Application Data
246	24.009755	[redacted]	[redacted]	TCP	54	57948 → 9993 [ACK] Seq=1 Ack=758 Win=8194 Len=0
247	24.159687	[redacted]	[redacted]	TLSv1.2	94	Application Data
248	24.212308	[redacted]	[redacted]	TCP	54	57949 → 9993 [ACK] Seq=1 Ack=786 Win=8191 Len=0
284	26.979733	[redacted]	[redacted]	TLSv1.2	94	Application Data
286	27.032225	[redacted]	[redacted]	TCP	54	57948 → 9993 [ACK] Seq=1 Ack=798 Win=8194 Len=0
287	27.179361	[redacted]	[redacted]	TLSv1.2	94	Application Data
288	27.220000	[redacted]	[redacted]	TCP	54	57949 → 9993 [ACK] Seq=1 Ack=826 Win=8190 Len=0
302	29.998759	[redacted]	[redacted]	TLSv1.2	94	Application Data
303	30.041817	[redacted]	[redacted]	TCP	54	57948 → 9993 [ACK] Seq=1 Ack=838 Win=8194 Len=0
304	30.199357	[redacted]	[redacted]	TLSv1.2	94	Application Data
305	30.245054	[redacted]	[redacted]	TCP	54	57949 → 9993 [ACK] Seq=1 Ack=866 Win=8190 Len=0

Figure 4: Encrypted communications between desktop client and broker

Bibliography

C. M. Ellison. The nature of a useable pki. *Computer Networks*, 31(8):823–830, 1999.

N. Narkhede, G. Shapira, and T. Palino. *Kafka: the definitive guide: real-time data and stream processing at scale*. " O'Reilly Media, Inc.", 2017.

Oracle. keytool — docs.oracle.com. *https://docs.oracle.com/javase/8/docs/technotes/tools/unix/keytool.html*, 2020.

M. Pandey. Kafka SSL: Setup with self signed certificate—Part 1 — medium.com. *https://medium.com/jinternals/kafka-ssl-setup-with-self-signed-certificate-part-1-c2679a57e16c*, 2020.

I. Ristić. Library: OpenSSL Cookbook 3ed — Feisty Duck — feistyduck.com. *https://www.feistyduck.com/library/openssl-cookbook/*, 2023.

A. Silvelo, D. Garabato, R. Santoveña, and C. Dafonte. A first approach to authentication based on artificial intelligence for touch-screen devices. *Proceedings*, 54(1), 2020.

S.-Y. Tan, W.-C. Yau, and B.-H. Lim. An implementation of enhanced public key infrastructure. *Multimedia Tools and Applications*, 74(16):6481–6495, Aug 2015.

D. Wood. Pki, the what, the why, and the how. *GSEC–SANS Institute*, 2002.

Method for Pricing Renewable Energy Certificates

María A. Baamonde-Seoane, María del Carmen Calvo-Garrido and Carlos Vázquez

Department of Mathematics, Faculty of Computer Science, and Research Centre CITIC, Universidade da Coruña, 15071 A Coruña, Spain
Correspondence: maria.baamonde1@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: In this work we present one valuation method for Renewable Energy Certificates (RECs). Starting from a system of FBSDEs and using Ito lemma, we propose a mathematical model based on a semilinear PDE with two stochastic factors. The main novelty comes from the use of the Bermúdez-Moreno algorithm to deal the non-linear convective term in the PDE. This duality algorithm is based on the Yosida regularization of non-linear maximal monotone operators. The resulting linear problem is discretized by using a characteristics method combined with a second order implicit finite differences scheme. We show illustrative results of the performance of the proposed model and the numerical method.

1 Introduction

In recent years, several governments have developed environmental policies for promoting renewable energy sources. Many countries have already adopted Renewable Portfolio Standards (RPSs) and trading of renewable energy certificates (RECs). Markets for tradable RECs can be used to encourage the growth of a particular type of renewable energy, as in the case of Solar Renewable Energy Certificates (SRECs) (see M. Coulon, J. Khazaei, W. B. Powell (2015) and A. Shrivats, S. Jaimungal (2020)). In the present work, assuming that the price of the certificate depends on two stochastic factors which are the accumulated green certificates and the renewable energy production rate, we present the PDE model that governs the valuation of such financial instruments and we propose an appropriate numerical method for its solution.

For the numerical solution of the nonlinear PDE problem, we first apply the Bermúdez-Moreno algorithm proposed in A. Bermúdez, C. Moreno (1981) to deal with the nonlinear convective term. This duality method is based on the approximation of a nonlinear maximal monotone operator by means of its Yosida regularization. In order to solve the obtained linearized problem, we use numerical methods based on semi-Lagrangian schemes in the direction without diffusion while an implicit second order finite differences scheme is applied in the direction with diffusion term. Finally, several numerical examples are presented to illustrate the good performance of the method and model.

2 Mathematical modelling

In what follows, for a fixed time horizon representing the end of the compliance period $T > 0$, we assume that the source of randomness in the model is given by one-dimensional Wiener process $W = (W_t)_{0 \leqslant t \leqslant T}$. We assume that this Wiener process is defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and it is adapted to the filtration $\mathbb{F} = \{\mathcal{F}_t,\ t \geqslant 0\}$.

Assuming that the price of a green certificate depends on two stochastic factors, which are the renewable generation rate, G_t, and the number of accumulated green certificates, B_t, we aim to address the model for pricing a REC.

First, let us denote by $\tilde{G}_t = \ln(G_t)$ the Ornstein-Uhlenbeck (OU) process, which satisfies the following stochastic differential equation (SDE):

$$d\tilde{G}_t = \alpha_g \left(f(t) + \frac{\beta_g}{\alpha_g} P_t - \tilde{G}_t \right) dt + \sigma_g dW_t, \tag{32.1}$$

assuming that $G_{t_0} = g_0$, and where α_g is the mean reversion speed of the process, P_t is the certificate price, β_g is the parameter which controls the level of immediate feedback from the price of the certificate, σ_g is the volatility and W_t is the Wiener process governing the stochastic part of the equation. Moreover, taking into account that weather conditions strongly affect the production of energy in the renewable power generation, we introduce a deterministic function $f(t)$ representing the seasonality effect. A common choice is to use a combination of trigonometric functions as it is proposed in M. Coulon, J. Khazaei, W. B. Powell (2015). Furthermore, in the OU process (32.1), the mean reversion level is linear in P_t plus seasonality. Now, the production of renewable energy G_t can be written as

$$G_t = \exp\left(\tilde{G}_t\right).$$

Secondly, we introduce the dynamics of the number of accumulated renewable energy certificates, B_t,

$$dB_t = G_t dt.$$

Note that B_t is non-negative and non-decreasing, and we assume that $B_{t_0} = 0$, where t_0 is the beginning of the compliance period.

3 Statement of the pricing PDE

If we denote by $P = P(t, \tilde{G}, B)$ the price of the renewable energy certificate at time t, by using a dynamic hedging technique and applying Itô's Lemma (see K. Itô (1951)), we can derive the following nonlinear PDE whose solution is the price of the REC:

$$\mathcal{L}[P] = \frac{\partial P}{\partial t} + \frac{1}{2}\sigma_g^2 \frac{\partial^2 P}{\partial \tilde{G}^2} + \alpha_g \left(f(t) - \tilde{G}\right) \frac{\partial P}{\partial \tilde{G}} + \beta_g P \frac{\partial P}{\partial \tilde{G}} + \exp(\tilde{G}) \frac{\partial P}{\partial B} - rP = 0, \tag{32.2}$$

where r is the constant risk free interest rate.

Assuming that the number of life years of the certificate is denoted by γ and the maturity of the certificate is T, the PDE problem associated to (32.2) is initially formulated in the unbounded domain $(T-\gamma, T) \times (-\infty, \infty) \times (0, \infty)$.

For the particular case of one single period (i.e. one year, $\gamma = 1$), the payoff at the expiry date of the certificate, T, is a decreasing function in the number of accumulated green certificates at maturity, B, and depends on the requirement on the percentage of energy obtained from renewables at maturity, R_T. Thus, in order to state the PDE problem, Equation (32.2) is completed with the final condition

$$P(T, \tilde{G}, B) = \pi_T \mathbb{1}_{\{B < R_T\}}, \tag{32.3}$$

where π_T is the penalty amount π at time T.

Moreover, there exists the possibility of extending the problem to multiple periods. In that case, at each compliance date, T^i, for $i = 1, ..., \gamma - 1$, a jump condition must be applied. Thus, when the obligation is set the value of the certificate is given by

$$P(T^i, \tilde{G}, B) = \max\left(\pi_{T^i} \mathbb{1}_{\{B < R_i\}}, P\left(T^i_+, \tilde{G}, \max(0, B - R_i)\right)\right). \tag{32.4}$$

Note that T^i corresponds to the end of the i-th life year of the certificate and R_i is the requirement at that time.

The study of the existence and uniqueness of solution for the nonlinear PDE problems defined by (32.2)-(32.3) or by (32.2)-(32.4) remains as an open problem as it is mentioned in M. A. Baamonde-Seoane, M. C. Calvo-Garrido, M. Coulon, C. Vázquez (2021) and M.A. Baamonde-Seoane, M.C. Calvo-Garrido, C. Vázquez (2023).

4 Numerical techniques

4.1 The duality algorithm

As previously pointed out, the PDE problem (32.2) presents a non-linear convective term. One possibility to deal with this non-linearity is based on the Bermúdez-Moreno algorithm involving the Yosida regularization of non-linear maximal monotone operators (see A. Bermúdez, C. Moreno (1981)).

For this purpose, let us introduce the maximal monotone operator m, defined by

$$m(P) = \begin{cases} 0, & \text{if } P < 0 \\ P^2, & \text{if } P \geqslant 0, \end{cases}$$

so that

$$P\frac{\partial P}{\partial \tilde{G}} \approx \frac{1}{2}\frac{\partial m(P)}{\partial \tilde{G}}.$$

Therefore, the equation (32.2) can be written in the form:

$$\frac{\partial P}{\partial t} + \frac{\sigma_g^2}{2}\frac{\partial^2 P}{\partial \tilde{G}^2} + \alpha_g\left(f(t) - \tilde{G}\right)\frac{\partial P}{\partial \tilde{G}} + \frac{\beta_g}{2}\frac{\partial m(P)}{\partial \tilde{G}} + \exp(\tilde{G})\frac{\partial P}{\partial B} - rP = 0. \tag{32.5}$$

Following the duality technique introduced in A. Bermúdez, C. Moreno (1981), in terms of the constant parameter $\omega > 0$, we introduce the new additional unknown θ, defined by

$$\theta = (m - I\omega)(P)$$

where I denotes the identity operator.

Next, by using the Bermúdez-Moreno lemma, we have the equivalence

$$\theta = m(P) - \omega P \Leftrightarrow \theta = m_\lambda^\omega(P + \lambda\theta), \tag{32.6}$$

where m_λ^ω denotes the Yosida approximation of $m - I\omega$ with parameter λ. For convergence purposes, we impose the relation $2\lambda\omega = 1$ in the choice of the parameters λ and ω. Under this constraint, the Yosida approximation can be analytically computed and is given by

$$m_\lambda^\omega\left(P + \frac{\theta}{2\omega}\right) = \begin{cases} -\theta - 2\omega P, & \text{if } P + \frac{\theta}{2\omega} \leqslant 0, \\ \theta + 2\omega P + \omega^2 - \omega\sqrt{4\theta + 8\omega P + \omega^2}, & \text{if } P + \frac{\theta}{2\omega} \geqslant 0. \end{cases}$$

Next, if we introduce the linear differential operator

$$\mathcal{L}[P] = \frac{\partial P}{\partial t} + \frac{\sigma_g^2}{2}\frac{\partial^2 P}{\partial \tilde{G}^2} + \alpha_g\left(f(t) - \tilde{G}\right)\frac{\partial P}{\partial \tilde{G}} + \frac{\beta_g\omega}{2}\frac{\partial P}{\partial \tilde{G}} + \exp(\tilde{G})\frac{\partial P}{\partial B} - rP, \tag{32.7}$$

the equation (32.5) can be rewritten in the form:

$$\mathcal{L}[P] = -\frac{\beta_g}{2}\frac{\partial\theta}{\partial\tilde{G}}. \tag{32.8}$$

Moreover, from the equivalence stated in (32.6), the equation (32.8) is coupled with the following non-linear equation:

$$\theta = m_\lambda^\omega(P + \lambda\omega).$$

4.2 Formulation of the PDE problem in a bounded domain

In order to apply numerical discretization using finite differences, it is necessary to define the bounded domain of the PDE problem.

Let $\Omega = (T-\gamma, T) \times \mathbb{R} \times (0, +\infty)$ be the initial unbounded domain. Moreover, let $\bar{\Omega} = (T-\gamma, T) \times (0, \hat{b}) \times (-\bar{g}, \bar{g})$ be the truncated bounded domain where $\hat{b}$ and $\bar{g}$ are real numbers, which are influenced by the requirement of the payoff function and the jump conditions at compliance dates. Now, we introduce the changes $\hat{B} = \frac{B}{\tilde{b}}$ and $\hat{G} = \frac{\tilde{G}}{\tilde{g}}$ with $\hat{g} = 2\tilde{g}$, so the bounded spatial domain $\Omega^* = (0,1) \times (0,1)$ in the new variables $(t, \hat{B}, \hat{G})$, whose boundary can be decomposed as $\Gamma = \bigcup_{i=1}^{2}(\Gamma_i^- \cup \Gamma_i^+)$ where

$$\Gamma_i^- = \{(y_1, y_2) \in \Gamma |\ y_i = 0\},\ \ \Gamma_i^+ = \{(y_1, y_2) \in \Gamma |\ y_i = 1\},\ i = 1, 2.$$

Next, as in M. A. Baamonde-Seoane, M. C. Calvo-Garrido, M. Coulon, C. Vázquez (2021), we follow the methodology introduced by O. A. Oleinik (1973) to obtain the boundaries where it is necessary to impose boundary conditions. On those boundaries, we will impose homogeneous Neumann boundary conditions.

4.3 Discretization of the PDE

In order to choose an appropriate time discretization scheme for the PDE (32.8), we note that the linear differential operator (32.7) is degenarate. Thus, we follow the idea first proposed in Y. d'Halluin, P. A. Forsyth, G. Labahn (2005), which consists of choosing a semi-Lagrangian method in the direction without diffusion combined with a Crank-Nicolson finite differences scheme in the direction with diffusion.

For the time discretization, we first consider the change of time variable $\tau = T - t$, where τ represents the time to maturity. Therefore, equation (32.8) can be equivalently written in the domain as follows

$$\frac{\mathrm{D}P}{\mathrm{D}\tau} - \mathcal{A}P = 0, \tag{32.9}$$

where

$$\frac{\mathrm{D}P}{\mathrm{D}\tau} = \frac{\partial P}{\partial\tau} - \hat{b}\exp(\hat{G}\hat{g} - \bar{g})\frac{\partial P}{\partial\hat{B}},$$

$$\mathcal{A}P = \frac{\hat{g}^2\sigma_g^2}{2}\frac{\partial^2 P}{\partial\hat{G}^2} + \hat{g}\alpha_g\left(f(T-\tau) - \left(\hat{G}\hat{g} - \bar{g}\right) + \frac{\beta_g\omega}{2\alpha_g}\right)\frac{\partial P}{\partial\hat{G}} + \frac{\hat{g}\beta}{2}\frac{\partial\theta}{\partial\hat{G}} - rP.$$

Next, we introduce the approximation for the material derivative:

$$\frac{\mathrm{D}P}{\mathrm{D}\tau} \approx \frac{P\left(\tau^{n+1}, \hat{B}, \hat{G}\right) - P\left(\tau^n, \chi^n\left(\hat{B}, \hat{G}\right), \hat{G}\right)}{\Delta\tau},$$

where $\chi^n\left(\hat{B},\hat{G}\right)=\chi(\tau^n)=\hat{B}+\Delta\tau\hat{b}\exp(\hat{G}\hat{g}-\bar{g})$ is the solution for $n=0,1,\ldots,N_T-1$ and represents the position at time τ^n of the point placed at $(\hat{B},\hat{G})$ at time τ^{n+1} and moving according to the velocity field $v=-\hat{b}\exp(\hat{G}\hat{g}-\bar{g})$.

By using a Crank-Nicolson scheme ($\hat{\theta}=0.5$ in the so called $\hat{\theta}$-method) for the second order differential term $\mathcal{A}P$ in equation (32.9), we obtain:

$$\begin{aligned}
&\frac{P^{n+1}-P^n\circ\chi^n}{\Delta\tau}-\frac{\hat{\theta}\hat{g}^2\sigma_g^2}{2}\frac{\partial^2P^{n+1}}{\partial\hat{G}^2}-\frac{(1-\hat{\theta})\hat{g}^2\sigma_g^2}{2}\frac{\partial^2\left(P^n\circ\chi^n\right)}{\partial\hat{G}^2}\\
&\quad-\hat{\theta}\hat{g}\alpha_g\left(f(T-\tau)-\left(\hat{G}\hat{g}-\bar{g}\right)+\frac{\beta_g\omega}{2\alpha_g}\right)\frac{\partial P^{n+1}}{\partial\hat{G}}\\
&\quad-(1-\hat{\theta})\hat{g}\alpha_g\left(f(T-\tau)-\left(\hat{G}\hat{g}-\bar{g}\right)+\frac{\beta_g\omega}{2\alpha_g}\right)\frac{\partial\left(P^n\circ\chi^n\right)}{\partial\hat{G}}\\
&+r\hat{\theta}P^{n+1}+r(1-\hat{\theta})\left(P^n\circ\chi^n\right)=\frac{\hat{\theta}\hat{g}\beta_g}{2}\frac{\partial\theta^{n+1}}{\partial\hat{G}}+\frac{(1-\hat{\theta})\hat{g}\beta_g}{2}\frac{\partial\theta^n}{\partial\hat{G}}.
\end{aligned}\tag{32.10}$$

At each time step, the equation (32.10) is coupled with the following non-linear relation between P^{n+1} and θ^{n+1}:

$$\theta^{n+1}=m_\lambda^\omega(P^{n+1}+\lambda\omega^{n+1}).\tag{32.11}$$

Next, we propose a fixed point algorithm to approximate the solution of the non-linear problem (32.10)-(32.11).

At each fixed point iteration, the full discretization of problem can be written as follows:

$$\begin{aligned}
&\frac{P_{i,j}^{n+1,k+1}-P_{i,j}^n\circ\chi^n}{\Delta\tau}-\frac{\hat{\theta}\hat{g}^2\sigma_g^2}{2}\left(\frac{P_{i,j+1}^{n+1,k+1}-2P_{i,j}^{n+1,k+1}+P_{i,j-1}^{n+1,k+1}}{\left(\Delta\hat{G}\right)^2}\right)\\
&\quad-\frac{(1-\hat{\theta})\hat{g}^2\sigma_g^2}{2}\left(\frac{P_{\chi^n,j+1}^n-2P_{\chi^n,j}^n+P_{\chi^n,j-1}^n}{\left(\Delta\hat{G}\right)^2}\right)\\
&\quad-\hat{\theta}\hat{g}\alpha_g\left(f(T-\tau^{n+1})-\left(\hat{G}_j\hat{g}-\bar{g}\right)+\frac{\beta_g\omega}{2\alpha_g}\right)\left(\frac{P_{i,j+1}^{n+1,k+1}-P_{i,j-1}^{n+1,k+1}}{2\Delta\hat{G}}\right)\\
&\quad-(1-\hat{\theta})\hat{g}\alpha_g\left(f(T-\tau^n)-\left(\hat{G}_j\hat{g}-\bar{g}\right)+\frac{\beta_g\omega}{2\alpha_g}\right)\left(\frac{P_{\chi^n,j+1}^n-P_{\chi^n,j-1}^n}{2\Delta\hat{G}}\right)\\
&\quad-\frac{\hat{\theta}\hat{g}\beta_g}{2}\left(\frac{\theta_{i,j+1}^{n+1,k}-\theta_{i,j-1}^{n+1,k}}{2\Delta\hat{G}}\right)-\frac{(1-\hat{\theta})\hat{g}\beta_g}{2}\left(\frac{\theta_{i,j+1}^n-\theta_{i,j-1}^n}{2\Delta\hat{G}}\right)\\
&\quad+r\hat{\theta}P_{i,j}^{n+1}+r(1-\hat{\theta})P_{\chi^n,j}^n=0.
\end{aligned}$$

where $\hat{\theta}=0.5$, $P_{r,s}^{l,m}=P^m(\tau^l,\hat{B}_r,\hat{G}_s)$, $P_{\chi^l,s}^{l,m}=P^m(\tau^l,\chi^l,\hat{G}_s)$ and $\theta_{r,s}^{l,m}=\theta^m(\tau^l,\hat{B}_r,\hat{G}_s)$.

5 Numerical examples

5.1 Academic test

As a sanity check of the code and numerical methods, in the first example we show an academic test with known analytical solution. For this purpose, we consider the

following non homogeneous non-linear PDE:

$$\mathcal{L}[P] = h,$$

where the differential operator $\mathcal{L}$ is defined by (32.2) and h is given by

$$h(t,B,\tilde{G}) = \left[-B\tilde{G} + \frac{1}{2}\sigma_g^2 t^2 B^2 - tB\alpha_g \left(f(t) + \frac{\beta_g}{\alpha_g}\exp\left((T-t)B\tilde{G}\right) - \tilde{G}\right) - \exp(\tilde{G})t\tilde{G} - r\right] P(t,B,\tilde{G}), \qquad (32.12)$$

with $P(t,B,\tilde{G}) = \exp\left((T-t)B\tilde{G}\right)$ is the analytical solution of the PDE (32.12).

By choosing $\hat{b} = 1$ and $\tilde{g} = 0.5$ for the change of variables, we pose the PDE problem in the bounded domain $\tilde{\Omega} = [0,1]\times[0,1]\times[0,1]$ with Dirichlet boundary conditions on Γ_1^+, Γ_2^- and Γ_2^+ that are given by the evaluation of the solution at the corresponding boundaries.

Table 1: Parameters in the PDE model for the academic test.

Parameter	T	γ	α_g	β_g	σ_g	r	ω	ϵ
Value	1	1	2	1.27×10^{-3}	0.1863	0.02	2	10^{-5}

In this academic test we do not include the seasonality effect, so that we take $f = 0$. Parameters in the PDE are collected in Table 1 and mostly taken from M. Coulon, J. Khazaei, W. B. Powell (2015).

Table 2: Relative errors and empirical convergence order in academic test.

Time steps	**Space steps**	**Error**	**R**	**Order**
40	32	0.0108651	-	-
80	64	0.0055962	1.9415	0.9572
160	128	0.0028748	1.9466	0.9610
320	256	0.0014710	1.9543	0.9667
640	512	0.0007474	1.9681	0.9768
1280	1024	0.0003780	1.9771	0.9834
2560	2048	0.0001906	1.9832	0.9879

Table 2 shows the errors, convergence ratio and empirical order of convergence with different time and spatial discretizations computed as in Y. d'Halluin, P. A. Forsyth, G. Labahn (2005). Thus, we can conclude that a first order convergence is achieved.

5.2 Real case

In this example, we have used the real New Jersey market data presented for SREC markets, i.e., markets for solar renewable energy certificates, in M. Coulon, J. Khazaei, W. B. Powell (2015). In this market, the energy year refers to the 12-month period ending on May 31. We assume that the maturity is $T = 13$, i.e., May 31, 2013. For convenience, the energy year 2013 is defined as the time interval $(12, 13]$. Thus, we consider the requirement schedule 2010-2013, i.e., $\gamma = 3$, with initial year $t = t_0 = T - \gamma = 10$ and the first compliance date at $t = 11$ (at the end of

the year 2011) which corresponds to $\tau = 2$. We consider the rest of the PDE parameters as those in the Table 2. Moreover, the seasonality function is chosen as $f(s) = -0.1209\sin(4\pi s) + 0.0900\cos(4\pi s) + 0.2151\sin(2\pi s) + 0.3859\cos(2\pi s)$ and represents the influence of weather conditions.

The requirement, R_i, and penalty, π_{T^i}, values at the end of each year for $i = 1, 2, 3$ are indicated in M. A. Baamonde-Seoane, M. C. Calvo-Garrido, M. Coulon, C. Vázquez (2021), M.A. Baamonde-Seoane, M.C. Calvo-Garrido, C. Vázquez (2023) and M. A. Baamonde-Seoane, M. C. Calvo-Garrido, C. Vázquez (2023). For the numerical methods, we start by choosing $\hat{b} = 7 \times 10^5$ and $\bar{g} = \ln(7 \times 10^5)$. Concerning the discretization parameters, we consider $\Delta\tau = \frac{1}{1200}$, and a uniform mesh with $\Delta\hat{B} = \Delta\hat{G} = 1/32$.

On the one hand, in Figure 1 we can observe that when the accumulated supply (B) and renewable energy (G) are low, the price of the certificate tends to the penalty value. On the other hand, when the value of both state variables increases, the price of the certificate decreases.

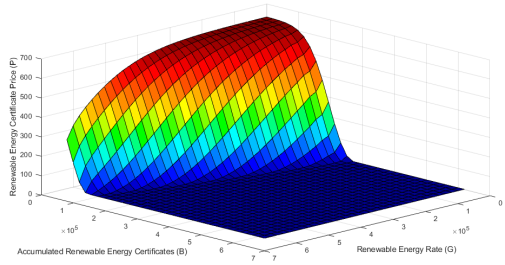


Figure 1: Renewable energy certificate price at time $t = T - 2/3$ in the real test.

In Figure 2 we can observe that, for values of accumulated certificates nearer to requirement, low values of generation rate are linked to certificate prices equal to the penalty amount. Additionally, for lower values of banked certificates, high enough values of generate rate are associated with prices equal to the penalty.

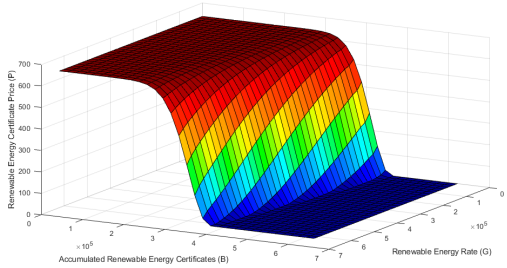


Figure 2: Renewable energy certificate price at time $t = T - 1/3$ in the real test.

Finally, in Figure 3 we represent the price of the certificate versus the number of accumulated renewable energy certificates for different times, obtaining some cross-sectional plots. At $t = T$, the REC price matches the penalty if the requirement is not met, otherwise the price is zero. Then, as we move backwards in time, the curves

take lower values and move to the left, due to the diffusion of the final value. This behaviour can be also observed in M.A. Baamonde-Seoane, M.C. Calvo-Garrido, C. Vázquez (2023) and M. Coulon, J. Khazaei, W. B. Powell (2015).

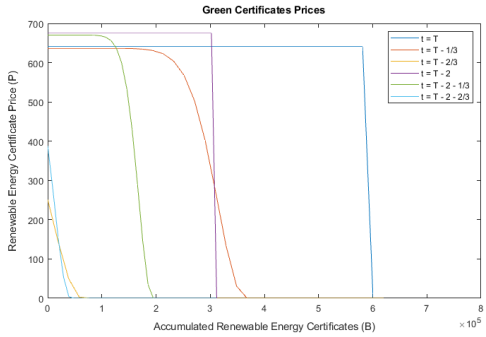


Figure 3: Price curves for different times in the real test.

Acknowledgements

CITIC is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS).

Bibliography

A. Bermúdez, C. Moreno. Duality methods for solving variational inequalities. *Computers & Mathematics with Applications*, 7:43–58, 1981.

A. Shrivats, S. Jaimungal. Optimal generation and trading in solar renewable energy certificate (SREC) markets. *Applied Mathematical Finance*, 27(1-2):99–131, 2020.

K. Itô. On stochastic differential equations. *Memoirs of the American Mathematical Society*, 4:1–51, 1951.

M. A. Baamonde-Seoane, M. C. Calvo-Garrido, C. Vázquez. Model and numerical methods for pricing renewable energy certificate derivatives. *Communications in Nonlinear Science and Numerical Simulation*, 118:107066, 2023.

M. A. Baamonde-Seoane, M. C. Calvo-Garrido, M. Coulon, C. Vázquez. Numerical solution of a nonlinear PDE model for pricing Renewable Energy Certificates (RECs). *Applied Mathematics and Computation*, 404:126199, 2021.

M. Coulon, J. Khazaei, W. B. Powell. SMART-SREC: A stochastic model of the new jersey solar renewable energy certificate market. *Journal of Environmental Economics and Management*, 73:13–31, 2015.

M.A. Baamonde-Seoane, M.C. Calvo-Garrido, C. Vázquez. Pricing renewable energy certificates with a Crank–Nicolson Lagrange–Galerkin numerical method. *Journal of Computational and Applied Mathematics*, 422:114891, 2023.

E. R. O. A. Oleinik. *Second order equations with nonnegative characteristic form*. American Mathematical Society, Providence, Rhode Island and Plenum Press, New York-London, 1973.

Y. d'Halluin, P. A. Forsyth, G. Labahn. A semi-langrangian approach for American Asian options under jump-diffusion. *SIAM Journal on Scientific Computing*, 27:315–345, 2005.

Development and forensic study of a ransomware for Android 8.1 devices.

Alfonso Torralba Mantiñán, Cristina López Bravo, José Luis Rivas López

Ciberseguridad AA.PP., Ingeniería de Sistemas para la Defensa de España S.A. S.M.E. M.P., 28040 Madrid, España
Grupo de Tecnologías de la Información (GTI), atlanTTIc, Universidade de Vigo, 36310 Vigo, España
Laboratorio de Informática Forense Europeo - LIFe, 36202 Vigo, España
Correspondence: a.torralba@udc.es & atorralba@isdefe.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: The world of technology is under attack all the time. The reasons for this range from economic to political and, as a result, there is a need for global awareness of the risks involved. With this, there is also a need for continuous training of cybersecurity professionals. Of all the attacks that cause the most damage to society, especially in the economic sphere, ransomware is the one that leads the ranking. This fact defined the first objective of the Master's thesis presented in this article: the design of a mobile ransomware for devices with Android 8.1 operating system. The aim was to investigate the functioning of ransomware-type viruses at a low level, as well as other related aspects. From the first objective arose the second: to carry out computer forensic studies targeting the previously designed virus. These reports are intended to be used for educational purposes, serving as a procedural guide for university professors or professionals in the sector who are interested in virus forensics.

Both objectives were successfully achieved. A ransomware virus was developed, hidden behind a so-called image gallery application. It encrypts certain images on the victim device and sends the encryption key to its own remote server. In addition, two forensic reports were produced in accordance with the appropriate standards. In these reports, each step of the virus analysis was explained in detail. A range of alternative tools to be used by the analyst during the analysis was also included.

This document is a brief summary of my Master's Thesis entitled «*Desarrollo y estudio forense de un ransomware para dispositivos Android 8.1*». The original document can be accessed via the following link: *https://github.com/torralba98/ransomware*

1 Introduction

The world of technology has been and will continue to be a great advance for civilization in general, allowing for the progress of globalization, as well as facilitating the day-to-day life of the world's population. However, good always goes hand in hand with evil. The countless number of cyberattacks that have occurred to date is well known. There are many reasons for these attacks, including economic factors and cyber warfare. There are numerous types of malware used by cyberattackers for such attacks, and new ones continue to emerge every day. Of these, special emphasis should be placed on ransomware. This type of malware has been classified as one of the most dangerous and has caused the greatest social repercussions

in recent years. So far in 2023, in the compilation of the most dangerous malware by Safety-Devices (Glamoslija, 2023), it can be seen that this type of malware already occupies the first position. Among the different repercussions that a malware causes to society, we highlight the dissemination of information, loss of data, blocking of equipment and economic loss. As it is one of the most common and at the same time most harmful malware, it is essential to train professionals with knowledge about it. It is of vital importance to know the different techniques used by cybercriminals to spread this malware and how they infect computers in order to have an advantage in the fight against them.

With this in mind, the first objective of the project was to **design a mobile ransomware** for Android 8.1. This version has been chosen taking into account that it is neither a very old nor a relatively new version. The motivation is to acquire new knowledge related to this type of virus, as the publicly available information is either practically non-existent or comes from unreliable sources. With this, we also proceeded to carry out the second objective, the **preparation of a forensic expertise**, with didactic purposes, aimed at detecting this type of malware, with the aim of making known the correct way to carry out this type of forensic report. This expertise is aimed at professionals in the sector and teachers, so it had to be complete and explanatory, in such a way that it would serve as a procedural guide.

2 Mobile ransomware

Of all the existing malware available today, as mentioned above, the design of a mobile ransomware is chosen, as it is considered to be a malware with a certain level of complexity. It would be very interesting to design this type of virus within the world of telephony in order to study, in addition to its operation, various techniques for its propagation and infection. There are numerous references to attacks related to this type of malware, among which we highlight the well-known **Wannacry** (Kaspersy, -) and **Petya** (Ivan Belcic, 2019) *ransomwares*. In our case, an encrypting ransomware was designed that will attack specific files. In this section, we will go into detail on the most important aspects of the designed ransomware. Topics such as its design, operation, functionalities, and the tools used for its development will be discussed.

2.1 Design

Broadly speaking, the design of our scenario can be summarized as shown in figure 1. In the image on the left, we can see the general architecture. It shows a user with his mobile device that, through an internet connection, communicates with the Raspberry. The latter will be in charge of providing the appropriate services from the server to the user. This reflected design applies to the following situations:

- **Situation 1**: Victim user accesses the web server with the intention of downloading the malicious application. The server provides the APK.
- **Scenario 2**: Infected mobile device communicates with the server to send certain information.

The image on the right shows how the technologies used on the server side interact with each other. Our Flask server, container of the web application, will be able to interact with the database making the relevant requests when necessary. Also, thanks to the Gradle tool, the server itself will be able to compile the source code of the malicious APK.

2.2 Operation

On the one hand, we have a server implemented with Python in combination with the **Flask** framework. This server supports a website from which the download of the malicious APK

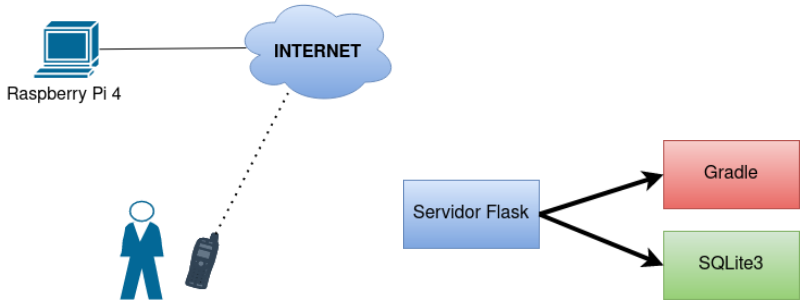


Figure 1: Project design.

(entry vector of the ransomware) will be offered. This server will also communicate with the malware. The latter will send to the former the encryption keys of the infected devices' information to the server, which will store this data in a local database (**SQLite3**), along with information about the victim device in order to identify it. Both the malicious APK and the server side have a common global configuration file that will facilitate the attacker's work. It was proposed to deploy our product on a **Raspberry Pi 4** so that, in case an attacker gains access to our server, he will not gain direct access to the files on our local machine, but to those on the low-cost board.

The implemented malware hides behind a fake application of **image gallery**. No vulnerability is exploited for its execution. In versions prior to *Android 10.0*, simply by the user accepting access permissions to the device's storage, it will grant full access to the contents of the device, which will allow browsing through its different directories and performing all the desired actions on the files contained therein. This is why it has been chosen to supplant this type of application, as these permissions are the ones requested as a general rule for its proper functioning, so it will not arouse suspicions. The ransomware itself, works by listing all the directories and files on the victim device to locate specific files, in our case, image files [1] These images will be encrypted with **AES128**, namely AES/CBC/PKCS5Padding. The encryption key, along with mobile device reference information, will be sent to the Flask server for storage, as mentioned above. Finally, the application will modify the mobile device's wallpaper to indicate to the victim that they have been infected and the conditions for paying the ransom to regain access to their encrypted files. In the event that the victim pays the ransom [2], the decryption key will be sent to the victim with the appropriate instructions.

Anti-analysis functionalities implemented

Various techniques are implemented in the malware itself to prevent its analysis and detection. All of them are explained below.

- **Virtual machine detection**. A function was designed to detect whether the malware is being executed in an emulator or in a virtual machine, in which case it will stop its execution.
- **Anti-debugging**. The use of debuggers was restricted to avoid inspections of the execution of the malware.

[1] Images are located, as it is considered that they may have sentimental value for the victim. If they are not backed up, it could force those affected to pay the ransom.

[2] It should be noted that in a real situation, in case of payment of the ransom, there is no guarantee that access to the encrypted files will be returned.

- **Failure to back up the application and its data**. This technique was implemented in order to prevent third party access to application data. The reason is that the use of encrypted file systems limits access to the device if it is turned off for an external attacker, however, it does not prevent other applications or processes on the device from reading data through the file system.
- **Obfuscation**. Implemented to increase the security of the code. It converts the original code into another one that is more difficult for humans to understand. It achieves the goal by applying encryption mechanics and patterns to prevent access to critical sections of the code.
- **Secure communication with the server**. All communications established between the malicious application and the server are encrypted, specifically through the use of **TLS**.
- **Polymorphism**. We want the malicious APK to be able to self-mute its own implementation, so that its hash signature will be different every time. This is intended to fool all those antiviruses that work on the basis of file signatures. In our project, polymorphism was implemented on the server side, so that for each new download of the virus, the source code self-modifies certain internal parameters resulting in a new hash signature.

2.3 Tools used in its development

Some technologies used during development are shown below.

- **Flask**. It is a *micro* framework implemented in Python that facilitates the development of web applications under the model-view-controller pattern.
- **Database management system**. Of vital importance is the storage and management of information by our server. As a database management system, it was decided to use **SQLite3**. It is a simple, efficient, powerful and fast database management system which, in addition to the fact that it is developed as a relational database, makes it the ideal system for our situation.
- **Android Studio**. Tool chosen for the development of the malicious application, as it has all the functionalities and tools necessary to carry out the correct implementation of a mobile application in its entirety.
- **Gradle**. Included in the Android Studio toolkit. This is a suite of advanced build tools to automate and manage the application build process.
- **OpenSSL**. Used for the generation of the certificate that allows the establishment of secure communications between the server and the server.

Some development languages used were **SQL** (functions related to the management of the information stored in the database), **HTML, CSS and JavaScript** (used for the development of the server-side web interface), **Java** (to carry out the complete development of the malicious application) and **Python** (for the implementation of the entire server-side).

3 Forensic expertise

As the last point to be dealt with in the project, forensic tests have been carried out on the previously developed malware. Specifically, two have been carried out, one for a simple version of the virus (without anti-analysis techniques implemented) and another for a more complex version (with all the anti-analysis techniques implemented). The aim is to serve as a procedural guide for professionals in the computer forensic sector, providing information on the steps to follow for a proper analysis of the malware. Both reports have been drafted in accordance with the standard **UNE 197010:2015 - «Normas Generales para la elaboración**

de informes y dictámenes periciales sobre TIC» (Jorge Navarro Clérigues, 2016), as well as other standards of good practice in forensic reports, and reference was made to expertise such as **«Investigación, informe y certificación de validez de la firma digital generado por la aplicación GestionaDocs»** by Adrián Ramirez Correa (Correa, 2018).

When conducting forensic examinations, it is very important to extract and preserve a logical backup image of the infected mobile device itself. This will avoid problems such as, for example, someone modifying the device and thereby altering its content, which could cause erroneous results during forensic analysis. To perform such cloning and analysis, there are many tools available today, some of them free and some of them for a fee. Examples of the latter on the market today are **Atola Taskforce** (Atola Technology, -) and **Cellebrite Forensic Workstation** (Cellebrite, -). In our case, the use of free tools was chosen, in addition to the limited resources available, to demonstrate that it is not necessary to spend a lot of money to be able to carry out complete forensic expertise.

3.1 Tools used

A multitude of tools have been used to produce these reports, some of which are mentioned below.

- **MOBILEdit** (MOBIledit, -). For the cloning of the logical image of the evidence. It generates a summary report of the information contained in the logical image. This is a paid forensic software, of which the company provided us with a free trial version.
- **Santoku**. Linux distribution oriented towards forensic reporting. It will be used to obtain the APK from the infected device itself and its subsequent analysis.
- **Wireshark**. Network packet analyzer used to identify and analyze connections established by malware.
- **Windows XP Professional Service Pack 3 (32 bits)**. A sandbox is installed inside it. The reason is to test the malicious website to prevent it from infecting us, for example, in case it exploits a vulnerability in our web browser.

4 Results

It was possible to design a fully functional malware, as well as its server part, satisfying all the initially set objectives, on which various types of functional tests (white box and black box) have been carried out with successful results. Tests were carried out to try to identify the APK as malicious with various antivirus tools, for example with the **VirusTotal** tool, all yielding negative analyses. On the other hand, both expert documents have been drafted to a very high level of detail. A multitude of forensic tools for carrying out various types of forensic facets have been introduced. Likewise, the whole process has been explained in as much detail as possible.

Bibliography

Atola Technology. Atola taskforce. *https://atola.com/products/taskforce/*, -. [Online; accessed 12-September-2023].

Cellebrite. Cellebrite forensic workstation. *https://cellebrite.com/es/cellebrite-forensic-workstation-es/*, -. [Online; accessed 12-September-2023].

A. R. Correa. Investigación, informe y certificación de validez de la firma digital generado por la aplicación gestionadocs. *https://gestionadocs.com/wp-content/uploads/2021/05/gestioandocs-informe-pericial.pdf* , 2018. [Online; accessed 12-September-2023].

K. Glamoslija. 10 most dangerous virus & malware threats in 2023. *https://www.safetydetectives.com/blog/most-dangerous-new-malware-and-security-threats/*, 2023. [Online; accessed 12-September-2023].

Ivan Belcic. Ransomware petya: Cómo funciona y cómo protegerse. *https://www.avast.com/es-es/c-petya*, 2019. [Online; accessed 12-September-2023].

Jorge Navarro Clérigues. Guía actualizada para futuros peritos informáticos. Últimas herramientas de análisis forense digital. caso práctico. *https://www.pensamientopenal.com.ar/system/files/2016/05/doctrina43429.pdf*, 2016. [Online; accessed 12-September-2023].

Kaspersy. ¿qué es el ransomware wannacry? *https://www.kaspersky.es/resource-center/threats/ransomware-wannacry*, -. [Online; accessed 12-September-2023].

MOBIledit. Mobiledit forensic. *https://www.mobiledit.com/forensic-express*, -. [Online; accessed 12-September-2023].

Highway Remodeling: Harnessing Georeferenced Data for Procedural Modeling

Mario Noriega, José Ribelles, Antonio Seoane, Margarita Martínez, Javier Taibo, and José A. Iglesias-Guitián

Computer Vision Center, 08193 Bellaterra, Spain
Institute of New Imaging Technologies, Universitat Jaume I, 12071 Castellón de la Plana, Spain
Civil Engineering Department, Universidade da Coruña, 15071 A Coruña, Spain
Civil Engineering Department, Universitat Politècnica de Catalunya, 08034 Barcelona, Spain
CITIC, Universidade da Coruña, 15071 A Coruña, Spain
Correspondence: j.iglesias.guitian@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: This paper introduces a novel procedural modeling system for generating 3D highway models, leveraging real-world data inputs such as highway layouts and manual annotations of essential elements. Our highly parameterized system facilitates easy customization, including modifications to the number of lanes and other key features, broadening its applicability across diverse domains, from urban planning and transportation engineering to virtual simulations.

1 Introduction

Procedural geometry modeling (Ebert et al., 2002), is a powerful approach that excels in modeling objects with diverse variations governed by shared rules, distinct from user-driven editing. Procedural modeling in computer graphics typically leverages algorithms to systematically generate complex, rule-compliant 3D models, freeing designers and engineers from manual constraints. For instance, procedural modeling has been successfully applied in various domains such as plant modeling (Gasch et al., 2022), architecture (Schwarz and Müller, 2015), buildings (Müller et al., 2006), cities (Parish and Müller, 2001), and landscapes (Teoh, 2009). Regarding highways specifically, they also consist of a series of elements arranged in accordance with established rules and regulations (refer to Figure 1). Nevertheless, the prospect of locating two identical highway segments in the real world is highly improbable due to their inherent uniqueness. Highway modeling is of significant importance across diverse applications, such as civil engineering for quality assurance, cost-effective design, and creating digital twin counterparts. Its relevance also extends to multiple domains, including video game development, driving simulators, and virtual or augmented reality experiences. Consequently, the development of automated procedural systems capable of generating all essential highway geometry, complemented by robust parametrization to facilitate the creation of diverse variations based on an original layout, assumes paramount importance.

This paper introduces a novel procedural system for the automated generation of highway models. Our system is inspired by a set of requirements that motivated us to develop our own procedural modeling:

- Integration of real-world georeferenced data for enhanced realism.

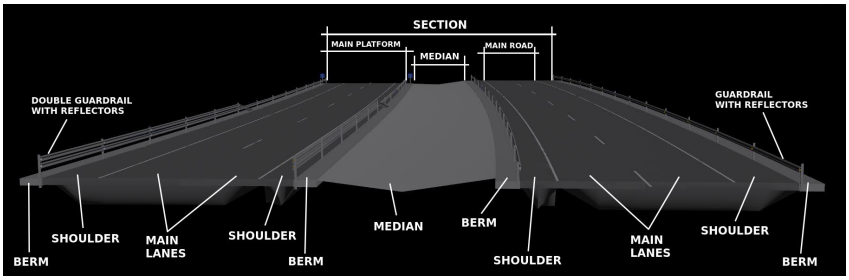


Figure 1: Principal 3-D geometric components comprising a standard highway road.

- Comprehensive modeling of highway environments, including safety and signaling elements, motivated by civil engineering construction standards.
- Achieving a high level of realism in a cost-effective manner, leveraging specialized software in procedural modeling (Houdini, 1996–2023).
- Wide applicability across various domains, from civil engineering to driving simulations and digital twin generation.

2 System overview

In this section, we provide an overview of our proposed system. Figure 2 illustrates the system's pipeline, highlighting the essential input data requirements and the subsequent generation of highway geometry. This systematic approach adheres to civil engineering standards and real-world specifications, ensuring accuracy in modeling and analysis within various applications. As input data, two primary components are required. Firstly, the highway layout, represented as a sequence of geographic coordinates defining the road's shape, is essential. Secondly, a set of annotations pinpointing the locations of signals, panels, barriers, and other pertinent elements is also necessary. The generation of highway geometry begins with the creation of road platforms, encompassing lanes and shoulders, as well as beacons, gantries, shoulders, horizontal and vertical signage, safety barriers, and the central median. The outcome of this process yields a complete geometric representation of the highway segment provided as input to the system. This systematic approach ensures the faithful recreation of highway geometry, adhering to civil engineering standards and real-world specifications, thus enabling accurate modeling and analysis within the intended applications.

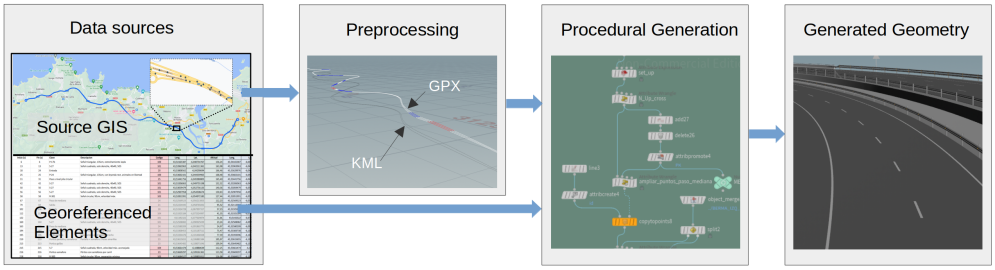


Figure 2: Our procedural modeling pipeline. Here, we represent a 23 km segment of Spain's A-8 highway, comprising 508 and 502 georeferenced points for one and the opposite directions respectively.

2.1 Data sources

The highway layout is represented as a sequence of georeferenced points, providing a precise definition of the road's configuration. As shown in Figure 2, we can describe a segment of a highway by using multiple georeferenced points to define the highway's shape and characteristics. Our system expects the highway layout data in the form of a KML file, as delivered e.g. from Google Maps. Thus, it could be automatically retrieved through third-party applications or services, simplifying the data acquisition process.

The input georeferenced elements involve in our case the annotation of georeferenced localizations of various relevant items present in a highway like e.g. signage, beacons, information panels, safety elements, and more. As one source to obtain this information, we utilized video footage captured with a GoPro camera installed in the cockpit inside a car vehicle. This annotation process is manual and entails marking the point of interest along with the second in the video when it appears and also when it disappears from view. Since the GoPro-recorded video incorporates GPS information, it allows for the automatic conversion of a timestamp into a geographical location. Annotating elements based on time intervals might appear coarse at first, especially given that a vehicle can cover a considerable distance in a single second. However, we have not identified the need for more precise annotations. The finalized annotations have been exported into a CSV file. In the case of the 23 km route shown in Figure 2, a total of 340 elements were manually annotated, requiring approximately three hours to complete this task.

2.2 Preprocessing

Since the obtained KML file lacks elevation information, we leveraged GPS Visualizer to procure elevation data for each of the georeferenced points. This versatile tool seamlessly converts the KML file into a GPX file, incorporating both positional and elevation data throughout the entire highway layout. Then, the GPX file is transferred to Houdini for the procedural generation of the highway, as illustrated in Figure 2.

2.3 Modeling

The modeling process initiates by generating individual polylines, one for each direction of the roadway, and based on the GPX file provided as input. Then, a third polyline representing the median is generated between these two (refer to Figure 3). At each specific point along the polylines, we assign the relevant attributes using information from the manually annotated elements found in the provided CSV file. This comprehensive dataset serves as the foundation for constructing the roadway model, which includes all the elements listed below (refer to Figure 4):

a) Roadway Platforms. Undoubtedly, this is the most crucial element of a roadway, serving as the foundation for lanes and shoulders.

b) Gantry. A gantry is a structure consisting of a pillar at each end, supporting vertical signaling elements above the road. These elements include information panels, traffic lights, dynamic variable message signs, etc. In this project, only the type of gantry that supports information panels has been implemented.

c) Shoulders. The shoulders are the areas situated between the outer edge of the shoulder and the gutter. The shoulders have been automatically generated without requiring manual annotation.

d) Horizontal Signage. This category encompasses all the road markings, such as continuous lines separating the shoulders, dashed lines between lanes, or between a lane and the on-ramp or off-ramp, directional arrows, etc. This signage is automatically generated without the need for manual annotation.

e) Vertical Signage. Vertical signage includes signs such as speed limits, lane merge warnings, entry notifications, etc.

f) Safety Barriers. Simple and double safety barriers have been included. Each side of the road is treated independently.

g) Median. The median is an area constructed between the inner edges of the shoulders on both sides of the road. Its width is not fixed and depends on the separation between the shoulders. Furthermore, it features a slope toward the center to prevent rainwater from flowing onto the roadway.

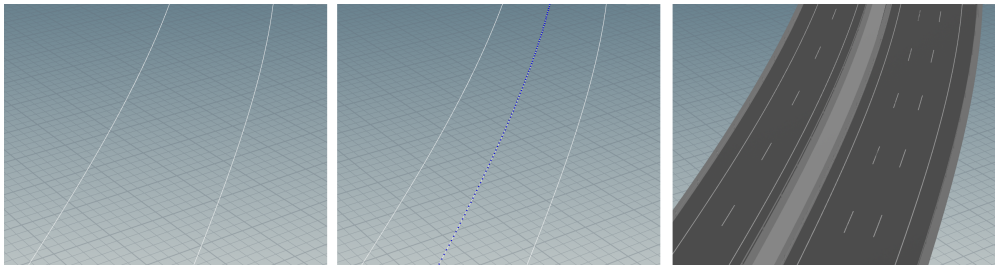
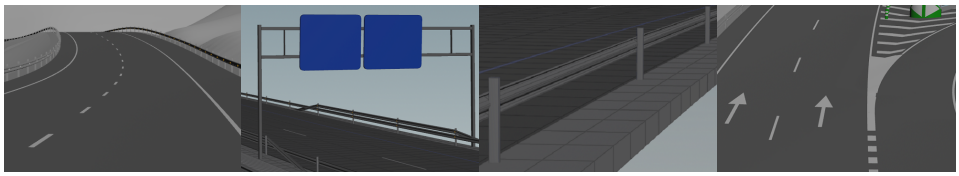
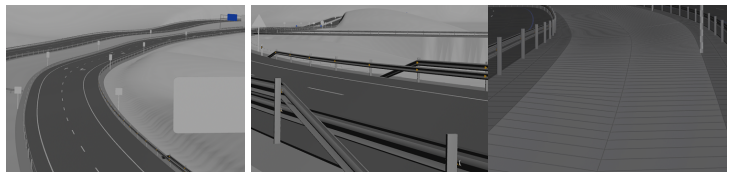


Figure 3: Polylines of both directions (left), plus the median (center), and road geometry (right)



(a) Roadway Plataforms (b) Gantry (c) Shoulders (d) Horizontal Signage



(e) Vertical Signage (f) Safety Barriers (g) Median

Figure 4: Elements included for the construction of the roadway model

3 Results

Figure 5 shows some results of the geometry that the procedural system generated from the data provided as input. We had to limit the length of the track to 3 km since otherwise Houdini would generate an amount of geometry too large, over half a million vertices, for the memory capacity of the computer, 8GB in our case. Regarding the processor, an Intel i5 processor was used, without an external graphics adapter, and there was no problem interacting with the scene.

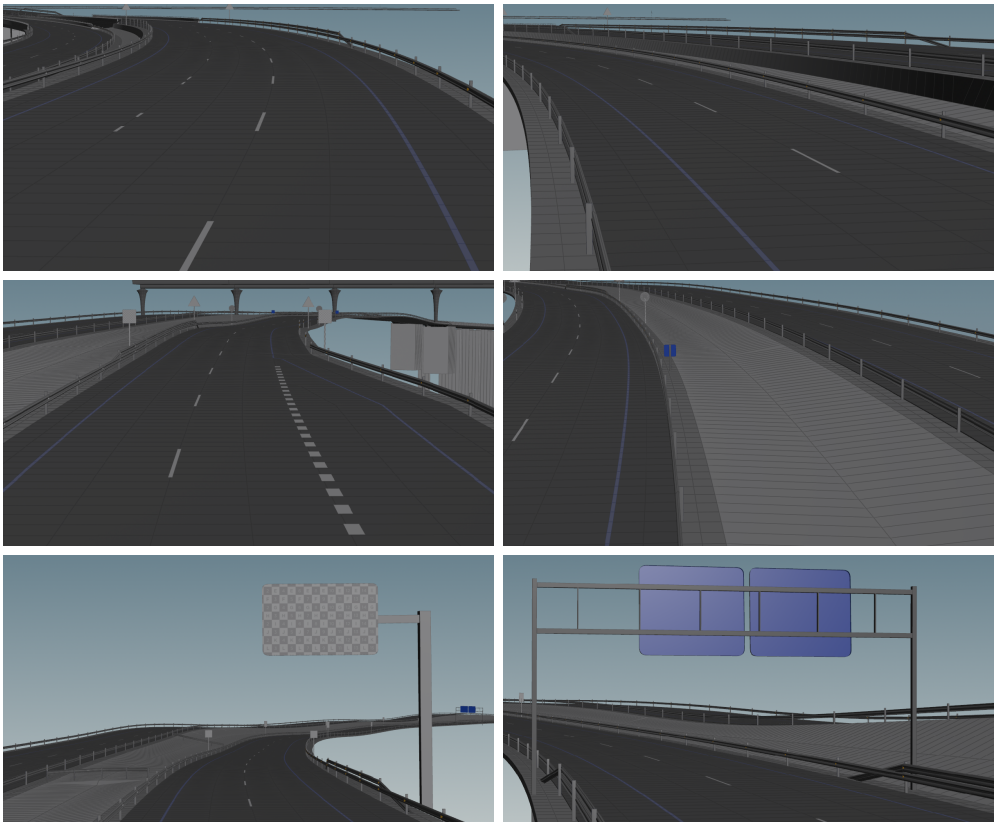


Figure 5: Results achieved using our proposed procedural system

A visual comparison between real images captured with a GoPro camera installed in the cockpit inside a car vehicle (left column) and views of the geometry generated using our procedural system from a very similar point of view (right column) is provided in Figure 6.

Figure 6: Real images captured with a GoPro camera (left column) and views obtained using our procedural system from a very similar point of view (right column)

Finally, Figure 7 shows the some cut of the highway where we have changed some parameters. For instance, images on the top row show two lanes, while images on the bottom row show three lanes. Furthermore, images on the top row show changes on the safety barriers, while images on the bottom row show changes in the width of the lanes.

4 Conclusions and future work

This article presents a procedural system that allows generating the geometry of a typical highway of the real world. The reference data (basic layout) of the road used as a case study has been extracted from Google Maps, while the signaling elements have been annotated by hand from videos obtained with a GoPro camera installed in a vehicle. The proposed system successfully models all the geometry corresponding to the platforms, signals and safety elements. As future work, it is expected to include in the model structures such as viaducts, level crossings and tunnels, as well as entrances (merges and on-ramps) and exits (diverges and off-ramps) of the road, so that all the traditional elements of a highway are characterized. Additionally,

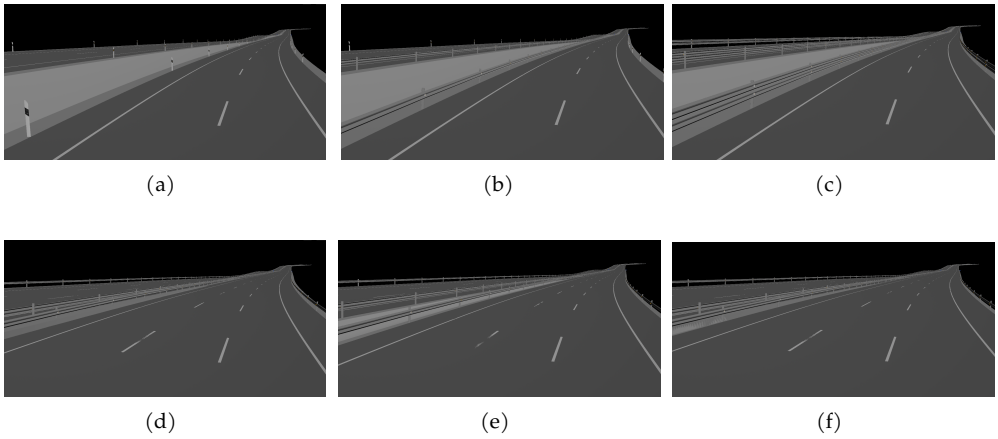


Figure 7: Examples of parameter changes: (a) no safety barriers, (b) simple safety barrier on the left side, (c) double safety barrier on both sides, (d) 3.5 meters wide lanes, (e) 3 meters wide middle and left lanes, (f) 4 meters wide right lane

work is being done to increase the degree of realism of the images and the aim is to incorporate terrain and vegetation models into the geometric model. The ultimate goal would be to provide the generation with the greatest possible degree of realism and at the same time versatile results.

Acknowledgments

This work has been supported by the Spanish Ministry of Science and Innovation (AEI/PID2020-115734RB-C22). The whole team also wants to acknowledge the support provided by Side Effects Software Inc. in developing our work. Jose Ribelles was supported by Generalitat Valenciana (CIBEST/2022/139). Javier Taibo and J.A. Iglesias-Guitian were supported by Xunta de Galicia (ED431F 2021/11). Additionally, J.A. Iglesias-Guitian also acknowledges the UDC-Inditex InTalent programme and the Ministry of Science and Innovation (AEI/RYC2018-025385-I).

Bibliography

D. S. Ebert, F. K. Musgrave, D. Peachey, K. Perlin, and S. Worley. *Texturing and Modeling: A Procedural Approach*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 3rd edition, 2002. ISBN 1558608486.

C. Gasch, J. M. Sotoca, M. Chover, I. Remolar, and C. Rebollo. Procedural modeling of plant ecosystems maximizing vegetation cover. *Multimedia Tools and Applications*, 81(12):16195–16217, 2022.

S. Houdini. 3d procedural software for film, tv and gamedev. https://www.sidefx.com/products/houdini/, last viewed July 2023, 1996–2023.

P. Müller, P. Wonka, S. Haegler, A. Ulmer, and L. Van Gool. Procedural modeling of buildings. In *ACM SIGGRAPH 2006 Papers*, pages 614–623. ACM Inc., 2006.

Y. I. Parish and P. Müller. Procedural modeling of cities. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 301–308, 2001.

M. Schwarz and P. Müller. Advanced procedural modeling of architecture. *ACM Transactions on Graphics (TOG)*, 34(4):1–12, 2015.

S. T. Teoh. Riverland: An efficient procedural modeling system for creating realistic-looking terrains. In *International Symposium on Visual Computing*, pages 468–479. Springer, 2009.

Automation Proposal for the Intermediate Steps in the 16S FFPE Samples Analysis Pipeline

Elsa Martin-De Arribas, Kelly Conde-Pérez, Pablo Aja-Macaya, Juan A. Vallejo, Margarita Poza, and Susana Ladra

Universidade da Coruña, Centro de Investigación CITIC, 15071 A Coruña, Spain
meiGAbiome, INIBIC-CICA - Universidade da Coruña - CIBERINFEC-ISCIII. Hospital Universitario, A Coruña, Spain.
Correspondence: elsa.mdearribas@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract:
In the day-to-day work of bioinformatics, the use of integrated software packages, which encompass a wide range of tools, enables the development of pipelines for omics data analysis. Within the various existing pipelines, we focus on the analysis of the 16S rRNA gene as it allows for the study of diversity and taxonomy of prokaryotic microorganisms such as Bacteria and Archaea. However, these pipelines often involve a sequence of multiple tools that require intermediate steps before further processing can proceed, as in the case between Cutadapt and DADA2. In fact, in a typical pipeline, the values for DADA2 input arguments 'trunc-len-f' and 'trunc-len-r' are extracted from the output of Cutadapt. The best approach for selecting optimal values (aka the trimming positions) is graphically visualizing Cutadapt output and manually selecting the most accurate trimming position length. Therefore, we propose the automation of this specific intermediate step between Cutadapt and DADA2 tools, by selecting values displayed in the graphs that meet the filtering criteria. This automation has been incorporated into a custom pipeline for the analysis of the microbiome in 16S paired-end samples from colorectal cancer patients, and could potentially serve as a standardization approach in these processes.

1 Introduction

In the realm of biological sciences, the study of microbial communities persists as a crucial element, with relevant spanning diverse fields, such as agriculture, marine environment, and medicine. While any microbiome-related field of study is of interest, our specific focus lies within the biomedical sphere, particularly in cancer research. This paper, in particular, centres on the examination of formalin-fixed paraffin-embedded (FFPE) microbiome samples from colorectal cancer (CRC) patients (Conde-Pérez et al., 2023). Sequencing of these samples was conducted using the 16S Amplicon Metagenomic Sequencing technique, a next-generation, high-throughput methodology that emphasizes amplifying specific, short, targeted regions known as amplicons. This methodology is a standard in molecular laboratories, supported by several dedicated bioinformatics tools and platforms for the analysis of 16S rRNA gene amplicon data (Straub et al., 2020).

One of the best-known and widely used examples is the QIIME2 platform (Bolyen et al., 2019), which stands for Quantitative Insights Into Microbial Ecology. Within this platform, numerous integrated tools are essential for such studies. In our case, we developed a specific pipeline for the analysis of these FFPE samples, adding a new qiime2-based implementation

called Sidle (Debelius et al., 2021). Sidle was a Python version developed within the QIIME2 environment, based on a pre-existing algorithm originally created by Fuks et al. in 2018 and called Short MUltiple Reads Framework (SMURF) algorithm (Fuks et al., 2018). The original purpose was to enable the reconstruction of multiple short, fragmented amplicons against a known database, but Sidle adds a novel tree-building solution. This enhancement significantly improves the resolution of the reconstructed community compared to single amplicons.

Inside the Sidle pipeline, two pivotal tools play an important role in the analysis: Cutadapt and Divisive Amplicon Denoising Algorithm (DADA2). As described in its paper, Cutadapt finds and removes unwanted sequences from high-throughput sequencing reads, such as adapters, primers, poly-A tails, and more (Martin, 2011). Subsequently, DADA2 (Callahan et al., 2016) takes over, returning the reads that have been retained after quality filtering and merging steps. However, transitioning from one tool to another involves the manual selection of the positions of the forward and reverse sequences that will be trimmed by DADA2 for subsequent read processing. That manual selection requires incrementing the time of selecting and testing the best optimal positions to pass the filtering, merging, and removing chimeras sequences in the sample reads. For that reason, we propose a novel automatic selection and testing of optimal positions in the DADA2 tool.

2 Background

2.1 16S Amplicon Metagenomic Sequencing

Over the last decade, the ribosomal 16S rRNA gene has been established as a phylogenetic marker due to its ubiquitous presence among members of both Bacteria and Archaea domains. Its structural composition is characterized by a combination of "conserved" and "variable" regions: the conserved regions are universally shared across all microorganisms, while the variable regions serve as distinctive "tags", facilitating the precise taxonomic or phylogenetic classification of each microorganism. Consequently, universal primers are designed based on the adjacent conserved regions to the targeted variable regions, enabling differentiating the taxa present (Straub et al., 2020).

Hence, one of the best-known approaches in metagenomic studies is sequencing the 16s rRNA gene or the 18S rRNA gene. However, since the 18S rRNA gene is present in eukaryotic organisms and fungi, it will not be considered in this study. Although there are various types of next-generation sequencing, we will focus on amplicon sequencing. Amplicons are small fragments of DNA or RNA obtained through Polymerase Chain Reaction (PCR) or other techniques that yield multiple copies of these fragments. The PCR amplicon-based sequencing has been employed in multiple biological fields, from agriculture to tumour immunology (Straub et al., 2020).

In our study, we employ a set of five primers specifically designed for amplicon sequencing. These primers play a crucial role in PCR, where they target and amplify the 16S rRNA gene from the bacterial DNA in our samples. It generates paired-end sequences, which provide valuable information about the amplicon fragments at both ends. These fragments are distinguished thanks to our carefully selected primer sets, which enable us to differentiate and analyse the sequences generated from the forward and reverse strands of the target gene regions. This approach enhances our ability to study microbial communities with greater accuracy and depth.

2.2 Colorectal cancer, microbiome and FFPE samples

Colorectal cancer (CRC) ranks as the second leading cause of cancer-related death globally (Wang et al., 2012; Wong et al., 2019), with the highest incidence in Spain in 2023 with 42,721 cases, followed by breast cancer with 35,001 cases, and lung cancer with 31,282 cases (Sociedad Española Oncológica Médica (SEOM), 2023). Traditionally, CRC treatments include surgery and chemotherapy (sometimes combined with radiotherapy). However, recent developments

focus on gut microbiome-targeted therapies due to the gut microbiota's role in CRC (Conde-Pérez et al., 2023; Wang et al., 2012; Wong et al., 2019).

The gut microbiota, residing in the gastrointestinal tract, influences energy metabolism, immunity, and carcinogenesis responses. Through functional studies, alterations in microbial composition and ecology have been revealed in individuals with CRC, pinpointing the roles of several bacteria in colorectal carcinogenesis, such as *Fusobacterium nucleatum*, certain *Escherichia coli* strains, and *Bacteroides fragilis* (Wang et al., 2012; Wong et al., 2019). Additionally, a recent study has proposed the translocation of particular bacteria, *Parvimonas micra* from the subgingival cavity to the gut (Conde-Pérez et al., 2023). This discovery is groundbreaking, as the data suggest that this bacteria may migrate as a part of a synergistic consortium with other periodontal bacteria.

Our research delves into the bacterial microbiota of CRC patients using the 16S amplicon sequencing, deepening our CRC understanding. The microbiota is extracted from the formalin-fixed paraffin-embedded (FFPE) samples, which entails preserving and processing colon or rectum tissue sections. FFPE's preservation protocol, while precise, affects DNA integrity during long-term storage, prompting the need for improved preservation methods (Guyard et al., 2017). These samples play a pivotal role in CRC diagnosis, characterization, and research through microbiota analysis, further advancing our CRC insights (Conde-Pérez et al., 2023).

2.3 Workflow

The principal aim of a bioinformatic pipeline is maximizing the reliability and confidence in community data analysis, which becomes particularly challenging when dealing with microbial communities. Among the studies comparing the existing technologies applied in bioinformatic pipelines, Straub's study assessed a popular list of analysis pipelines, i.e., Mothur, QIIME (version 1 and 2), and MEGAN against 24 datasets (Straub et al., 2020). QIIME2 was identified as the most accurate option due to its denoising capabilities (DADA2 and Deblur), surpassing OTU clustering methods used by the other pipelines. Despite the extra computational time required by DADA2, it effectively retains sequences below specific length thresholds and captures unique site-specific sequences, highlighting the critical importance of precise pipeline selection in microbiome research.

Taking all of these criteria into consideration, the FFPE samples pipeline was developed following the recommendations of the Sidle author (Debelius et al., 2021). In a nutshell, our pipeline workflow begins by importing FASTQ format samples into the QIIME2 environment. Given the unique characteristics of FFPE samples, where our objective is to detect the maximum number of microorganisms, we define five variable regions. Next, we demultiplex the samples into the five defined regions (in our case) using the Cutadapt tool, which leverages primers to differentiate and remove them from the sequences. Following this, manual trimming positions are selected based on Cutadapt's output for each region, which can be computationally intensive and time-consuming. These positions, defining parameters such as 'trunc-len-f' and 'trunc-len-r' for right-side trimming and 'trim-left-f' and 'trim-left-r' for left-side trimming, are then used in DADA2. Subsequently, DADA2 is executed, producing three output files for each region. One of these files contains statistics extracted for each sample during the tool's step, resulting in a final percentage of reads that have passed all selection criteria. The final steps involve aligning the region in order to reconstruct a portion of the original 16S rRNA targeted gene and detect the microorganisms present in the samples (Conde-Pérez et al., 2023).

2.4 Related studies

Upon conducting a comprehensive review of the literature related to this topic, a new tool called FIGARO was identified in the form of a preprint (Weinstein et al., 2019). Its primary objective is to optimize cut-off positions for maximizing read retention post-filtering, resembling DADA2 in approach. However, FIGARO's GitHub repository has not seen updates since its initial re-

lease, potentially causing installation challenges. Additionally, we found the nf-core/ampliseq framework (Ewels, *et al.*, 2020; Straub et al., 2020), a bioinformatics analysis pipeline. It supports denoising across various amplicons (16S, ITS, CO1, 18S), allows taxonomic assignment, and works with both paired-end and single-end data from multiple sequencing platforms. Notably, no instances of analysis with FFPE samples were found.

In a brief comparative benchmarking of these two options and the approach presented in this paper, it becomes evident that FIGARO (Weinstein et al., 2019) is not a readily applicable tool within the QIIME2 environment. Conversely, the rationale behind their optimal positions of trimming selection approach is well-founded. On the other hand, the nf-core offers the flexibility to develop a pipeline tailored to a specific sample type using Nextflow or one of the multiple pipelines shared by its contributors. However, it is important to note that no existing pipelines were identified that refer to the sample type utilized in our study, and proficiency in Nextflow is a prerequisite for its utilization (Ewels, *et al.*, 2020; Straub et al., 2020).

3 Proposal

The previous section emphasizes the underlying issue with FFPE samples due to DNA degradation over time due to the tissue preservation techniques employed (Guyard et al., 2017). To maximize the usability of reads from FFPE samples, this study focuses on the DADA2 tool. Due to the complexity of the samples, the study was designed using five different sets of primers to amplify five regions of the 16S rRNA gene, which would later be reconstructed thanks to QIIME2-implementation Sidle. The pipeline used the QIIME2 ecosystem platform installed in a Conda environment, where all the required tools were available. Firstly, paired-end samples in FASTQ format were imported to QIIME2 and were separated into the five regions by Cutadapt, thanks to the five sets of primers. From here on, the general process will be counted, but it was carried out in each of the regions. Cutadapt usually returns a qza extension file output, which would be the input file for other tools, but, for dynamically visualizing the results and data, QIIME2 provides a qzv extension file. These types of files are designed to be opened in the web-based interactive viewer called QIIME2 View, which offers a user-friendly manner of exploring data.

Henceforth, assuming that the qzv file is actually a compressed file consisting of a set of files containing the graphics and data to be displayed, files containing the data for forward and reverse directions were analysed. These two files showed the same structure: a data frame with eight columns and rows as length sequence, representing the data in two interactive graphical plots with quality values as the y-axis and sequence length as the x-axis. Various Python scripts were created to analyse the data and extract the most optimal positions for forward and reverse directions. Some filter criteria were established during the data exploration (each direction was evaluated independently):

1. Eliminate low-quality sequences by multiplying the maximum count by 0.2.
2. Select a minimum starting length, typically set at 100.
3. Get statistical analysis of each quality columns (2%, 9%, 25%, 50%, 75%, 91% and, 98 %) and count column.
4. Transform quality values within columns to a scale of 0, 1, or -1.
5. Detect local maxima points within columns 50% and 75% using the Python library scipy.signal and the find_peaks function.
6. Test optimal positions for DADA2 input parameters, ensuring they are greater than the minimum length requirement (here 100).
7. Set left-side values to zero based on previous results.
8. Ensure optimal positions satisfy DADA2's overlap formula.

Figure 1: Forward strand graph for 16S rRNA gene amplified R1 region

Upon completing the testing of optimal positions for both directions and obtaining three DADA2 output files, we analyse the denoising_stats output file. This file provides denoising statistics, detailing the entire DADA2 denoising process, including total input reads, passing reads after quality filtering, merged sequences, and the number of non-chimeric reads retained. The aim is to retain as many non-chimeric reads as possible for each sample, although the specific target may vary depending on the study's goal and characteristics.

4 Preliminary Results

This study proposes an ad hoc approach to automate an intermediate step between Cutadapt and DADA2 tools, detecting changes along the quality values of base positions in both read directions. Traditionally, manual selection of optimal positions requires careful examination of data to identify fluctuations in quality levels. The proposal detects small and large changes and indicates the most optimal positions for trimming in the right section of the graph, ensuring they are higher than position 100, as in Sidle, post-dada2-trimming is performed before alignment. This makes it easier to perform the step of alignment and reconstruction of the final sequence of the 16S rRNA gene, formed by the regions under study (more than 60% of the region is reached in the reconstruction).

In a brief example of how it works, for the R1 region, manually selected position values for the forward and reverse strands were 114 and 115, respectively. However, the algorithm suggested positions of [114, 122, 130] for the forward strand and [112, 126] for the reverse strand. Test results indicated that the optimal trimming positions were 130 for the 'trunc-len-f' argument (right side of the forward strand) and 126 for the 'trunc-len-r' argument (right side of the reverse strand). As shown in Figure 1, the blue cross represents all potential optimal positions, while the highlighted purple cross signifies the final optimal values for the forward direction, with the x-axis representing the index position and the y-axis indicating detected fluctuations. In this pipeline, the 'trim-left-f' and 'trim-left-r' parameters were set to zero for both strands. It is evident that the manually selected positions were overly conservative. This process was applied to five regions, and the automated selection proved more accurate, except for the R2 and R4 regions, where there were closely spaced fluctuations.

5 Discussion

The study of the microbiome is still a pending subject. Although a great deal of progress has been made in a short time that it has really been booming, it is still necessary to standardize the processes. In this case, the work focused on a complicated type of sample, since due to the treatment it receives before sequencing, the starting DNA is highly degraded. However, these types of samples can now be analysed thanks to the SMURF tool implementation within QIIME2, called Sidle. Despite this, throughout the workflow defined in Sidle, a great loss of information has been observed simply because the values of the trimming parameters in DADA2 have been incorrectly defined.

In a promising development, this limited-scale study introduces an initial ad hoc approach to streamline an intermediate step between the Cutadapt and DADA2 tools. Our primary goal is to efficiently identify quality changes in both read directions, going beyond traditional manual selection by detecting subtle and significant changes. The approach determines the most suitable trimming positions within the right section of the quality profile graph, with a criterion set beyond position 100. This choice aligns with the Sidle workflow, where a pre-alignment step called post-dada2-trimming is applied, splitting sequences at position 100 across all samples from all regions. This approach streamlines subsequent alignment and reconstruction steps for the final 16S rRNA gene sequence, covering over 60% of the targeted gene during reconstruction.

There remains a substantial amount of work ahead, considering that the results presented are in their preliminary stages. To ensure the robustness and applicability of our approach, several key tasks are on our agenda. Firstly, we will evaluate the tool's performance on datasets containing FFPE samples, as well as normal datasets, i.e. V3V4 samples. Secondly, we have future plans to expand our testing to encompass 18S and ITS samples, though this lies further down the road. Additionally, we aim to benchmark the optimization outcomes against other established tools in the field. Furthermore, we will explore the utility of our approach in the context of the Deblur tool, which is akin to DADA2 but selected based on specific process requirements and sample types. Lastly, our strategy involves the meticulous analysis of DADA2's verbose output during runtime, enabling us to proactively halt the process if anomalies such as sample removal in filtering steps emerge, ensuring the integrity of the workflow.

6 Conclusions

Microbiota's role in diseases like autoimmunity, cancer, neurological, and renal diseases has been studied over the past two decades. Tools like Sidle, integrated into QIIME2, allow for a pre-established pipeline for analysing FFPE paired-end sample sequences. The pipeline involves other tools as Cutadapt for primer removal and DADA2 for denoising, with parameters 'trunc-len-f' and 'trunc-len-r' determining read truncation. We suggest a new automated approach, based on the data output (normally visualized graphically to choose the positions) and following the quality-filtering indications from Cutadapt. In conclusion, this paper presents preliminary findings, and our ongoing work aims to optimize this procedure, incorporate enhancements, and potentially develop machine learning or deep learning models for full automation, with plans for benchmarking against other tools.

Funding

The development of the project is supported in part by CITIC as a Research Center of the Galician University System and is financed by the Consellería de Cultura, Educación e Universidade de la Xunta de Galicia through the ERDF (80%), Operational Program ERDF Galicia 2014–2020, and the remaining 20% by the Secretaría Xeral de Universidades (Grant. ED431G 2019/01), as well the Xunta de Galicia/ERDF-EU [ED431C 2021/53], and the Recovery, Transformation, and

Resilience Plan (PRTRC17.I1), The present study also received funding from the Instituto de Salud Carlos III (ISCIII, Spain) through the project PI20/00413 and co-funded by the European Union (EU) to M.P., while K. C-P. was granted by Asociación Española Contra el Cáncer (AECC).

Acknowledgments

The authors of this paper extend their sincere appreciation to the collaborative efforts and contributions of the meiGAbiome Group, as well as the entire team of medical and anatomopathologists. Finally, we are deeply grateful to the patients whose selfless donations have made this and numerous other studies possible.

Bibliography

E. Bolyen et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology*, 37(8):852–857, 2019.

B. Callahan et al. DADA2: High-resolution sample inference from illumina amplicon data. *Nature Methods*, 13(7):581–583, 2016.

K. Conde-Pérez et al. Parvimonas micra can translocate from the subgingival sulcus of the human oral cavity to colorectal adenocarcinoma. *Molecular Oncology*, 2023.

J. Debelius et al. A comparison of approaches to scaffolding multiple regions along the 16S rRNA gene for improved resolution, 2021.

Ewels, *et al.* The nf-core framework for community-curated bioinformatics pipelines. *Nature Biotechnology*, 2020.

G. Fuks et al. Combining 16S rRNA gene variable regions enables high-resolution microbial community profiling. *Microbiome*, 6(1), 2018.

A. Guyard et al. DNA degrades during storage in formalin-fixed and paraffin-embedded tissue blocks. *Virchows Archiv*, 471(4):491–500, 2017.

M. Martin. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, 17(1):10–12, 2011.

Sociedad Española Oncológica Médica (SEOM). Las cifras del cáncer en españa 2023. *https://seom.org/images/Las_cifras_del_Cancer_en_Espana_2023.pdf*, 2023.

D. Straub et al. Interpretations of environmental microbial community studies are biased by the selected 16s rrna (gene) amplicon sequencing pipeline. *Frontiers in Microbiology*, 11, 2020.

T. T. Wang et al. Structural segregation of gut microbiota between colorectal cancer patients and healthy volunteers. *The ISME Journal*, 6(2):320–329, 2012.

M. Weinstein et al. Figaro: An efficient and objective tool for optimizing microbiome rrna gene trimming parameters, 2019.

S. Wong et al. Gut microbiota in colorectal cancer: mechanisms of action and clinical applications. *Nature Reviews Gastroenterology & Hepatology*, 16(11):690–704, 2019.

Prototype of an Entity Recognition System for Antimicrobial Resistance Data Management.

Francisco Prado-Valiño, Roi Santos-Ríos, Carlos Gómez-Rodríguez, and Jesús Vilares

Universidade da Coruña, Centro de Investigación CITIC - Grupo LYS, Facultad de Informática, Campus de Elviña, 15071 - A Coruña (España)
Correspondence: jesus.vilares@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: Often, a study or research process requires the analysis of large volumes of information in the form of unstructured text. This task consumes a large amount of time and resources of the human experts in charge of it. For this reason, there is great interest in developing automatic systems to support these activities by applying Text Mining techniques. A key task in this type of process is Entity Recognition: extracting the entities of interest contained in a text and classifying them into pre-established categories.
The work presented here is part of the GRALENIA project, which seeks to improve antimicrobial resistance data management in the hospital setting. The main goal of our work is the development of an entity recognizer prototype for the identification and labeling of indicators of interest present in clinical reports.

1 Introduction

In recent years, thanks to the progressive digital conversion of our health systems, the Biomedical field has experienced a significant increase in the amount of information available for research and analysis. In this context, electronic health records (EHR) constitute a major information source, since they contain a large amount and variety of information about our health (Fleuren and Alkema, 2015). However, although part of these data may be structured (laboratory tests, prescriptions, etc.), many other is formed by unstructured text (nursing notes, admission reports, etc.), making it difficult to analyze said data.

The study of the huge amount of information contained in these unstructured sources involves a great consumption of time and resources of trained personnel. Due to the complexity and cost of these analyses, there is great interest in developing support systems capable of automating these tasks as much as possible.

2 Main Goals

Our work consists of the application of Text Mining (TM) (Hotho et al., 2005) techniques for Entity Recognition (ER) in order to automate the detection of entities on unstructured clinical text. In our case, we will focus on the field of antimicrobial resistance (AMR), specifically on the resistance of bacteria to antibiotics. This work is part of the GRALENIA project,[1] which seeks to improve AMR data management in the hospital setting, specifically focusing on three bacteria: *MRSA*, *Klebsiella ESBL* and *Klebsiella carbapenemase*. For this purpose, the project proposes to

[1] *http://gralenia.es* (visited on Sept. 2023).

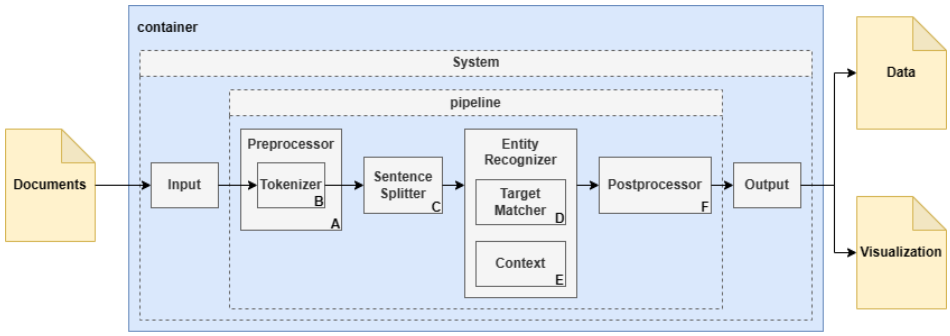


Figure 1: General architecture of the ER system: rule-based approaches.

create a digital platform that, using Natural Language Processing (NLP) techniques, analyzes patients' clinical reports in search of AMR-related indicators.

These indicators belong to one of the categories pre-defined by our experts: BACTEREMIA, CENTRAL NERVOUS SYSTEM (CNS) INFECTION, FEVER, GASTROENTERITIS, GENITAL INFECTION, ORAL INFECTION, PHLEBITIS, RESPIRATORY INFECTION, SEPSIS, SURGICAL WOUND INFECTION, and URINARY INFECTION. Within the framework of the parent project, and in a later phase, the detected indicators will serve as input features to an AI system for the detection of resistances and possible sources of bacterial contagion.

In this work, we intend to develop a prototype ER system to identify possible AMR indicators (named *symptomatic expressions*) in patients' clinical reports.

3 Development

At the time of writing this article, a suitable *dataset* for analysis and system evaluation in Spanish was not yet available. As a temporary solution, we decided to work with the MIMIC-III (Johnson et al., 2016) English database in the meantime, since the nature of its content adapts to the needs of our project. This database contains anonymized information on medical data belonging to more than 40,000 patients at *Beth Israel Deaconess Medical Center* (*BIDMC*).

SpaCy[2] open-source library has been used for building the core of our Natural Language Processing (NLP) system. SpaCy provides us with a framework in the form of a modular pipeline, this being the architecture that we will use for our ER system. As shown in Figure 1, it consists of the following modules:

1. **Tokenizer:** Segments the text into tokens. It also generates the *Doc* object on which the other SpaCy components will work.
2. **Preprocessor:** Preprocesses the tokenized text prior to creating the *Doc*.
3. **Sentence Splitter:** Segments the text into sentences, marking them in the *Doc*.
4. **Entity Recognizer:** Performs the ER task, properly speaking, on the *Doc*.
5. **Postprocessor:** Postprocesses the *Doc*.

3.1 First approach: Baseline rule-based system

Instead of generating the analysis rules of the system manually, we have automated their creation through code. In this way, it is enough to add, delete or modify any of the symptomatic expressions and the system will generate the new updated rules automatically. In addition,

[2] *http://spacy.io* (visited on Sept. 2023)

this allows us to easily add new categories by simply adding new sets. These rules are applied through a pattern matching process, labeling the detected category on the text, as can be seen in Example 1.

The following sentence, extracted from a MIMIC-III document, presents several indicators that we are interested in identifying and labeling. Specifically, it contains the symptomatic expressions *"abdominal pain"*, *"vomit"* and *"diarrhea"*, belonging to the category GASTROENTERITIS, along with *"dysuria"*, belonging to URINARY INFECTION. Once we apply the rules on this phrase, we obtain the following output:

"Denies specific complaints when asked, including chest pain, cough, ***abdominal pain GASTRO-*** ***ENTERITIS*** *, nausea and vomits,* ***diarrhea GASTROENTERITIS*** *, and* ***dysuria URINARY_*** ***INFECTION*** *."*

(Example 1)

Preprocessing

To facilitate establishing correspondences, those abbreviations and acronyms present in the input text are replaced by their components: *"rx'd"* for *"prescribed"* or *"tº"* for *"temperature"*, for instance. This process is carried out during the *Preprocessing* phase, before applying the rules.

Results

The results for this first approach (*base*) are shown in Table 1, which details the number of labeled entities (i.e. that matched some pattern), the number of documents with at least one label, and the percentage that these documents represent over the total. Total coverage was 41.01%, so there is still ample room for improvement.

3.2 Second approach: Fuzzy matching and term expansion

One of the main problems with our initial approach is the negative impact of spelling errors, since they prevent matching: *"fevre"* instead of *"fever"* or *"taychardia"* instead of *"tachycardia"*, for example. To reduce this problem as much as possible, we will introduce *fuzzy matching* mechanisms in the rules in order to make the patterns more flexible. Specifically, we will use the *Levenshtein editing distance* (Levenshtein et al., 1966) on the words contained in a symptomatic expression. Given two words, the Levenshtein edit distance is the minimum number of single-character edits (insertions, deletions, and substitutions) required to convert one word into the other.

Similarly, another of the main problems detected is the use of morphological variants (e.g. changing from singular to plural) and semantic variants (e.g. synonyms) of the original symptomatic expressions. As a solution, inflectional morphemes accounting for number variation (e.g. *"expectoration"*-*"expectorations"*) and verbal forms (e.g. *"vomit"*-*"vomiting"*) have been integrated. Synonyms commonly used in the application domain have also been added (e.g. *"secretion"*-*"discharge"*) after being obtained from the OpenMD online medical dictionary.[3]

Rules for complex symptomatic expressions

As explained before, our automatically generated rules provide flexibility when it comes to expanding the set of symptomatic expressions to be detected by the system but, at the same time, they are currently limited in terms of their complexity. Some of these expressions, such as measurements, present more complex structures than those initially proposed by experts and, then, automatically encoded into the rules. To define rules capable of correctly identifying

[3] *http://openmd.com/dictionary/* (visited on Sept. 2023).

and labeling this kind of complex structures, we have chosen to manually refine the patterns involved. Example 2 includes one of these complex expressions.

> In order to label expressions related to blood pressure along with their associated numerical value, we defined *ad hoc* rules with more elaborate patterns. These were based on a list of dictionaries containing (*characteristic*, *value*) pairs. Once these new rules are applied, the following output is obtained:
>
> *"His* ***blood pressure was 223/125 SEPSIS*** *, otherwise Hemodialysis stable."*
>
> (Example 2)

Postprocessing

By allowing matches with words within edit distance 1 of those considered within a rule, it is inevitable that expressions containing correct terms that are very similar to those of a pattern, may be erroneously labeled as entities (e.g. *"fever"-"never"*).

To avoid this, those entities identified by the rules are then checked by the system using a dictionary of the language from which we have previously removed our symptomatic expressions. Furthermore, frequent erroneous expressions detected in previous analyses of the system output have been added, as in the case of that one shown in Example 3.

> *"We have* ***not found SURGICAL_WOUND_INFECTION*** *a particular bacteria that is causing your recurrent infections."*
>
> (Example 3)

Results

As we can see in Table 1, the coverage for our second approximation (*fuzzy*) has increased to 58.41%, with the categories FEVER and RESPIRATORY INFECTION standing out. The use of fuzzy matching rules now make it possible to capture symptomatic expressions of these categories containing variants and spelling errors. Simultaneously, other categories, such as CNS INFECTION and GASTRENTERITIS, have lost labeled entities, but this is because they were redistributed to other categories. GENITAL INFECTION continues to be the category with the fewest entities, due to the small number of cases of this type in the corpus.

3.3 Third approach: Transformers

Despite the limitations described, the performance achieved through the previous improvements made it possible to generate a set of automatically labeled examples of sufficient size. The availability of this new *silver corpus*, enabled us to implement a new approach based on *transformers* (Vaswani et al., 2017). To do this, we have maintained our *pipeline* architecture, although modifying some of its modules. Specifically, the *Sentence Splitter* and *Entity Recognizer* modules have been removed, since they are not necessary in this new approach, while otheres have been introduced, as shown in Figure 2:

- **Transformer**: Contains the pretrained *language model* used by the system to obtain highly contextualized *embeddings* of the input words.
- **Perceptron**: Responsible for assigning labels from the output of the previous module.
- **Remove Transformer Data**: Responsible for optimizing the memory consumption by removing unnecessary data from the *embeddings*.

Table 1: Analysis of results for the base system (*base*) and its improved version (*fuzzy*) using fuzzy correspondence and term expansion.

Entity type	**No. Tagged Entities**		**No. Documents with tags**		**% Documents with tags**	
	base	*fuzzy*	*base*	*fuzzy*	*base*	*fuzzy*
Bacteremia	6,016	7,488	4,069	4,704	6.80	7.86
CNS Infection	17,999	11,464	9,656	6,426	16.64	10.74
Fever	4,054	34,638	2,628	15,954	4.39	26.67
Gastroenteritis	13,630	10,483	6,730	5,798	11.25	9.69
Genital Infection	3	9	3	9	0.01	0.02
Oral Infection	991	1,108	771	852	1.29	1.42
Phlebitis	100	269	80	231	0.13	0.39
Respiratory Infection	997	29,854	946	15,585	1.58	26.05
Sepsis	31,347	36,045	15,549	16,958	25.99	28.35
Surgical Wound Infection	26	3,334	25	2,374	0.04	3.97
Urinary Infection	1,246	1,298	1,044	1,088	1.75	1.82
Total	76,409	135,990	24,534	34,945	41.01	58.41

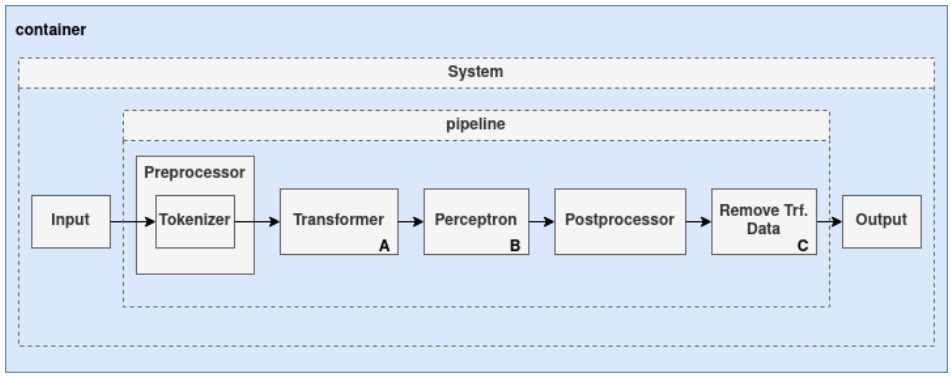


Figure 2: General architecture of the ER system: *transformer*-based approach.

Language models and training

As previously mentioned, we have used the output of our improved rule-based system to automatically generate a *silver corpus*, taking as training examples those sentences that contained labeled cues. Next, the resulting examples set was shuffled in order to balance its distribution and, then, splitted into three (sub)sets: 70% for training, 20% for validation and 10% for testing purposes.

When analyzing the performance of this last approach, several language models have been considered:[4]

- **DistilBERT** (Sanh et al., 2019): *Distilled* version of the base BERT model Devlin et al. (2018). It is smaller (reduced from 100M to 67M parameters) and faster (60%) than the original model, while maintaining a similar linguistic understanding capacity.
- **Bio-DistilBERT** (Rohanian et al., 2022): Model obtained by training DistilBERT on the PubMed dataset[5] in *batches* of size 192 for 200k training cycles.

[4] Downloaded from HuggingFace: *http://huggingface.co* (visited on Sept. 2023).
[5] *http://pubmed.ncbi.nlm.nih.gov* (visited on Sept. 2023).

Table 2: Global comparison of the outputs for our different approaches.

Approach	No. Tagged Entities	No. Documents with tag	% Documents with tag
Base rules	76,409	24,534	41.01
Improved rules	135,990	34,945	58.41
DistilBERT	170,503	41,456	69.30
Bio-DistilBERT	183,968	**44,148**	**73.80**
Bio-ClinicalBERT	166,081	39,386	65.84
PubMedBERT	**192,652**	42,860	71.65

- **Bio-ClinicalBERT** (Alsentzer et al., 2019): Model obtained by training BioBert (Lee et al., 2020) on the MIMIC-III dataset in *batches* of size 32 for 150k cycles.
- **PubMedBERT** (Gu et al., 2020): Model obtained by training BERT on the PubMed dataset in *batches* of size 8,192 for 62.5k cycles.

Systems Comparison

As shown in Table 2, we can see a significant increase in performance with respect to our previous rule-based approaches. In the case of our *transformer*-based approach, the PubMedBERT model obtained a higher number of labeled entities, while Bio-DistilBERT achieved higher document coverage. This increase in performance is mainly due to the generalization capacity of the language models with respect to the use of rules, which is a great advantage when working with unstructured free text content, as in the case of the documents managed by our system.

Regarding the type of symptomatic expressions detected, the new *transformer*-based system is able to identify expressions with large variations and errors, although it might also identify words incorrectly, thus resulting in incorrect labelling, as shown in Example 4. However, some of these incorrectly labeled expressions may also be of interest to the task, thus making the work of annotators easier for the construction of a future *gold standard*: replacing its original label by the correct one should be enough.

> For the symptomatic expression *"hypotension"*, with category SEPSIS, the PubMedBERT model also allows variants such as *"hypotensive"* to be identified. However, it also incorrectly labels terms such as *"hyponatremia"* or *"hypoglycemia"*, since they also contain the prefix *"hypo-"*.
>
> (Example 4)

Acknowledgements

This work was funded by the GRALENIA project (Ref. 2021/C005/00150055) supported by the Spanish Ministry of Economic Affairs and Digital Transformation, the Spanish Secretariat of State for Digitization and Artificial Intelligence, Red.es and by the NextGenerationEU funding. We also acknowledge the European Research Council (ERC), that has also partially funded this research under the Horizon Europe research and innovation programme (SALSA, grant agreement No. 101100615), ERDF/MICINN-AEI (SCANNER-UDC, Ref. PID2020-113230RB-C21), Xunta de Galicia (ED431C 2020/11), and Centro de Investigación del Sistema Universitario de Galicia (SUG) CITIC, funded through the collaboration agreement for the reinforcement of the research centres of the Galician University System (CIGUS) between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities.

Bibliography

E. Alsentzer, J. R. Murphy, W. Boag, W.-H. Weng, D. Jin, T. Naumann, and M. B. A. McDermott. Publicly available clinical bert embeddings, 2019.

J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.

W. W. Fleuren and W. Alkema. Application of text mining in the biomedical domain. *Methods*, 74:97–106, 2015. Text mining of biomedical literature.

Y. Gu, R. Tinn, H. Cheng, M. Lucas, N. Usuyama, X. Liu, T. Naumann, J. Gao, and H. Poon. Domain-specific language model pretraining for biomedical natural language processing, 2020.

A. Hotho, A. Nürnberger, and G. Paas. A brief survey of text mining. *Journal for Language Technology and Computational Linguistics*, 20(1):19–62, 2005.

A. E. Johnson, T. J. Pollard, L. Shen, L.-w. H. Lehman, M. Feng, M. Ghassemi, B. Moody, P. Szolovits, L. Anthony Celi, and R. G. Mark. Mimic-iii, a freely accessible critical care database. *Scientific Data*, 3(1):160035, May 2016.

J. Lee, W. Yoon, S. Kim, D. Kim, S. Kim, C. H. So, and J. Kang. Biobert: a pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics*, 36(4):1234–1240, 2020.

V. I. Levenshtein et al. Binary codes capable of correcting deletions, insertions, and reversals. *Soviet physics doklady*, 10(8):707–710, 1966.

O. Rohanian, M. Nouriborji, S. Kouchaki, and D. A. Clifton. On the effectiveness of compact biomedical transformers, 2022.

V. Sanh, L. Debut, J. Chaumond, and T. Wolf. Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter. *ArXiv*, abs/1910.01108, 2019.

A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. u. Kaiser, and I. Polosukhin. Attention is all you need. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.

Performance of Xiaomi Mi Band 5 Against Polysomnography in Measuring Sleep

Patricia Concheiro-Moscoso, Betania Groba, María del Carmen Miranda-Duro, and Diego Alvarez-Estevez

Universidade da Coruña, CITIC, TALIONIS Group, Faculty of Health Sciences, Campus de Oza, 15071, A Coruña, Spain
Universidade da Coruña, CITIC, 15071 A Coruña, Spain
Centro de Investigación CITIC, Universidade da Coruña, 15071 A Coruña, Spain
Correspondence: patricia.concheiro@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: This study sought to compare the effectiveness of one of the top-selling activity wristbands, the Xiaomi Mi Band 5, with polysomnography in the measurement of sleep stages. Conducted at a hospital in A Coruña, Spain, this study recruited individuals already participating in a polysomnography study at a sleep unit. The study encompassed a total of 45 adults.The Xiaomi Mi Band 5 demonstrated an overall accuracy. Notably, it tended to overestimate total sleep time, light sleep, and deep sleep compared to polysomnography. Conversely, it underestimated wake after sleep onset and REM sleep measured by polysomnography. The Xiaomi Mi Band 5 holds potential as a tool for monitoring sleep and identifying alterations in sleep patterns, especially among individuals without sleep disorders. However, further research is warranted to evaluate its efficacy when used with individuals suffering from various types of sleep disorders.

1 Introduction

Wearable devices have marked a milestone in the revolution of participatory population health by providing unprecedented access to health data with the potential to transform health promotion and well-being (Nieto-Riveiro et al., 2018). These devices are shaping a culture of participatory health, empowering individuals to actively manage and control their own health. Among these devices, wristbands and activity watches have emerged as leaders in this revolution by offering real-time information that enhances awareness and understanding of crucial biomedical parameters, such as sleep (Concheiro-Moscoso et al., 2023).

Sleep has gained increasing recognition in society in terms of its importance and implications for health (Sadek et al., 2019). This growing awareness has spurred interest in monitoring and measuring various aspects and parameters of sleep (de Zambotti et al., 2019). Sleep assessment can be subjective or objective, with subjective tools reflecting the user's or professional's perception of sleep quality and quantity, while objective measurements have sought more practical and cost-effective alternatives to polysomnography (PSG), which has been the gold standard for measuring sleep patterns (Kubala et al., 2020).

This paradigm shift has led to intense research on the validity and applicability of wearable devices in monitoring sleep and related biomedical parameters in diverse populations (Sadek et al., 2019). The technology industry has responded to this ever-growing demand by improving the accuracy and reliability of sleep stage detection, such as light sleep, deep sleep, and REM sleep, as well as the accessibility and comprehension of sleep, activity, and heart rate data for consumers (de Zambotti et al., 2019).

The benefits and positive impact of these devices on the population are becoming increasingly evident as they promote awareness of the importance of sleep and encourage the management of healthy sleep habits and lifestyles (Chong et al., 2020). As society embraces these devices, their influence on the healthcare and research domains becomes more prominent. Healthcare professionals and healthcare systems have begun integrating these devices into clinical practice, enabling more comprehensive user assessments and more accurate monitoring of sleep and other biomedical parameters (Concheiro-Moscoso et al., 2021).

Moreover, large companies and healthcare organizations have shown interest in using these devices to monitor and enhance the health of their employees and customers (Chong et al., 2020). The validation of sleep data quality provided by these devices has become a necessity to ensure their reliable use in daily life, clinical practice, and research (Concheiro-Moscoso et al., 2022). Despite the growing adoption of wearables, there are few validation studies comparing sleep data to standard methods, such as PSG (de Zambotti et al., 2019).

Therefore, this study aimed to assess the quality of sleep data from the Xiaomi Mi Band 5 wearable device. The findings of this research contribute to our understanding of the reliability and accuracy of sleep data provided by wearable technology, further emphasizing the potential impact of wearables on healthcare, research, and overall population health.

2 Objectives

The primary goal of this study was to evaluate how well the Xiaomi Mi Band 5 measures sleep-wake stages in comparison to polysomnography (PSG) conducted in a hospital sleep unit. Additionally, the study had several secondary objectives:

1. To establish the level of agreement between sleep measurements obtained from PSG and those from the Xiaomi Mi Band 5.
2. To assess the Xiaomi Mi Band 5's accuracy, specificity, and sensitivity in classifying sleep and wake stages in comparison to PSG.
3. To evaluate the Xiaomi Mi Band 5's performance in detecting specific sleep stages, including wakefulness, light sleep, deep sleep, and REM sleep, in comparison to PSG results.

3 Methodology

3.1 Study Setting and Participant Profile

This project is an observational and prospective study conducted within a sleep unit at a hospital in A Coruña. The study targeted individuals over 18 years of age who were undergoing polysomnography (PSG) for clinical purposes unrelated to this research.The study was approved by the A Coruña-Ferrol research ethics committee (2020/318).

Calculations indicated a required sample size of 43 participants for achieving statistically significant results, yet the study enrolled 58 participants. However, data from only 45 participants (predominantly males with a mean age of 53.24 ± 15.44 years) were included in the analysis due to missing or invalid data from the remaining 13 participants.

Participants were categorized into two groups based on the presence of sleep disorders: participants without sleep problems, $n = 20$) and SDis (with sleep problems, $n = 25$). Thus, data analysis was conducted both on the entire participant cohort and separately for the established groups based on sleep disorder presence.

3.2 Data Processing

In accordance with established standards, data processing followed the recommendations and guidelines set by the American National Standards Institute and the Consumer Technology

Association (ANSI/CTA Standard). These guidelines stipulate that the validation analysis of a device should be carried out using sleep stages segmented into 30-second epochs. While PSG records sleep states in 30-second epochs throughout the entire sleep cycle, the data provided by the Xiaomi Mi Band, whether in its application or in CSV files, consisted of the total duration of each sleep stage within its cycle (Kemp and Olivan, 2003).

To ensure the highest level of precision in the analysis, the research team reprocessed the data from both devices. This involved transforming the data into the European Data Format+ (EDF+) using a Python script developed with the PYEDFlib library (Concheiro-Moscoso et al., 2023). Specifically, the acquisition of Xiaomi Mi Band data in 30-second epochs was synchronized with the PSG-defined sleep onset and the recorded light on/off events, which were documented in the sleep unit's report.

3.3 Statistical Analysis

Statistical analysis was conducted using R software (version 4.1.2; R Foundation for Statistical Computing) with the entire sample size (n=45). Comparisons between summary measures of PSG and the Xiaomi Mi Band 5 equivalents were executed utilizing either the paired 2-tailed t-test (t) or the Mann-Whitney Wilcoxon test (z), contingent upon the normality of the data as determined by the Shapiro-Wilk normality test. Parametric 2-tailed t-tests were applied to normally distributed data, while non-normally distributed data underwent analysis via the Mann-Whitney Wilcoxon test. Moreover, the Bland-Altman method was employed to assess the concordance between PSG and the Xiaomi Mi Band 5 concerning each sleep parameter. Calculations included the mean difference (or bias) between the two methods, the standard deviation (SD), the 95% confidence interval (CI), and the Bland-Altman 95% limits of agreement (mean observed difference $\pm$ 1.96 $\times$ SD of observed differences). A positive bias indicated that the Xiaomi Mi Band 5 tended to underestimate a variable in comparison to the gold standard (PSG), while a negative bias indicated overestimation.

4 Results

The first test that marked the validation process involved comparing the summarized sleep measures obtained from both devices using a t-test. In this initial analysis, it is noteworthy that similar results were observed between the Xiaomi device and PSG in sleep-related variables such as initial sleep onset, total sleep period duration (TSPD), and sleep onset latency (SOL) across the entire participant group. Specifically, comparable outcomes were noted between both devices in total sleep time (TST), wake time after sleep onset (WASO), and deep sleep time in the group without sleep disorders, while there was similarity in SOL and light sleep time among participants with sleep disorders.

Table 1: Comparison of PSG and Xiaomi Mi Band 5 sleep measures in the total sample

Name	t	z	P
Initial sleep onset (hh:mm)	1.12	-	0.266
TSPD (minutes)	0.28	-	-0.778
SOL (minutes)	-	-1.07	0.288

Overall, the results of the Bland-Altman analysis (see Figure 1) indicated that the Xiaomi Mi Band 5 device tended to overestimate measures of TST, Sleep Efficiency (SE), time spent in light sleep, and time spent in deep sleep when compared to PSG. Conversely, Xiaomi Mi Band 5 showed a tendency to underestimate measures of WASO and Rapid Eye Movements (REM) sleep in comparison to PSG measurements.

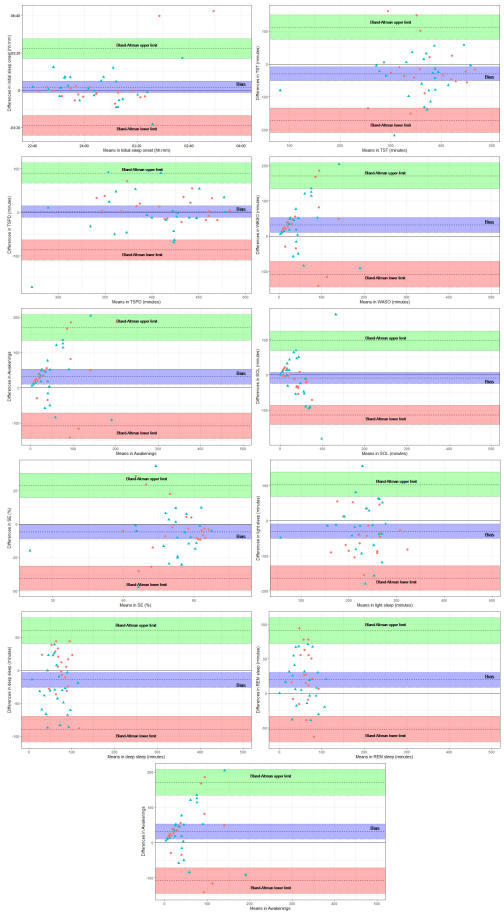


Figure 1: Bland Altman Plot

With regard to the results obtained from the Epoch by Epoch (EBE) detection test using a Python script, overall, the Xiaomi wristband exhibited greater accuracy (the ability to detect sleep-wake stages), as well as higher sensitivity (the ability to detect sleep stages) compared to its specificity (the ability to detect wake stages) in detecting the phases of the sleep-wake cycle.

Table 2: Comparison of PSG and Xiaomi Mi Band 5 sleep measures in the total sample

		Xiaomi	
		Wake	Sleep
PSG	Wake	3019	5525
	Sleep	3238	27786

5 Discussion and Conclusions

In comparison to existing literature, the results obtained with the Xiaomi Mi Band 5 wristband and PSG are consistent with findings from previous comparative studies involving devices such as Fitbit HR, Fitbit Charge 2, or Jawbone UP (de Zambotti et al., 2019). Specifically, minor differences were observed between the Xiaomi device and PSG in certain summarized measures when compared to these devices. Furthermore, it is noteworthy that the results obtained in the group without sleep disorders generally exhibited greater accuracy compared to the group with sleep disorders. In this regard, the Xiaomi Mi Band 5 could serve as a valuable tool for health monitoring in the population without sleep disorders (Concheiro-Moscoso et al., 2023). Nonetheless, further research comparing this device to the gold standard, PSG, in diverse populations with and without sleep disorders is needed to assess its accuracy and reliability.

Funding

CITIC is financed by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidade and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS). The research carried out by TALIONIS group is financed by the Xunta de Galicia (Aid from the Consellería de Cultura, Educación Formación Profesional e Universidades for the consolidation and structuring of competitiveresearch units) ED431B 2022/39. DAE received funding from the project ED431H 2020/10 of the Xunta de Galicia. The publication is part of the project TED2021-130127A-I00, funded by MCIN/AEI/10.13039/501100011033 and by the European Union "NextGenerationEU"/PRTR.

Bibliography

K. P. L. Chong, J. Z. Guo, X. Deng, and B. K. P. Woo. Consumer perceptions of wearable technology devices: Retrospective review and analysis. *JMIR Mhealth Uhealth*, 8(4):e17544, Apr 2020. URL *http://mhealth.jmir.org/2020/4/e17544/*.

P. Concheiro-Moscoso, F. J. Martínez-Martínez, M. d. C. Miranda-Duro, T. Pousada, L. Nieto-Riveiro, B. Groba, F. J. Mejuto-Muiño, and J. Pereira. Study protocol on the validation of the quality of sleep data from xiaomi domestic wristbands. *International Journal of Environmental Research and Public Health*, 18(3), 2021. URL *https://www.mdpi.com/1660-4601/18/3/1106*.

P. Concheiro-Moscoso, B. Groba, F. J. Martínez-Martínez, M. del Carmen Miranda-Duro, L. Nieto-Riveiro, T. Pousada, and J. Pereira. Use of the xiaomi mi band for sleep monitor-

ing and its influence on the daily life of older people living in a nursing home. *DIGITAL HEALTH*, 8:20552076221121162, 2022.

P. Concheiro-Moscoso, B. Groba, D. Alvarez-Estevez, M. d. C. Miranda-Duro, T. Pousada, L. Nieto-Riveiro, F. J. Mejuto-Muiño, and J. Pereira. Quality of sleep data validation from the xiaomi mi band 5 against polysomnography: Comparison study. *Journal of Medical Internet Research*, 25:e42073, 2023.

M. de Zambotti, N. Cellini, A. Goldstone, I. M. Colrain, and F. C. Baker. Wearable sleep technology in clinical and research settings. *Medicine and science in sports and exercise*, 51(7): 1538—1557, July 2019. URL *https://europepmc.org/articles/PMC6579636*.

B. Kemp and J. Olivan. European data format 'plus' (edf+), an edf alike standard format for the exchange of physiological data. *Clinical Neurophysiology*, 114(9):1755–1761, 2003. URL *https://www.sciencedirect.com/science/article/pii/S1388245703001238*.

A. G. Kubala, B. Barone Gibbs, D. J. Buysse, S. R. Patel, M. H. Hall, and C. E. Kline. Field-based measurement of sleep: Agreement between six commercial activity monitors and a validated accelerometer. *Behavioral sleep medicine*, 18(5):637—652, 2020. URL *https://europepmc.org/articles/PMC7044030*.

L. Nieto-Riveiro, B. Groba, M. C. Miranda, P. Concheiro, A. Pazos, T. Pousada, and J. Pereira. Technologies for participatory medicine and health promotion in the elderly population. *Medicine*, 97(20):e10791, May 2018. URL *https://europepmc.org/articles/PMC5976343*.

I. Sadek, A. Demarasse, and M. Mokhtari. Internet of things for sleep tracking: wearables vs. nonwearables. *Health and Technology*, 2019.

Rationale and Design of a Scope 3 Capital Charge

Davide Trevisani, José Germán López, Chris Kenyon, Carlos Vázquez, and Mourad Berrahoui

Department of Mathematics and CITIC, University A Coruña, Spain
Department of Mathematics, University College London, UK
Lloyds Banking Group, UK
Correspondence: davide.trevisani@udc.es, jose.lsalas@udc.es, c.kenyon@ucl.ac.uk, carlos.vazquez.cendon@udc.es, mourad.berrahoui@lloydsbanking.com

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: Climate change is caused by greenhouse gas emissions, and governments have introduced over seventy carbon pricing instruments (CPIs). Banks finance a significant fraction of global emissions, and many have committed to reduce their facilitated, or Scope 3, emissions to (net) zero by 2050. However, it is possible that governments will introduce a CPI impacting banks on their Scope 3 emissions earlier. Here we design a Scope 3 capital charge to make banks resilient against the possibility, albeit not certainty, that governments could introduce such a Scope 3 CPI. Based on interest rate swaps, our numerical examples are financially significant for counterparties with significant emissions. The contribution of this work is to provide a technical basis for banks to be sufficiently resilient.

1 Introduction

Climate change is caused by greenhouse gas emissions (IPCC, 2021), and governments have introduced over seventy carbon pricing instruments (CPIs) (World Bank, 2023). Banks finance a significant fraction of global emissions and some of them publish their Scope 3 emissions. It is possible that governments will introduce a CPI impacting banks on this type of emissions before 2050. This introduction for banks would be a significant stress as it would have the characteristics of a correlated market event, thus creating an impact for every bank counterparty with significant facilitated emissions. In contrast to this possible scenario, in the financial industry, capital requirements already exist, and from a regulatory point of view, they are set to ensure that banks are "sufficiently resilient to withstand losses in times of stress" (BCBS, 2017). Regulators are highly aware of climate-related risks (BCBS, 2022; BoE, 2021a; European Central Bank, 2020; FSB, 2017) but have generally stated that the introduction of specific climate-related capital charges is not needed (BoE, 2021b, 2023). They expect climate effects to be already included through existing credit risk frameworks, although they remain open to proposals (Federal Reserve System, 2022). However, the aforementioned potential of governments acting directly to curb Scope 3 emissions via CPIs has not been considered (Holscher et al., 2022; Oehmke, 2022). In this work, we aim to consider this possibility (not certainty) and propose a design of a new Scope 3 capital charge. We also present some numerical examples that could be of interest from the regulators' perspective.

The plan of this work is as follows. In Section 2 we present the mathematical modeling, and the derivation of derivative pricing from regulators perspective by mean of semi-replication.

Section 3 is devoted to risk evaluation to define the CPI-capital measure (CPIC). Next, in Section 4 we present numerical examples based on vanilla IRSs. Finally, some conclusions are presented in Section 5.

2 Mathematical modelling

We denote by V_t the risk-free value of the derivative, and $\widehat{V}_t$ the value from the point of view of regulators. We start by considering a timeline $[T_0, T]$, and a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[T_0,T]}, Q)$. Such space is assumed to support three-dimensional correlated Brownian motions (W, W^F, W^S) and three Poisson processes J^B, J^C, and J^{co}, independent of each other. Given a larger maturity $\bar{T} \geqslant T$, the economy we consider is given by two blocks of Markovian dynamics. The first block is

$$dB_t/B_t = r_t\,dt \qquad \text{Riskless bank account,} \tag{38.1}$$

$$dP_{t,\bar{T}}/P_{t,\bar{T}} = r_t dt + \sigma_{t,\bar{T}} dW_t \qquad \text{Default-free bond,} \tag{38.2}$$

$$dP^B_{t,\bar{T}}/P^B_{t,\bar{T}} = r^B_t dt + \sigma^B_{t,\bar{T}} dW_t + (1 - R^B_t) dJ^B_t \qquad \text{Own bond,} \tag{38.3}$$

$$dP^C_{t,\bar{T}}/P^C_{t,\bar{T}} = r^C_t dt + \sigma^C_{t,\bar{T}} dW_t + (1 - R^C_t) dJ^C_t \qquad \text{Counterparty bond,} \tag{38.4}$$

while the second block is

$$dF_t/F_t = \mu^F_t dt + \sigma^F_t dW^F_t \qquad \text{Carbon Future contract,} \tag{38.5}$$

$$dS_t/S_t = \mu^S_t dt + \sigma^S_t dW^S_t \qquad \text{Counterparty share price.} \tag{38.6}$$

Furthermore, additional costs are due to CPI, that is the jump-diffusion process

$$\mathcal{T}_t := g(V_t, \mathrm{BS}_t, F_t) J^{co}_t, \tag{38.7}$$

where the charge function g represents the mitigation cost of carbon emissions caused by V. In the absence of collateral, we have

$$g(V_t, \mathrm{BS}_t, F_t) := \frac{\max(V_t, 0)}{\mathrm{BS}_t} E_t F_t, \tag{38.8}$$

where E_t is the counterparty's emission in tonnes per year, and BS_t is the counterparty's balance sheet given by

$$\mathrm{BS}_t := S_t \times (\#\text{ outstanding shares})_t\,. \tag{38.9}$$

2.1 Pricing by semi-replication

We derive a model to price derivatives from the regulators' perspective. The formulas here obtained extend the xVA approach of (Burgard and Kjaer, 2013), and they naturally extend to other types of derivatives. Namely, we consider a semi-replicating and self-financing portfolio Π of the form

$$\Pi = \widehat{V} + \delta P + \alpha^C P^C, \tag{38.10}$$

and financed by $\alpha^B P^B$ firm's bonds, and a mixed cash account β. Given our framework, the self-financing condition reads as

$$d\widehat{V} + \delta dP + \alpha^C dP^C = \alpha^B dP^B + d\beta + \mathcal{T} dt, \tag{38.11}$$

where

$$\mathcal{T} = g(V, \mathrm{BS}, F) J^{co}.$$

represents instantaneous CPI costs. By a standard hedging argument (see Kenyon et al. (2023) for more details) we find that the cost of carbon taxes G is given by the solution of the PDE

$$\begin{cases} \frac{\partial G}{\partial t} + \mathcal{A}_t G = (r^B + \lambda^C)G + g(V, \mathrm{BS}, F)J_t^{co}, \\ G_T(P) = 0, \end{cases} \tag{38.12}$$

with the differential operator $\mathcal{A}_t := rP\frac{\partial}{\partial P} + \frac{1}{2}\sigma^2 P^2 \frac{\partial^2}{\partial P^2}$. In particular, for the above-stated PDEs a dependence on the path-realization of $\mathcal{T}_t$ emerges. Accordingly, we introduce the notation $\mathcal{T}^t := \{\mathcal{T}_s | \, s \in [t, T]\}$ to denote the set of paths of $\mathcal{T}$ starting at time t. Under the assumption of existence and regularity, by Feynman-Kàc representation Theorem, the solution of (38.12) at time $t = T_0$ can be expressed as the random variable, defined for $\omega \in \Omega$ as

$$G(T_0, T, P, \mathcal{T}^{T_0}(\omega)) = \mathbb{E}\left[\int_{T_0}^{T} e^{-\int_{T_0}^{u} r_s^B + \lambda_s^C \, ds} \mathcal{T}_u(\omega)\, du \,\Big|\, P_{T_0} = P\right]. \tag{38.13}$$

3 Risk evaluation

From the previously outlined results, the random CPI introduction (38.13) determines an unexpected loss. We then start with a thorough examination of the expected-unexpected loss. By denoting with M_{T_0} the quadruple $(P_{T_0}, F_{T_0}, \mathrm{BS}_{T_0}, J_{T_0}^{co})$ at present time T_0 we have

$$\begin{aligned} \mathrm{EL}(T_0, T, M_{T_0}) &:= \mathbb{E}[G(T_0, T, P, \mathcal{T}^{T_0}(\omega))|M_{T_0}] \\ &= \mathbb{E}\left[\int_{T_0}^{T} e^{-\int_{T_0}^{u} r_s^B + \lambda_s^C ds} g(V_u, \mathrm{BS}_u, F_u) J_u^{co}\, du \Big| M_{T_0}\right], \end{aligned} \tag{38.14}$$

while, for an arbitrary $t \in [T_0, T]$, the expected loss accruing between t and T becomes

$$\mathrm{EL}(t, T, M_t) = \mathbb{E}\left[\int_{t}^{T} e^{-\int_{T_0}^{u} r_s^B + \lambda_s^C ds} g(V_u, \mathrm{BS}_u, F_u) J_u^{co} du \,\Big|\, M_t\right]. \tag{38.15}$$

Note that the expectations in (38.14) or (38.15) neglect the presence of potential fat tails of the CPI-profile. Hence, while this measure does offer insight into the scale of the issue, it is unsuitable for shielding the firm from extreme circumstances. The expected (unexpected) loss EL of (38.14) can be seen as an analogous counterpart to the expected exposure, as it is defined in (BCBS, 2023). Accordingly, we define the following CPI-capital measure (CPIC):

$$\boxed{\mathrm{CPIC}(T_0, T, M_{T_0}) := \alpha \int_{T_0}^{(T_0+1y)\wedge T} \max_{t\in[T_0,s]} \mathbb{E}\left[\mathrm{EL}(t, T, M_t)\big|M_{T_0}\right] ds,} \tag{38.16}$$

where α denotes the capital scaling factor. For CCR capital, $\alpha = 1.4$ in the RWA formula, but note that this CPIC formula is a capital formula, not a RWA formula — in this, it is similar to the SA-CVA capital formula.

Indeed, an analogy of the measure (38.16) can be found for CCR-capital in CRE 53.13 (BCBS, 2023). To simplify the computation of CPIC, it follows by the Tower propriety of conditional expectation (see Mikosch (1999)) that

$$\mathbb{E}[\mathrm{EL}(t, T, M_t)|M_{T_0}] = \mathbb{E}\left[\int_{t}^{T} e^{-\int_{t}^{u} r_s^B + \lambda_s^C ds} Q(J_u^{co} = 1) g(V_u, \mathrm{BS}_u, F_u)\, du \Big| M_{T_0}\right]. \tag{38.17}$$

4 Numerical example

We focus on a vanilla interest rate swap (IRS). For this product, we compute the expected-expected losses, the CO_2eVA for different dates, and the resulting CPIC capital. The key quantity to consider is the participation charge g. As a function of the state variables we set

$$\tilde{g} = \tilde{g}(r_u, F_u, S_u) = x_t \beta_t \frac{\max(V(r_u), 0)}{\mathrm{BS}(S_u)} E_u F_u = x_t \beta_t \frac{\max(\mathrm{IRS}(r_u), 0)}{\mathrm{BS}(S_u)} E_u F_u.$$

Factors x_t and β_t scale mitigation costs. The former, x_t, represents the gradual introduction of CPIs ; the latter, β_t denotes a reduction of the counterparty's emissions by its own efforts or technological improvements. We then consider a stylized regional low-cost airline and a stylized shipping company as possible counterpart of the IRS, which differ among themselves in terms of emissions and balance sheet (see Table 1).

Table 1: Characteristic of stylized airline and shipping companies. Note that the "shares" represent equity plus debt.

Example	CO2 million tonnes / year	Balance Sheet (milions USD)
Airline	2.7	2,600
Shipping	10.8	3,700

For both counterparties in our examples, β decreases linearly from $\beta_{T_0} = 100\%$ to $\beta_{2050} = 50\%$. In the respect of x_t, we consider $x_{T_0} = 0\%$, it increases linearly to 100% in 2040 and is constant thereafter. The other scenario uses 2040 as the reference point for x_t and β_t. We assume that the full annual mitigation cost is to be paid yearly after the reference point.

4.1 Stochastic dynamics

For simplicity, we consider a single-curve short-rate model. For the same reason, the counterparty stock price and the future carbon price are modelled through geometric Brownian motions (GBM).

1. For the short-rate we use a classical Hull-White dynamic

$$dr_u = [\vartheta(u) - a r_u]du + \sigma dW_u, \tag{38.18}$$

   with constant $a, \sigma \in \mathbb{R}^+$ and initial short-rate r_0. For simplicity and reproducibility of our results, we consider an hypothetical flat forward curve with a constant value of 3.2%. Likewise, the mean reversion parameter a and the volatility σ are equal to 5% and 2%, respectively.

2. Futures carbon prices are calibrated to inflation-adjusted NGFS scenarios, expressed in USD rebased from 2010. Future inflation is constant and equal to 3%.

3. Both the growth rate and volatility of counterparties' stock are constant and equal to $\mu^S = 10\%$ and $\sigma^S = 20\%$ respectively.

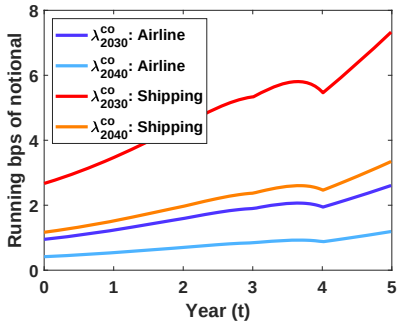
4. Counterparts' probability of default in 1-year is 1%, and no funding costs are considered.

5. Jumps intensity λ^{co} is chosen so that $Q(J^{co}_{2030} = 1) = 50\%$ or $Q(J^{co}_{2040} = 1) = 50\%$.

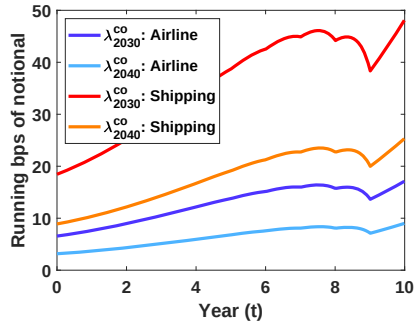
6. The correlation between the noises of the short rate and the carbon future is 50%. The other correlations are set equal to zero.

4.2 Results

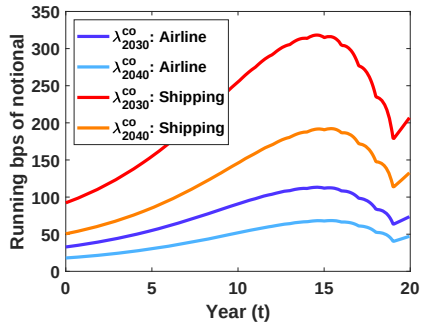
In this part, numerical results for ATM 5-10 and 20 years IRS are presented. We distinguish the cases given by the stylized airline and the shipping company, as well also different values of λ^{co}. For these results, the scaling factor x_t reaches 100% in 2040. During the life of the contract, the expected-expected CPIs costs evolve as in Figure 1.



(a) Maturity $T = 5$ years



(b) Maturity $T = 10$ years



(c) Maturity $T = 20$ years

Figure 1: Plots of expected-expected CPIs costs $t \mapsto \mathbb{E}[\mathrm{EL}(t, T, M_t)|M_{T_0}]$ (see (38.17)) for IRS with different maturities. The value of λ^{co} is the hazard rate for tax-introduction. In case CPIs costs arise with a cumulative probability of 50% before 2030, the hazard rate $\lambda^{co}_{2030} = 1042$ bps. Analogously, the hazard rate $\lambda^{co}_{2040} = 413$ bps. Because of higher chances of introduction, λ^{co}_{2030} gives higher CPIs costs.

Table 2 compares CPIC-capital levels to SA-CCR capital levels (i.e. SA-CVA RWA times 8% with 100% risk weight, so roughly BBB rating for AIRB). For short IRS, i.e. 5-year, the CPIC-capital is much smaller for the counterparties considered, at up to 24%. However, for longer IRS, 10-year and 20-year, the CPIC-capital levels are comparable to SA-CCR capital and become much larger. This non-linear effect is typical in climate pricing where there are increasing carbon costs over time, and increasing cumulative probability of Scope 3 CPI introduction.

Table 3 assesses the effectiveness of CPIC-capital. Particularly, in this table we report how often the CPIC-capital succeeds in covering the actual costs. Namely, for the random cost

$$X := \int_{T_0}^{T} e^{-\int_{T_0}^{u} r_s^B + \lambda_s^C \, ds} g(V_u, BS_u, F_u) J_u^{co} du, \tag{38.19}$$

we estimate the probability of $(X \leqslant \mathrm{CPIC})$.

Table 2: CPIC-capital compared with SA-CCR capital (i.e. SA-CCR RWA x 8% with 100% risk weight, so roughly BBB rating with AIRB). Hazard rate is $\lambda^{co}_{2040} = 413$ bps, and the capital scaling factor for CPI capital is set to $\alpha = 1$.

Example	Maturity (Years)	CPIC (bps of not.)	CPIC (run. bps of not.)	SA-CCR (bps of not.)	SA-CCR (run. bps of not.)	ratio CPIC / SA-CCR
A	5	2	0	25	5	0.08
A	10	33	3	44	4	0.75
A	20	379	19	71	4	5.35
S	5	6	1	25	5	0.24
S	10	92	9	44	4	2.09
S	20	1066	53	71	4	15.06

Table 3: Effectiveness of CPIC-capital. Given future histories we consider whether the capital is sufficient to absorb the CPI-introduction losses. Over 80% of the time, the CPI capital is effective. To increase CPIC effectiveness to 97.5% the capital required would need to be roughly 10x larger. Hazard rate is $\lambda^{co}_{2040} = 413$ bps, and the CPI capital scaling factor $\alpha = 1$.

Example	Maturity (Years)	CPIC (bps of not.)	(CPIs $\leqslant$ CPIC) (probability)	97.5-perc. (bps of not.)	ratio to get to 97.5-perc.
A	5	2	91%	25	13
A	10	33	86%	319	10
A	20	379	82%	2990	8
S	5	6	90%	70	12
S	10	92	86%	897	10
S	20	1066	82%	8404	8

5 Discussion and Conclusions

In this work we introudced a Scope 3 capital charge to make banks resilient against the possibility, though not certainty, that governments could introduce such a Scope 3 CPI. We focused our analysis on the Trading Book, i.e. derivatives, but the extension to the Banking Book, and thus loans, is a straightforward simplification. Our examples, based on interest rate swaps, show that as contracts increase from 5-years to 10-years or more the CPI-capital moves from a fraction of SA-CCR capital to multiples of SA-CCR capital. Thus, for the counterparties considered, the CPI-capital is highly necessary. Our Scope 3 CPI capital design is both effective, being sufficient in 80% of simulations, and parsimonious. To move the effectiveness to 97.5% would require an order of magnitude increase. This justifies our choice of a capital multiplier $\alpha = 1$. Of course, these numbers are indicative and any regulatory introduction would need more extensive justification.

Acknowledgements

The authors would like to gratefully acknowledge discussions with participants in recent WBS and ISDA conferences. Davide Trevisani, José Germán López, and Carlos Vázquez acknowledge the funding from the Galician Government (grant ED43IC 2022/047). Davide Trevisani and Carlos Vázquez acknowledge the funding from European Union through the ABC-EU-

XVA inside the call H2020-MSCA-ITN-2018 (Grant Agreement 813261). CITIC is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS).

Bibliography

BCBS. High-level summary of Basel III reforms, 2017. *https://www.bis.org/bcbs/publ/d424_hlsummary.pdf*.

BCBS. Frequently asked questions on climate related financial risks. *https://www.bis.org/bcbs/publ/d543.pdf*, 2022.

BCBS. Calculations of RWA for credit risk, 2023. *www.bis.org/basel_framework/standard/CRE.htm*.

BoE. Guidance for participants of the 2021 biennial exploratory scenario: Financial risks from climate change, 2021a.

BoE. Climate-related financial risk management and the role of capital requirements, 2021b. *https://www.bankofengland.co.uk/-/media/boe/files/prudential-regulation/publication/2021/october/climate-change-adaptation-report-2021.pdf*.

BoE. Bank of england report on climate-related risks and the regulatory capital frameworks, 2023. *https://www.bankofengland.co.uk/prudential-regulation/publication/2023/report-on-climate-related-risks-and-the-regulatory-capital-frameworks*.

C. Burgard and M. Kjaer. Funding strategies, funding costs. *Risk*, 26(12):82–87, 2013.

European Central Bank. Supervisory expectations relating to risk management and disclosure. Guide on climate-related and environmental risks., 2020.

Federal Reserve System. Principles for climate-related financial risk management for large financial institutions. Federal Register 12/08/2022, 2022. *https://www.federalregister.gov/documents/2022/12/08/2022-26648/principles-for-climate-related-financial-risk-management-for-large-financial-institutions*.

FSB. Recommendations of the task force on climate-related financial disclosures. Task Force on Climate-related Financial Disclosures, 2017.

M. Holscher, D. Ignell, M. Lewis, and K. Stiroh. Climate change and the role of regulatory capital: A stylized framework for policy assessment. Finance and Economics Discussion Series (FEDS), 2022. *https://www.federalreserve.gov/econres/feds/climate-change-and-the-role-of-regulatory-capital.htm*.

IPCC. Climate Change 2021: The Physical Science Basis, 2021.

C. Kenyon, A. Macrina, and M. Berrahoui. CO2eVA: pricing carbon externalities transition. *Available at SSRN 4136710*, 2023.

T. Mikosch. *Elementary stochastic calculus with finance in view*. World scientific, 1999.

M. Oehmke. Bank capital regulation and climate change. ASC Insight, No 3 / November, 2022. *https://www.esrb.europa.eu/pub/asc/insights/shared/pdf/esrb.asc.insight_03_11_22~c72a4ae30d.en.pdf*.

World Bank. State and trends of carbon pricing 2023, 2023.

Interactive Virtual Reality Public Speaking Training Simulator to Empower Confident Communicators

João Donga, Paulo Veloso Gomes, António Marques, and Javier Pereira

LabRP, Center for Rehabilitation Research, School of Media Arts and Design, Polytechnic of Porto, Porto, Portugal
LabRP, Center for Rehabilitation Research, School of Health, Polytechnic of Porto, Porto, Portugal
CITIC Research Center, University of A Coruña, A Coruña, Spain
Correspondence: jpd@esmad.ipp.pt

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: Public speaking is a valuable skill that plays a crucial role in personal, academic and professional success. Many individuals struggle with anxiety and lack confidence when facing an audience. To address this challenge, a virtual reality public speaking simulator has been developed to empower individuals and enhance their public speaking abilities. The virtual reality simulator offers a dynamic platform for immersive practice and training, incorporating distinct modes. Autonomous mode and interactive real-time intervention by physiological biofeedback mechanism and/or a therapist. By leveraging the immersive nature of virtual reality, this simulator aims to boost users confidence, enhance their articulation, persuasion skills and foster their ability to connect with diverse audiences.

1 Introduction

Public speaking is a fundamental skill that has played a pivotal role in human communication for centuries (Lillis, 2005). It is the art and science of addressing an audience, conveying a message, and inspiring, persuading, or informing others through spoken words. The ability to speak effectively in public is a skill that transcends professional fields and personal contexts (Ningsih et al., 2023), (OECD and Organization, 2017).

Public speaking is not merely about talking in front of a group; it is a powerful mean of connecting with people, sharing ideas, and influencing minds. It's a skill that can empower individuals to lead, educate, inspire change, and foster understanding. Moreover, it is a skill that, when honed, can boost self-confidence and personal growth.

Glossophobia, derived from the Greek words "glossa" meaning tongue or language and "phobos" meaning fear, is a common and deeply ingrained fear that afflicts a significant portion of the population. It refers to the fear of public speaking or speaking in front of an audience. This social anxiety disorder can range from mild nervousness to severe panic and can have a profound impact on a person's personal and professional life (Dansieh et al., 2021).

Glossophobia is more widespread than one might think. Surveys consistently rank the fear of public speaking as one of the most common fears, often surpassing even the fear of death in prevalence (Ambit and Pandayan, 2020). It affects people of all ages, backgrounds, and levels of experience. Even accomplished professionals and seasoned public speakers can experience glossophobia when facing a particularly daunting audience or topic.

The origins of glossophobia are multifaceted and can vary from person to person (Bodie, 2010). Some common causes and contributing factors (Nahliah and Rahman, 2018) include: Fear of Judgment: Many individuals fear being evaluated or criticized by their audience. This fear of judgment can be particularly paralyzing, as it taps into our innate desire for acceptance and approval. Performance Anxiety: Glossophobia often stems from a fear of making mistakes, forgetting one's lines, or losing one's composure in front of others. The pressure to perform flawlessly can be overwhelming. Traumatic Experiences: Past negative experiences with public speaking, such as embarrassing moments or harsh criticism, can leave lasting psychological scars and reinforce the fear. Lack of Preparation: Insufficient preparation and a lack of confidence in one's knowledge or content can exacerbate glossophobia. Feeling unprepared can amplify feelings of anxiety. Social Anxiety: For some individuals, glossophobia is a subset of a broader issue of social anxiety. The fear of speaking in public is just one manifestation of a more generalized social anxiety disorder.

Glossophobia is a treatable condition. Various strategies and interventions can help individuals overcome their fear of public speaking, including:

Practice: Repeated exposure and practice are effective in desensitizing individuals to the fear of public speaking (Ortiz et al., 2022). Cognitive-Behavioral Therapy (CBT): CBT can help individuals identify and challenge irrational beliefs and thought patterns related to public speaking (Anderson et al., 2005). Breathing and Relaxation Techniques: Learning to manage anxiety through techniques such as deep breathing and progressive muscle relaxation can be helpful. Public Speaking Courses: Enrolling in public speaking courses or joining organizations provide a structured and supportive environment for skill development.

Virtual reality can be a valuable tool for helping individuals overcome glossophobia by providing a safe and controlled environment for exposure therapy and skill development (Lindner et al., 2021).

This work aims to create an interactive immersive virtual reality system to practice public speaking techniques and that can also be used in glossophobia treatment programs.

The virtual environment aims to recreate a simulated environment in which the participant is in front of an audience that reacts to their speech. Throughout the presentation, the environment can change its behavior autonomously, through the intervention of the therapist, or through real-time biofeedback.

2 Methods

A virtual reality system was designed to allow the training of public speaking through practice and Cognitive-Behavioral Therapy.

It consists on a solution that replicates a conference auditorium (Figure 1) and can be used in two modes:

- Autonomous mode: For use independently by the user.
- Controlled Mode: For use with therapist monitoring and intervention.
- Real-time biofeedback mode: For use with therapist monitoring and intervention.

Figure 1: Conference auditorium

The system allows to choose several parameters such as: duration of the presentation, noise level, number of audience present (Figure 2).

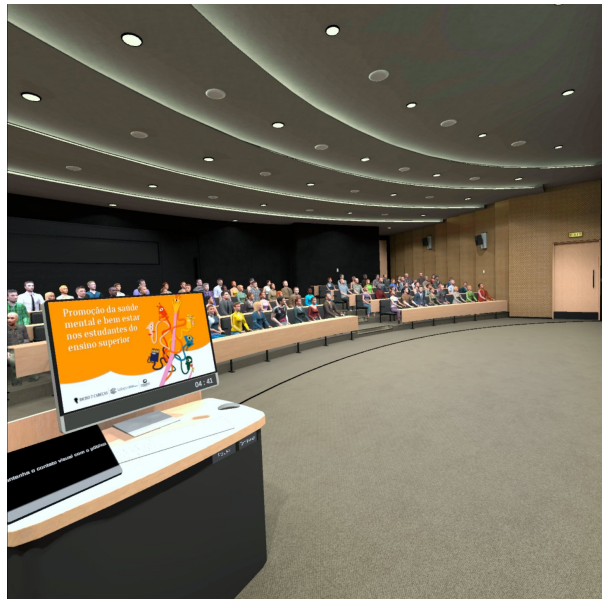


Figure 2: Parameters

In controlled mode, the therapist can vary the factors to reduce or increase the level of stress. He can send predefined or personalized messages, in real-time, to the patients during presentation to calm them down and suggestions to improve their performance (Figure 3).

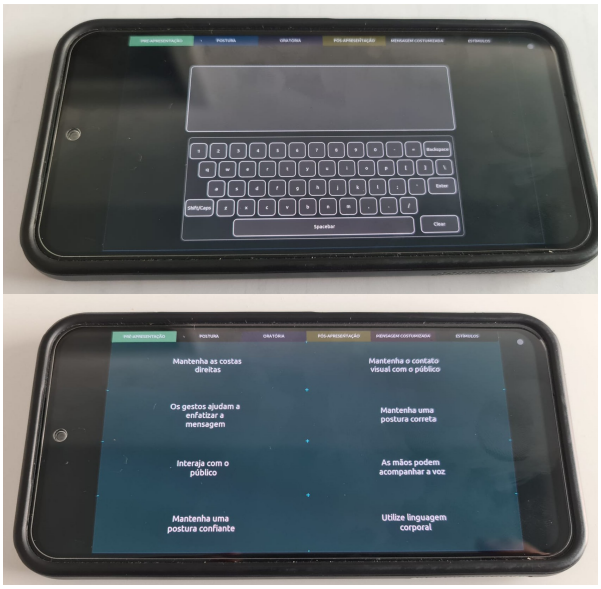


Figure 3: Therapist Module

3 Future Work

The use of a third mode using real-time biofeedback in conjunction with an AI algorithm will allow the system to react in real time recognizing the type and intensity of user emotions and adapt the variables to optimize the experience.

Also the development of other scenarios that allow the user to trains public speaking in different situations such as meetings, social events, interviews should be considered using the framework developed.

4 Conclusion

The developed VR environment allows users practice public speaking and in real time check the result of their performance. The therapist can track and monitor the values obtained in real time and change the environment conditions to send different stimuli to users so that they can adapt their attitudes to new situations.

In real-time biofeedback mode, the VR environment has an adaptive behavior according to the biodata collected during the presentation. Through its portability, the system allows personalized individual training and sessions accompanied by a therapist who may or may not influence the virtual environment.

Acknowledgments

CITIC is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS).

Bibliography

A. G. Ambit and M. R. O. Pandayan. Can't say a word with you: Glossophobia. *SSRN*, 2020.

P. L. Anderson, E. Zimand, L. F. Hodges, and B. O. Rothbaum. Cognitive behavioral therapy for public-speaking anxiety using virtual reality for exposure. *Depression and anxiety*, 22(3): 156–158, 2005.

G. Bodie. A racing heart, rattling knees, and ruminative thoughts: Defining, explaining, and treating public speaking anxiety. *Communication Education - COMMUN EDUC*, 59:70–105, 2010.

S. A. Dansieh, E. Owusu, and A. Seidu. Glossophobia: The fear of public speaking in esl students in ghana. *Language Teaching*, 1:22–35, 2021.

T. Lillis. Communicative competence. *Encyclopedia of Language & Linguistics*, 2005.

P. Lindner, J. Dagöö, W. Hamilton, A. Miloff, G. Andersson, A. Schill, and P. Carlbring. Virtual reality exposure therapy for public speaking anxiety in routine care: a single-subject effectiveness trial. *Cognitive Behaviour Therapy*, 50(1):67–87, 2021. PMID: 32870126.

N. Nahliah and F. Rahman. Glossophobia in training of speech. *ELS Journal on Interdisciplinary Studies in Humanities*, 1(1):28–36, 2018.

Y. Ningsih, N. Mariyati, and N. Susanti. Student's public speaking ability through presentation task. *Journal of English in Academic and Professional Communication*, 9:81–89, 07 2023.

OECD and I. L. Organization. Better use of skills in the workplace, 2017.

S. Ortiz, M. Deshais, R. Miltenberger, and K. Reeve. Decreasing nervous habits during public speaking: A component analysis of awareness training. *Journal of Applied Behavior Analysis*, 55, 09 2022.

An Intelligent IoT Wearable for Monitoring and Preventing Asthma Attacks

Pablo García-Rivada, Paula Fraga-Lamas, and Tiago M. Fernández-Caramés

Department of Computer Engineering, Faculty of Computer Science, Universidade da Coruña, 15071 A Coruña, Spain
Centro de Investigación CITIC, Universidade da Coruña, 15071 A Coruña, Spain
Correspondence: p.grivada@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: Asthma is a chronic disease that affects the respiratory tract, inflicting coughing, wheezing and difficulty for breathing. In order to mitigate these problems, this paper introduces an IoT wearable that is able to monitor the state of an asthma patient: breathing frequency, heart rate and oxygen saturation level. It also detects parameters that may worsen the situation of the patient: temperature, humidity and presence of gas or particles. The system also notifies the patient when he/she has to take his/her asthma medication. The value of the parameters is shared in real time by Bluetooth and historic values are stored in a database. The system uses color-based indications to calculate the risk level. With this information, the user can detect an asthma attack before it happens. To illustrate how the system operates, asthma attacks were simulated, showing that the system is able to determine that the attack is happening based on the specified input parameters.

1 Introduction

A wearable is an IoT device that is designed to be worn, like a smart watch, for different purposes like measuring data while providing wireless communications (Fernández-Caramés and Fraga-Lamas, 2023). In this paper, it is presented a wearable capable of measuring, showing and storing different parameters that can be crucial for asthma patients, warning them about threatening environments or irregular vital signs.

Asthma affects around 5% of people globally, and specially at young ages (Lancet, 2020). The more industrialized the country, the more asthmatics there are (Mancilla-Hernández et al., 2015). An asthma attack occurs when airways become inflamed and narrowed causing coughing, wheezing, tightness in the chest and less amount of air getting in and out the lungs. Additionally, the body produces mucus obstructing even more the airways (CDC, 2019).

There are some external factors that may impair patient breathing or trigger an asthma attack. Although such factors can vary between different people, some of the factors are: tobacco smoke, dust mites, atmospheric contamination, smoke from burning wood or grass, pets, cockroach allergens, mold, infections like flu or COVID, weather, sports which keep competitors constantly moving and even strong emotions (CDC, 2019).

2 Design and Implementation

The developed IoT system is able to measure the previously mentioned parameters through sensors. To notify the user when to take his/her medication, the system allows for configuring alarms that are notified by turning on an LED, which is turned off by pressing a button.

Real-time data are displayed on a smartphone by connecting it via Bluetooth Low Energy (BLE). The board also stores data on RAM (removing the oldest data when it runs out of memory) and can uploads them later opportunistically (i.e., once the ESP32 detects that an Internet connection is available) to a web application developed in Django, which stores all data on a database and provides an interface that shows them through a graph.

The wearable consists of an embedded board based on an ESP32 connected to a set of sensors using the I²C interface. The used sensor are: an RTC DS3231 is used to measure temperature and to configure alarms; a BME280 is used for measuring atmospheric pressure, temperature, relative humidity and altitude; a SEN0460 is used for measuring air quality (PM2.5, PM1.0 and PM10); a MAX30102 is to measure heart rate and oxygen levels on blood; MPL3115A2 is used for determining the breathing rate; and, finally, a CCS811 is used for detecting the presence of CO2 and TVOC (Total Volatile Organic Compounds).

Breathing frequency is calculated with the barometer MPL3115A2, potentially embedded in a smart mask that the user wears. For doing this, an algorithm was created to compare continuously the latest obtained value with the previous one (obtained roughly every 20 seconds), thus determining whether atmospheric pressure has increased or decreased. The number of times pressure has increased and decreased during a time interval indicates the breathing frequency.

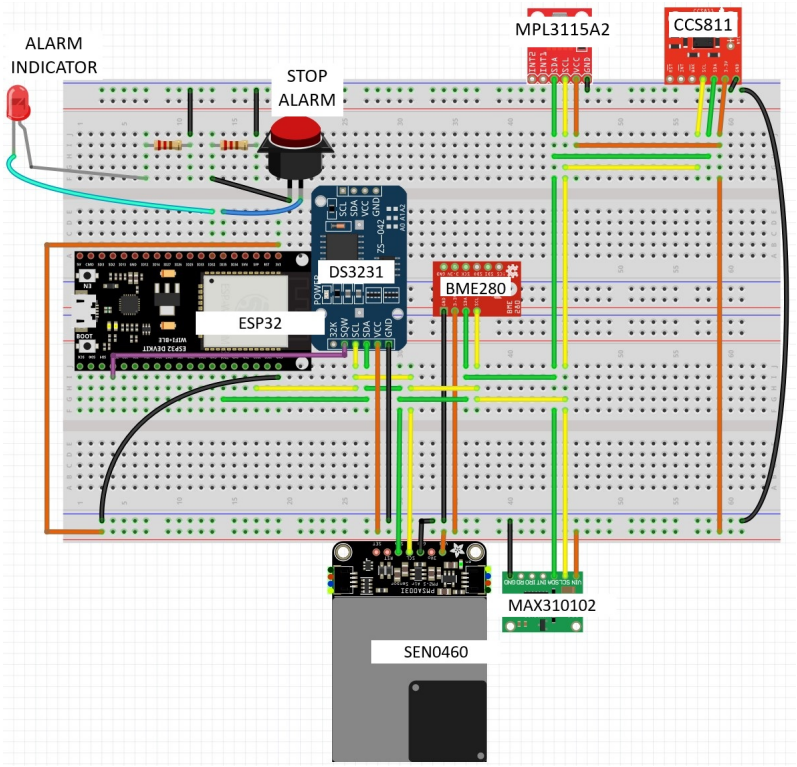
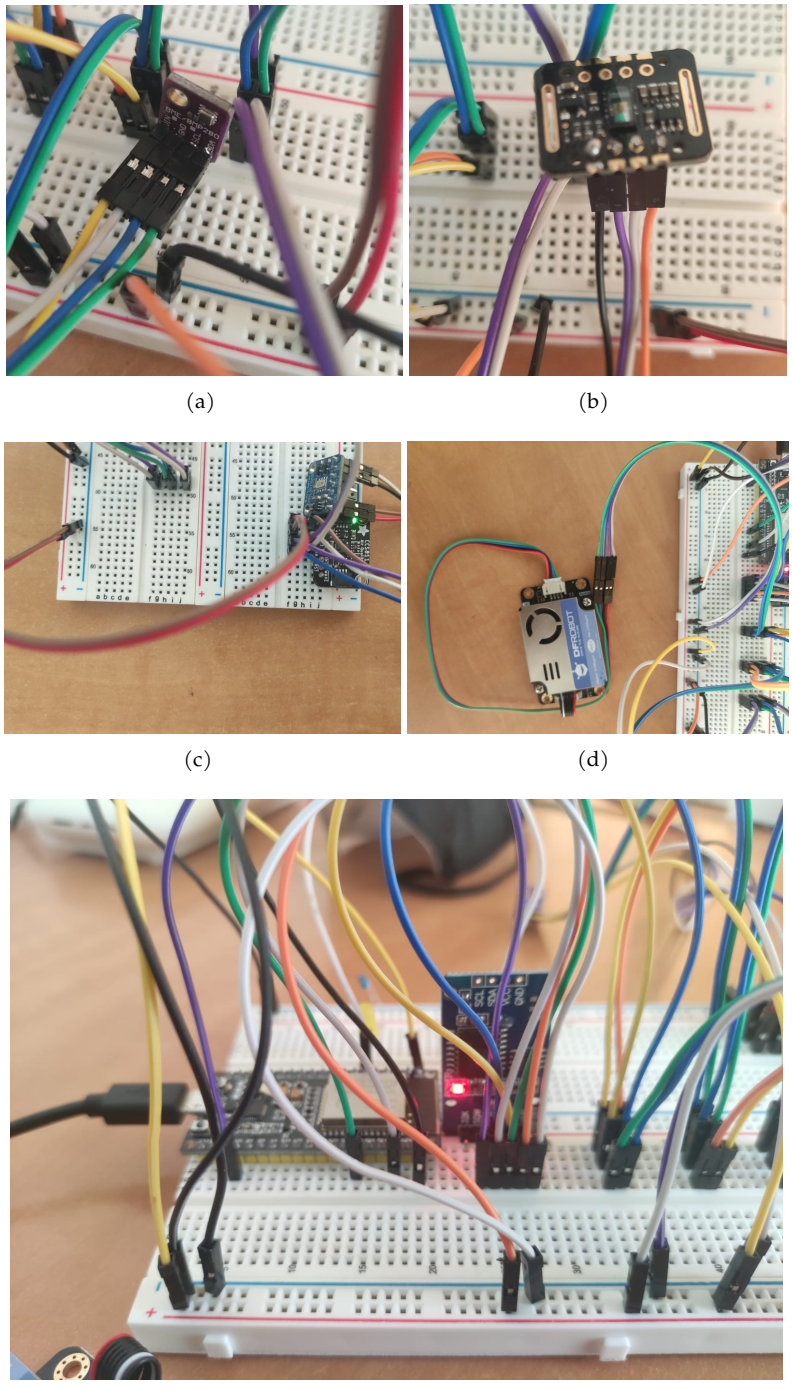


Figure 1: Connections and components of the prototype of the IoT wearable.

Table 1 shows the cost of the hardware used by the IoT wearable, including sensors, wires and other components (as of September 2023). Such a table does not include the cost of the final form-factor of the device. For instance, all the hardware described could be embedded, for example, in a smart mask. Figure 1 shows the internal components of the IoT wearable, while Figure 2 includes pictures of the different parts of the first working prototype. All sensors get power from the 3.3V pin of the microcontroller board, represented by an orange wire and use



(a) (b)

(c) (d)

(e)

Figure 2: (a) Barometer BME280, (b) Pulse Oximeter MAX30102, (c) Barometer for breathing rate MPL3115A2 & Air quality sensor CSS811, (d) Particle Counter SEN0460 and (e) DS3231 & ESP32.

the same ground (GND), represented as a black wire. Green and yellow wire are used for SDA and SCL connections, which are the connections used by I2C to share the communications bus with the microcontroller.

Table 1: Main components and cost of the IoT wearable.

Part	
Component	Price
Microcontroller Board	12€
Pulse oximeter Sensor	7€
Barometer Sensor	18€
RTC Module	8€
Breathing Frequency Sensor	15€
Particle Counter Sensor	36€
Air Quality Sensor	22€
External battery	14€
Resistors, LEDs, wires, etc.	19€
Total	151€

3 Results

After the development of the system, two scenarios were simulated:

- Detecting an asthma attack in real time. Using the vital signs parameters measured by the sensors, it is possible to detect when the user is suffering an asthma attack by detecting a noticeable increase in breathing rate even when the user is not doing physical activity (a regular value for breaths per minute is between 12 and 18 (MedlinePlus, 2008)). In Figure 3 it can be observed how the user starts with a normal breathing rate (around 18 breaths per minute) and then the asthma attack occurs, which increases up to 44 breaths per minute, as the beats per minute increase and the oxygen saturation in blood (SPO2) decreases from 95-100% to 85% (for this specific example, the SPO2 value has been simulated to illustrate the attack in a coherent way). Finally, the breaths per minute gradually decrease to 15 and the beats per minute and SPO2 return to their initial levels.
- Checking the environment via light signals. Besides the value for some parameters that indicate how dangerous the environment is, we use an heuristic value (score) which works like a traffic light, using green, yellow or red if the value is good, acceptable or bad for an asthma patient, respectively. For calculating this score, the system compares the parameter values to specific reference intervals: if the read value is between the reference values, it has a good score. For example, the reference value for temperature is between 15 and 25 Celsius degrees, but it can be easily configured to different scenarios. The color of the traffic light is internally represented by a number, so zero is a good value (green) and two is a bad value (red), while one is a value that is not optimal but it does not imply a dangerous situation for the patient. In a good scenario, most scores would be in green or yellow, while in a dangerous situation many red or yellow scores would be shown. As an example, Figure 4 shows the internal output of the system for different scenarios: Figure 40.4(a) illustrates a scenario where all scores are good or not bad, Figure 40.4(b) represents the outputs of a scenario where some parameters are bad for an asthmatic

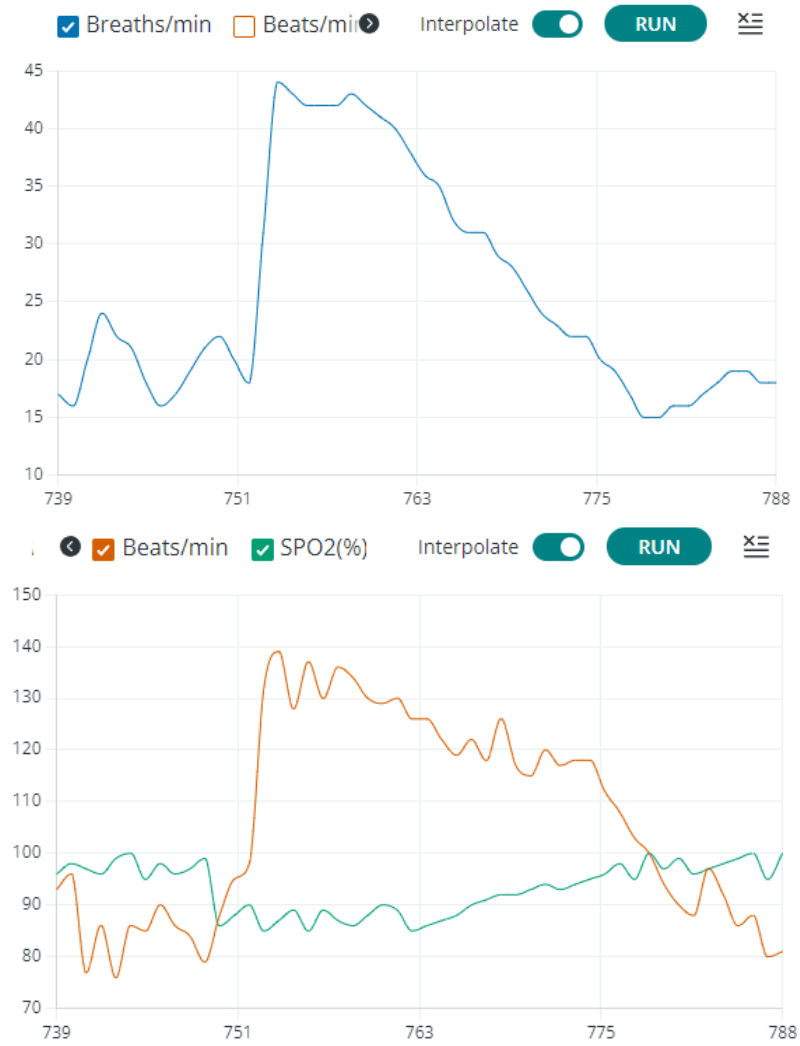


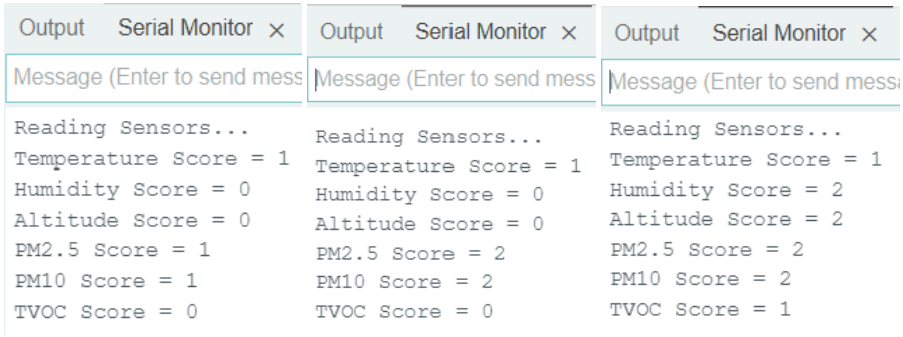
Figure 3: Breaths per minute, heartbeat and SPO2 during a simulated asthma attack.

and, finally, Figure 40.4(c) illustrates a scenario where most scores are dangerous for an asthmatic.

4 Conclusions

This paper has presented a prototype of an IoT wearable that is able to determine whether an asthma attack occurs and when the environment may trigger such an attack. The wearable makes use of multiple sensors to measure several important parameters for asthma patients, being able to monitor them in real time through a BLE device and to store them in its internal memory to dump them later opportunistically into a database when it is detected that an Internet connection is available.

The obtained results show that the system is able to perform the desired tasks adequately,



(a) (b) (c)

Figure 4: The system evaluates the situation using different scores for each parameter.

but the IoT wearable can still be considered as an early prototype. In the future, it can evolve towards a commercial system that integrates all the hardware into a smart mask. Moreover, the developed algorithm used for calculating the user's breathing rate might be refined to increase its accuracy.

Acknowledgment

This work has been funded by the Xunta de Galicia (by grant ED431C 2020/15), and by grants PID2020-118857RA-100 (ORBALLO) and TED2021-129433A-C22 (HELENE) funded by MCIN/AEI/10.13039/501100011033 and the European Union NextGenerationEU/PRTR. CITIC is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS).

Bibliography

CDC. Salud ambiental.contaminación del aire y salud respiratoria.asma, 2019. URL *https://www.cdc.gov/asthma/es/faqs.htm#*. Source: https://www.cdc.gov/nceh.

T. M. Fernández-Caramés and P. Fraga-Lamas. Design of a fog computing, blockchain and iot-based continuous glucose monitoring system for crowdsourcing mhealth. *Multidisciplinary Digital Publishing Institute Proceedings*, 4(1:37), 2023.

T. Lancet. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the global burden of disease study 2019. *The Lancet*, 396:1204–1222, 2020. URL *https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)30925-9/fulltext*.

E. Mancilla-Hernández, M. A. Medina-Ávalos, R. H. Barnica-Alvarado, D. Soto-Candia, and R. G.-V. andYahvéh Zecua-Nájera1. Prevalencia de asma y determinación de los síntomas como indicadores de riesgo. *Revista Alergia México*, 2015. URL *https://www.redalyc.org/articulo.oa?id=486755028006*.

MedlinePlus. Vital signs, 2008. URL *https://medlineplus.gov/ency/article/002341.htm*.

Reimagining Art: Virart, a Platform Using Virtual Reality for Creative Expression and Cultural Engagement

João Donga, João Azevedo, Marta Carrapa, and Paulo Veloso Gomes

School of Media Arts and Design, Polytechnic of Porto, Porto, Portugal
LabRP, Center for Rehabilitation Research, School of Health, Polytechnic of Porto, Porto, Portugal
Correspondence: jpd@esmad.ipp.pt

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: Virart is an immersive experience that showcase the potential of Virtual Reality in recreating artistic works. By leveraging the capabilities of VR, users can explore and interact with digital recreations of Júlio Maria dos Reis Pereira artistic creations. The VR environment enhances the sense of presence and enables viewers to explore the art from different angles, offering a deeper level of engagement compared to traditional mediums. By showcasing the possibilities of virtual reality in recreating artistic works, Virart opens new avenues for artists, art enthusiasts, and the general public to appreciate and engage with art in innovative ways. It demonstrates the potential of virtual reality technology in preserving and promoting cultural heritage, allowing art to be experienced beyond physical limitations.

1 Introduction

The fusion of art and technology has been a recurring theme throughout history, from the invention of the camera to the rise of digital art. However, virtual reality (VR) has taken this intersection to unprecedented heights. VR art galleries, immersive digital environments designed to showcase artworks and exhibitions, have opened up a new dimension in the art world. These digital spaces not only replicate the experience of visiting a physical gallery but also enhance it through interactivity and accessibility.

Virtual reality in the context of art has evolved rapidly over the past decade (Kim, 2016). Early experiments with VR art often involved simple 3D modeling and animation. However, advancements in hardware and software have enabled artists to create highly detailed and immersive virtual environments that rival physical galleries.

Artists have embraced VR as a medium for artistic expression. They can sculpt, paint, and manipulate digital materials in ways previously impossible, pushing the boundaries of creativity (Gräsler and Taplick, 2019), (Gong and Georgiev, 2020), (Graessler and Taplick, 2019). VR art provides a unique platform for exploring new artistic forms and experimenting with the multisensory aspects of artistic expression.

One of the most compelling aspects of VR art galleries is their ability to democratize access to art. Traditional art galleries are often limited by physical location, opening hours, and admission fees. VR galleries eliminate these barriers, allowing anyone with a VR headset and internet connection to explore art from around the world, 24/7 (Shehade and Stylianou-Lambert, 2020), (Lee et al., 2020), (Carrozzino and Bergamasco, 2010).

Furthermore, VR galleries can create inclusive environments that accommodate people with disabilities. Through customizable interfaces and assistive technologies, individuals who may

face challenges in visiting physical galleries can engage with art on their terms (Drigas et al., 2009).

Virart is a project developed in an academic context with the aim of better understanding the relationship and interdependence between the artistic, technological and conceptual areas of multimedia and painting, allowing an in-depth exploration of the selected tools and a growing capacity for project development.

An artistic reconstruction experience was developed in which the user finds himself in a gallery with the works of artist Júlio Maria dos Reis Pereira. Presents four works of art (Figure 1) by Júlio Maria dos Reis Pereira offering users the opportunity to explore and admire the paintings in a three-dimensional and interactive way.

This project assumes a multidisciplinary and transdisciplinary nature, and aims to offer a different experience, taking the user to an unknown context, through the use of technological devices that allow a new model of interaction with works of art.

Figure 1: Original Painting and 3D Recreation

2 Materials and Methods

The project was developed in the following steps:

1. Definition of Concept and Goals: The creation of an artistic reconstruction experience using works by Júlio Maria dos Reis Pereira, creating an immersive experience and demonstrating the possibilities of 3D art.
2. Research: Study of Júlio Pereira's artistic work and choice of works of art for 3D recreation
3. Study of existing VR museums to understand best practices and get inspiration.
4. Content Planning: Creation of a detailed content plan outlining the space, works of art, light, sound and interactive elements.
5. Choosing VR Technology: Use of the Oculus Quest 2 as they offer a combination of image quality, precise movement tracking and interactive features.
6. Develop 3D Models and Assets: Creation of models using Maya. Development of textures, animations, and audiovisual elements to enhance the experience.
7. Design the VR Environment: Design the virtual museum space, including layout, lighting, and navigation. Consideration as visitors will move through the exhibits and interact with the content.
8. Programming and Development: Use of game engine Unity.
9. Implementation of interactive elements, navigation systems, and user interfaces.
10. Content Integration: Integration of 3D models, textures, audio into the VR environment.

Figure 2: View of the art work

Figure 3: View of the gallery Dark Mode

Figure 4: View of the gallery

3 Results

When exploring perspectives as different ways of seeing a reality, it is similar to observing a certain event from several different angles. Through this approach, the Virtual Reality art gallery provided users with a unique immersion, allowing them to explore and appreciate works of art in diverse ways. Define a model that can serve as a basis for creating interactive virtual exhibitions, collaborations with other artists and cultural institutions, and even inspire new approaches in the field of art and technology. Furthermore, the continuity of the project may allow the expansion of the collection of works by Júlio Maria dos Reis Pereira and the creation of new immersive experiences, further enriching the artistic legacy.

4 Conclusion

In future work we plan to define a model that can serve as a basis for creating interactive virtual exhibitions, collaborations with other artists and cultural institutions, and even inspire new approaches in the field of art and technology. Furthermore, the continuity of the project may allow the expansion of the collection of works by Júlio Maria dos Reis Pereira and the creation of new immersive experiences, further enriching the artistic legacy.

When exploring perspectives as different ways of seeing a reality, it is similar to observing a certain event from several different angles. Through this approach, the Virtual Reality art gallery provided users with a unique immersion, allowing them to explore and appreciate works of art in different ways.

In conclusion, Virart has proven to be a dynamic and innovative platform, capable of broadening artistic horizons and stimulating creativity. By exploring different perspectives and offering a unique experience, it redefined the way works of art can be appreciated and interpreted, highlighting the power of technology to enrich and revolutionize the world of art.

Bibliography

M. Carrozzino and M. Bergamasco. Beyond virtual museums: Experiencing immersive virtual reality in real museums. *Journal of Cultural Heritage*, 11(4):452–458, 2010. URL *https:*

//*www.sciencedirect.com/science/article/pii/S1296207410000543*.

A. Drigas, L. Koukianakis, and J. Glentzes. An e-culture – e-museums environment for common citizens and disabled individuals. *International Journal of Digital Culture and Electronic Tourism*, 1(4):267–279, 2009.

Z. Gong and G. Georgiev. Literature review: Existing methods using vr to enhance creativity, 09 2020.

I. Graessler and P. Taplick. Supporting creativity with virtual reality technology. *Proceedings of the Design Society: International Conference on Engineering Design*, 1(1):2011–2020, 2019.

I. Gräsler and P. Taplick. Architecture of a virtual reality-based tool for the support of creativity, 04 2019.

B. Kim. Virtual reality as an artistic medium: A study on creative projects using contemporary head-mounted displays. Master's thesis, Aalto University. School of Arts, Design and Architecture, 2016. URL *http://urn.fi/URN:NBN:fi:aalto-201612085686*.

H. Lee, T. H. Jung, M. tom Dieck, and N. Chung. Experiencing immersive virtual reality in museums. *Information & Management*, 57(5):103229, 2020. URL *https://www.sciencedirect.com/science/article/pii/S0378720618310280*.

M. Shehade and T. Stylianou-Lambert. Virtual reality in museums: Exploring the experiences of museum professionals. *Applied Sciences*, 10(11), 2020.

The Use of Artificial Intelligence in Interactive Virtual Reality Adaptive Environments with Real-Time Biofeedback Applied to Phobias Psychotherapy

Paulo Veloso Gomes, Vítor J. Sá, João Donga, António Marques, Bárbara Gomes, Raquel Simões de Almeida, and Javier Pereira

LabRP, Center for Rehabilitation Research, School of Health, Polytechnic of Porto, Porto, Portugal
Centro ALGORITMI, University of Minho, Guimarães, Portugal
Universidade Católica Portuguesa, Braga, Portugal
LabRP, Center for Rehabilitation Research, School of Media Arts and Design, Polytechnic of Porto, Porto, Portugal
CITIC Research Center, University of A Coruña, A Coruña, Spain
Correspondence: velosogomes@ess.ipp.pt

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: Exposure therapy is a type of psychotherapy where the patient is gradually exposed to a fear situation. Patients may present different degrees of phobia, and the degree of phobia of each patient changes throughout the treatment. The use of Interactive Virtual Reality Adaptive Environments, where the interaction is personalized through real-time biofeedback mechanisms, allows the environment to adapt to the patients and their evolution throughout the treatment. The Artificial Intelligence affective algorithms continuously monitor the patient's behavioral data and physiological responses to adjust the intensity and number of the stimuli. This real-time adaptation aims to personalize and optimize the exposure process, gradually desensitizing patients to their fears.

1 Introduction

Phobias are characterized as irrational fears about situations, creatures, places or objects, affecting around 10% of the world's population at some point in their lives (Garcia, 2017), (Eaton et al., 2018), (Samra and Abdijadid, 2023). These disorders are identified by anxiety and/or fear in certain situations, which pose little or no real danger.

The American Psychiatric Association 2013 recognizes different types of phobias: (1) Agoraphobia - characterized by symptoms of anxiety and panic in situations where the person perceives their environment to be unsafe, with no easy way to escape. These situations may include, for example, going to shopping malls, busy streets, or public transportation; (2) Social Phobia - generally described as social anxiety and causing distress and impaired ability that negatively interferes with the routines of one's daily life. someone. One of the most common is the fear of public speaking; (3) Specific Phobia - consists of an intense fear and anxiety about some specific trigger, such as heights (acrophobia), spiders (arachnophobia), or small, enclosed spaces (claustrophobia).

Apart from pharmacological treatment, the most common form of treatment for phobias is in vivo exposure (Noordik et al., 2010), (Thng et al., 2020). This method consists of confronting the patient for an extended period of time (e.g. 60 minute sessions) with their feared stimulus until the distress subsides. The aim is to change the patient's response to the object or situation that is causing the irrational fear. Gradual and repeated exposure to the anxiogenic source and related thoughts, feelings and sensations can help control the patient's anxiety and fear (Raeder et al., 2020). However, exposure therapy has the negative factor of lack of environment full control, which can lead to negative reactions from the individual being treated.

An innovative strategy to intervene in this problem is the use of Virtual Reality (VR), which consists of a computer-generated environment to simulate the real world through an immersive experience (Bell et al., 2020). VR uses software applications, such as games or simulators, to create a virtual experience in a therapeutic environment, thus making it possible to recreate real everyday situations that cause anxiety in individuals undergoing treatment, in a therapeutically controlled environment (Rimer et al., 2021). Thus, it is possible for the patient to carry out exposure sessions to the feared difficulties/situations, exploring virtual scenarios that trigger emotions, physiological sensations, thoughts, and behaviors similar to real situations in a controlled environment. It is an intervention that enables individuals to find an additional self-regulation, ideal for themselves and their symptomatology, using inhibitory learning, i.e. control of impulsive responses (Albakri et al., 2022), (Reeves et al., 2022).

The literature has shown that VR can be as or more effective as in vivo exposure for the treatment of most in vivo exposure for the treatment of most phobias, with respect to the level of measured anxiety and avoidance (Wechsler et al., 2019). In addition to environmental control, virtual reality also allows for a high degree of confidentiality, since the exposure is done inside a room and there is no risk of potentiating negative reactions, either by the individuals or by possible observers. This intervention also enables the therapist to track the images the individual is viewing, allowing them to understand which specific images more accurately have caused the increases in anxiety and then therapeutically work through the response with the patient.

As there is no standardized protocol for treating phobias, treatment should always be tailored to each person to achieve more robust long-term results (Bergsnev and Sánchez Laws, 2022), (Thng et al., 2020).

The aim of this work is to apply Artificial Intelligence (AI) algorithms in the use of Interactive Virtual Reality Adaptive Environments, that adapt them self to the patients and their evolution throughout the treatment, where the interaction is personalized through real-time biofeedback mechanisms.

2 Methods

The process of building immersive environments is complex and involve a strong technological component and an important human component. The research was conducted using a sociotechnical approach based on Actor-Network Methodology, which focuses on the development of systems whose operation is based on the interaction between human and technological elements. These elements interact with each other during the construction process with the aim of obtaining a controlled system, which reproduces the characteristics necessary for an effective therapeutic exposure. However, this interaction remains and is fundamental throughout the treatment.

The research applies the state-of-the-art VR technology and Artificial Intelligence applications to create realistic and interactive immersive adaptative environments, as user-centered design methodologies involving user feedback and the necessary iterations to ensure the authenticity and effectiveness of the virtual environments.

The development of the algorithms considered the possibility of their use in three different modes of system operation:

- Autonomous Mode Strategy: For use independently by the patient or with the supervision of the therapist.
- Controlled Mode Strategy: For use with therapist monitoring and intervention.
- Real-time Biofeedback Mode Strategy: For use with therapist monitoring and supervision.

3 Results

A biofeedback system collects in real-time unimodal data from skin conductivity, respiratory and heart rate user devices (Figure 1). The system collects also the unimodal data from Headset movements and command buttons. These unimodal data are recorded and graded separately.

The algorithm analyzes the unimodal data collected during exposure from the patient's voluntary reactions, that reflect a conscious response (head movements, use of commands) and involuntary reactions which results from physical reflexes and biological changes (skin conductivity, respiratory and heart rate), and converts it into multimodal data to adapt the environment according to the Biofeedback obtained, readjusting the frequency, intensity and duration of the stimuli generated.

Three operating modes of the system were designed. Each of them has specific characteristics that allow their use in different therapeutic strategies.

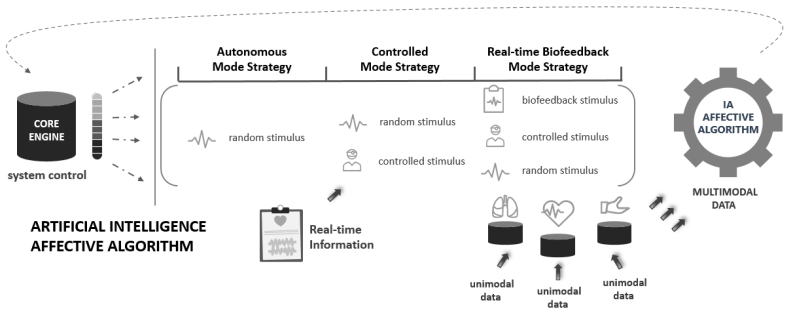


Figure 1: Artificial intelligence affective algorithm for virtual reality biofeedback systems.

3.1 Autonomous Mode Strategy

This mode allows the system to be used individually by the patient. Given its portability, it can also be used at home. The system offers different sections with an increasing degree of difficulty, the user has the autonomy to choose the session they want to carry out. By measuring the user's level of comfort, the system adapts the frequency, duration and intensity of stimuli to their emotional state. Users can repeat the same session as many times as they wish, however, the algorithm generates different sessions with the same degree of difficulty to ensure that repetition is not monotonous or predictable. This mode can also be used under the supervision of the therapist.

3.2 Controlled Mode Strategy

For use with therapist monitoring and intervention, this mode allows full customization of exposure. The therapist, using a tablet, visually monitors what the patient is seeing. Observing the patient's behavior and reactions to stimuli, the therapist can interact with the system, controlling the frequency, intensity and duration of the stimuli.

3.3 Real-time Biofeedback Mode Strategy

In this mode the therapist assumes the role of monitoring and supervision. The system adapts to the user's behavior using the AI algorithm. The AI affective algorithm uses multimodal data to recognize the type and intensity of user emotions during exposition and optimize the experience through the continuous adaptation of stimuli to his own emotional state.

4 Conclusion

The use of VR devices has some limitations as it is a recent technology, consequently users are not yet familiar with how it works and how to resolve minor technical problems that may arise during its use.

To make viable the use of Interactive Virtual Reality Adaptive Environments in Phobias Psychotherapy it was considered the use of AI affective algorithms in different real contexts. The complexity of the algorithms can be adapted to the needs of customizing adaptive immersive environments according to therapeutic objectives, taking into account technological limitations or restrictions on the use of certain devices.

The AI algorithms used guarantee great flexibility in managing the data that indicates user feedback. This feature allows data to be collected independently, which makes it possible to increase or decrease the number of devices and types of data collected according to the portability desired or the degree of accuracy desired. The concept of interactive VR adaptive environment with real-time biofeedback allows its use throughout the entire treatment process. The flexibility adopted in the development and the modular architecture of the AI affective algorithms allows them to be applied in the therapy of different types of phobias.

Acknowledgments

CITIC is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS).

Bibliography

G. Albakri, R. Bouaziz, W. Alharthi, S. Kammoun, M. Al-Sarem, F. Saeed, and M. Hadwan. Phobia exposure therapy using virtual and augmented reality: A systematic review. applied sciences. *Dialogues in clinical neuroscience*, 3(12):1672, 2022.

A. P. Association. *American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*. American Psychiatric Association, Arlington, VA, 2013.

I. H. Bell, J. Nicholas, M. Alvarez-Jimenez, A. Thompson, and L. Valmaggia. Virtual reality as a clinical tool in mental health research and practice. *Dialogues in clinical neuroscience*, 2(22): 169–177, 2020.

K. Bergsnev and A. L. Sánchez Laws. Personalizing virtual reality for the research and treatment of fearrelated disorders: A mini review. *Frontiers in Virtual Reality*, 3, 2022.

W. W. Eaton, O. J. Bienvenu, and B. Miloyan. Specific phobias. *The lancet. Psychiatry*, 8(5): 678–686, 2018.

R. Garcia. Neurobiology of fear and specific phobias. learning & memory. *Cold Spring Harbor, N.Y.*, 9(24):462–471, 2017.

E. Noordik, J. van der Klink, E. Klingen, K. Nieuwenhuijsen, and F. J. H. Dijk. Exposure-in-vivo containing interventions to improve work functioning of workers with anxiety disorder: a systematic review. *BMC Public Health*, 598(10), 2010.

F. Raeder, C. J. Merz, J. Margraf, and A. Zlomuzica. The association between fear extinction, the ability to accomplish exposure and exposure therapy outcome in specific phobia. *Scientific reports*, 1(10):4288, 2020.

R. Reeves, D. Curran, A. Gleeson, and D. Hanna. A meta-analysis of the efficacy of virtual reality and in vivo exposure therapy as psychological interventions for public speaking anxiety. *Behavior Modification*, 4(46):937–965, 2022.

E. Rimer, L. V. Husby, and S. Solem. Virtual reality exposure therapy for fear of heights: Clinicians' attitudes become more positive after trying vret. *Frontiers in psychology*, 12, 2021.

C. K. Samra and S. Abdijadid. Specific phobia. *In: StatPearls* [*Internet*]. *Treasure Island* (*FL*): *StatPearls Publishing; 2023 Jan-*, 2023. Available from: https://www.ncbi.nlm.nih.gov/books/NBK499923.

C. E. W. Thng, N. S. J. Lim-Ashworth, B. Z. Q. Poh, and C. G. Lim. Recent developments in the intervention of specific phobia among adults: a rapid review. *F1000Research*, 9(195):F1000 Faculty Rev–195, 2020.

T. F. Wechsler, F. Kümpers, and A. Mühlberger. Inferiority or even superiority of virtual reality exposure therapy in phobias?-a systematic review and quantitative meta-analysis on randomized controlled trials specifically comparing the efficacy of virtual reality exposure to gold standard in vivo exposure in agoraphobia, specific phobia, and social phobia. *Frontiers in psychology*, 10:1758, 2019.

Simulation of Virus Propagation and Acceptance of Socio-Sanitary Measures Through an Intelligent Model

Daniel J. García-Paz, Amparo Alonso-Betanzos, Bertha Guijarro-Berdiñas, and Alejandro Rodríguez-Arias

Universidade da Coruña, LIDIA, Faculty of Computer Science, 15071 A Coruña, Spain
Universidade da Coruña, Centro de Investigación CITIC, 15071 A Coruña, Spain
Correspondence: alejandro.rodriguez.arias@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: During the most critical moments of the SARS-COV-2 pandemic, various containment measures were enacted to hinder the virus's spread and mitigate its impact. This work focuses on studying the impact of the population's adherence level to socio-sanitary measures on the virus's spread, aiming to better understand its relevance in crisis situations. To achieve this goal, we use an agent-based model (ABM) that incorporates a special type of agent that represents social networks, for example, twitter, to analyze the influence of social networks on the agents' decision-making.

Internally, our model relies on two models that allow for simulation development. On the one hand, an epidemiological model based on an adaptation of the SIR model allows us to simulate the spread of the virus. On the other hand, a decision-making model is responsible for analyzing the levels of acceptance of containment measures by citizens and allows simulation of interactions between agents. On this basis, Twitter has been incorporated as a critical node, which allows information to be extracted about the opinions of the agents and how these affect the population's adherence to socio-sanitary measures. This information is obtained thanks to the application of sentiment analysis techniques on a set of tweets related to COVID-19.
As a result, a useful tool was obtained for policy makers to simulate the psycho-social behaviour of citizens in the face of different restrictive measures in order to evaluate their effectiveness.

1 Introduction

The SARS-CoV-2 crisis is still seen as a period of uncertainty and significant social change, even three years later. It was strongly marked by restrictions and the imposition of socio-sanitary measures on the citizens in many countries, aiming to mitigate the spread of infections and consequently, the advancement of the virus causing the pandemic.

Over the past three years, various studies such as Saez et al. (2020) or Aleta and Moreno (2020) have indeed demonstrated that all these measures positively contributed to the fight against COVID-19. However, despite their widespread implementation, it is known that these measures have been extremely controversial due to social movements that denied their effectiveness or even the existence of the virus itself, and consequently, refused to adhere to them. This raises the possibility of questioning the impact that individual actions can have on the spread of a virus or, in other words, whether the fact that an individual or a minority

refuses to adhere to the measures has a real effect on their effectiveness for the entire population.

In this context, social networks becomes particularly relevant as a source of numerous discussions in which both critics and supporters of socio-sanitary measures have used them to express their opinions and, at times, to try to persuade others of their stance. With this in mind, it's worth asking: Beyond the merit of socio-sanitary measures, what is the impact of these discussions on social networks on the population? Do they affect the effectiveness of the measures? Do they instill distrust in the measures among the public?

The search for answers to these questions indeed places us in a complex situation because the study of human behavior, the factors influencing it, and how all of this affects the evolution of an epidemic is not a trivial matter. This is where the use of computational models becomes particularly relevant, allowing us to study the behavior of a population in a controlled virtual environment. Agent-based models represent systems by simulating the actions of their individual components (agents, which act autonomously), and the interaction between them and their environment, allowing complex processes and systems to be solved. A key feature of ABM is that it allows us to tackle or study problems related to the concept of emergent phenomena. This means examining system dynamics that appear from the behavior and interaction of the individuals that make up the system, thus allowing the study of questions related to how the global behavior of society links to individual behavior. As a result, ABMs have been widely used in various fields, including epidemiology, Amouroux et al. (2010), and the social sciences, Groff et al. (2019).

This work proposes the application of an agent-based model to study how the effectiveness of socio-sanitary measures is influenced by the population's behaviour, which will be shaped by the information they receive through social networks. To achieve this, the city of A Coruña (Spain) will be used as the geographical virtual environment, based on Geographic Information System (GIS) data, and real socio-demographic data from the city will be used to populate the agents. These agents will exhibit certain behaviors that allow them to satisfy their individual needs, and these behaviors will be influenced by the information they receive through social networks. To accomplish this, a decision-making model (see section 2.3) will be implemented to simulate social behaviors, the acceptance of socio-sanitary measures, and communication among the agents. For modeling the virus's spread, an epidemiological model based on the SEIRD model will be implemented (see section 2.2).

With this foundation established, our aim in this work is to simulate the spread of information to agents through social networks, represented as special agents in the model. These agents will be fed by a set of tweets from the 6-month period to be simulated, between June and November 2021, which will be processed using sentiment analysis techniques to gain an understanding of the opinions that socio-sanitary measures elicit in the population and how this opinion impacts the behaviour of the agents and their acquaintances.

2 Proposed model

As mentioned before, the model relies on 4 pillars: the environments, the epidemiological model, the decision-making model, and the critical nodes representing social networks.

2.1 Enviroment

The environment (see figure 1) has been developed based on GIS data from the city of A Coruña, which has been used to implement both the city's structure based on its census zones and the location of leisure establishments, supermarkets, government buildings, or offices,

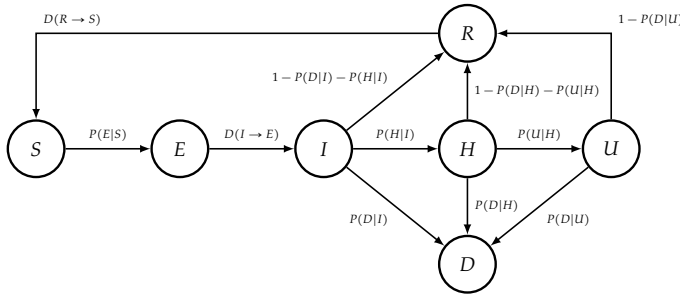
among others.

Figure 1: Representation of the city of A Coruña in the proposed model.

Influence network

Another fundamental part of the environment in which agents interact is the existence of networks of agents with whom they can interact. For the creation of influence networks, two different subgroups of agent contacts have been established. These groups will consist of two networks, on one hand, a network of friends composed of a random set of agents in the environment who are within a similar age range (maximum of 5 years difference) as the agent, furthermore, there is the possibility of adding random friendship relationships between agents. On the other hand, the citizen will have a social circle, Hamill and Gilbert (2009), a set of individuals who will be part of their influence network due to geographic proximity.

2.2 Epidemiological model



where:

$P(Y|X)$: Probability that an individual in state X transitions to state Y.

$D(X \to Y)$: Days that an individual in state X takes to move to state Y.

Figure 2: SEIRDS adaptation

As an epidemiological model, an adaptation of the SEIRDS model has been chosen (see Figure 2), an evolution of the SIR model proposed by Kermack and McKendrick (1927). The SEIRDS model addresses the limitations of the SIR model, which only considers the existence of three possible states (susceptible, infected, and recovered), by adding a latency period, the

possibility of reinfection, and the lethality of the disease.

Regarding this model, our adaptation takes into account two additional assumptions: the hospitalization of infected individuals and the admission to the ICU of hospitalized individuals, as the probability of death increases in each of these cases, reducing the probability of recovery.

2.3 HUMAT

Decision-making and agent interaction are crucial aspects in the development of this system. To achieve this, we have chosen to implement an HUMAT architecture, Antosz et al. (2019), which allows us not only to implement a decision-making model based on the individual needs of the agents but also to create networks of relationships among agents that enable the spread of information and the influence of some agents over others.

The HUMAT architecture addresses the problem of decision-making in a five-step cycle (see Figure 3) after initialization:

- **Evaluate:** The cycle starts with the evaluation of the two alternatives presented to the agent to satisfy their needs: 1) accepting the measures or 2) rejecting the measures.
- **Choice:** The agent selects its behavior based on its evaluation.
- **Action:** The agent acts accordingly, which will impact the probabilities of infection and communications with other agents in its circle of influence.
- **Experimentation:** The effects resulting from the agent's decision-making emerge.
- **Update:** The agent's needs are updated based on its experiences.

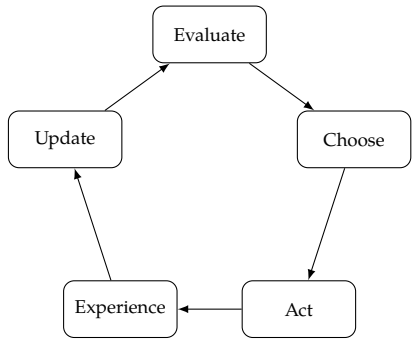


Figure 3: Decision-making strategy

2.4 Critical nodes

In our model, social network information is propagated through critical nodes, representing specific agents in the model capable of significantly altering agent behavior by establishing communication with them. In our case, these critical nodes represent social networks, specifically Twitter, which will influence agent behavior through the dissemination of tweets that have been previously processed using natural language processing techniques (see section 3). For the propagation of information, the critical node sends between 4 and 6 tweets daily to each of the users of the social network. These communications will involve 2 types of tweets (in favor or against the measures), which will modify the satisfaction of the individual needs of the agents.

3 Natural language processing

Sentiment analysis through Twitter data has been a topic of study since the beginning of the pandemic, being part of various studies such as Lopez and Gallemore (2021), Kaur et al. (2020), or Dubey (2020). In this line, the main objective of this study is to analyze whether this approach can be used to study the impact that information circulating on social networks has on the public opinion and, consequently, on the acceptance of socio-sanitary measures. For this purpose, it is necessary to obtain information from social networks, in our case, Twitter, and then process it to extract relevant information.

To obtain this information, the Twitter API has been used, which allows the collection of tweets through various filters, providing information related to a specific topic, user, or geographic location. In our case, we have chosen to collect tweets about Covid-19 within the Spanish-speaking community, as it is generally assumed that this community will post tweets that will reach our population. The result has been a set of 4,164 tweets spread over 24 months, from March 1, 2020, to February 28, 2022, of which we will use a subset corresponding to the 6-month period from June to November 2021, which coincides with a surge in COVID-19 infections in the city, according to data from the Galician Health Service (SERGAS).

The processing of tweets is carried out in two stages. First, a preprocessing step is performed because tweets, by their nature, often contain elements that do not provide information or cannot be processed by the language model. In this step, hashtags and URLs have been removed from the tweets, and emojis have been replaced with explanatory texts for each one.

On this basis, for the tweet processing, "pysentimiento", Pérez et al. (2021), has been used, a natural language processing framework designed for Twitter and based on RoBERTa, Liu et al. (2019), a pre-trained language model based on transformers designed for natural language processing tasks like sentiment analysis. This model allows us to obtain a series of values from a tweet that indicate its polarity, i.e., whether the tweet is positive, negative, or neutral (see figure 4). These polarity values are used for the propagation of information through the critical node, altering the satisfaction of individual agent needs, and consequently, their behavior. This allows us to study the relevance of these types of messages, both for Twitter users themselves and for the rest of the population.

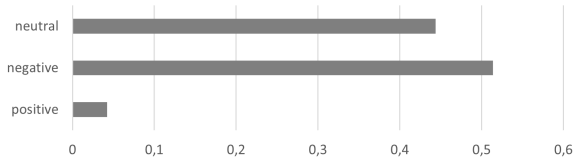


Figure 4: Percentages of tweets polarization.

4 Simulations

The model executions have been carried out with a total population of 215,000 agents, corresponding to the adult population of the city of A Coruña, of which 300 are infected at the beginning of the simulation And it will take place between June and November 2021. These agents are distributed across 187 census zones and move through approximately 4,000 different locations, including leisure establishments, supermarkets, government buildings, offices, among others.

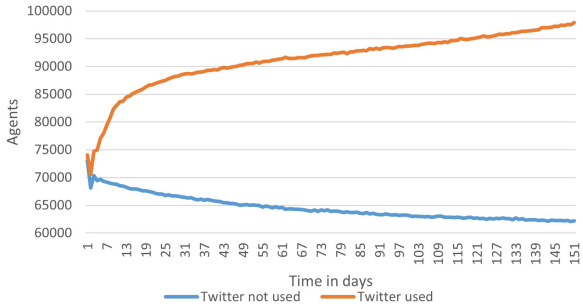
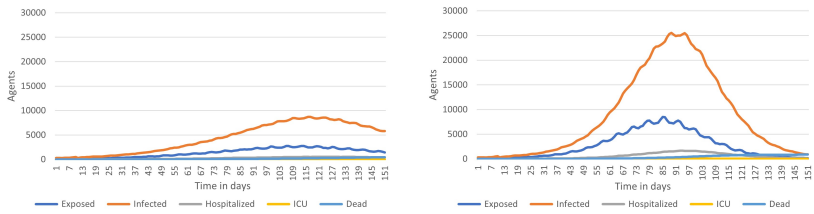


Figure 5: Reject behaviour evolution comparation

With the aim of observing the impact of information circulating on social networks on the public opinion, two types of model executions have been conducted, one in which critical nodes were not included and another in which they were included (see Figure 5).

Regarding the influence of critical nodes, we can observe that indeed, the information circulating on social networks has a significant impact on the public opinion. This translates into an increase in the number of agents rejecting socio-sanitary measures and, consequently, an increase in the number of infected agents (See figure 6). This is because, as can be seen in Figure 4, the information obtained on Twitter has a significant bias towards negativity, with positive polarization being around 5%. This leads to an increase in the rejection of socio-sanitary measures.



(a) SEIRD evolution without Twitter (b) SEIRD evolution with Twitter

Figure 6: SEIRD evolution comparation

5 Conclusions

Based on the results of the simulations, it can be concluded that the proposed model is capable of analyzing the variation in the levels of acceptance of socio-sanitary measures by the population and how this variation affects the spread of the virus. It is also evident that the addition of critical nodes to the model correctly alters the behavior of the agents, allowing us to conclude that the information circulating in the model is having the intended impact.

Furthermore, it can be observed that the information distributed by the critical nodes has had a significant impact on the decision-making of the agents. Indeed, there is a direct correlation between the number of agents rejecting socio-sanitary measures and the number

of infected agents. Therefore, we can conclude that both the decision-making model, the epidemiological model, and the critical nodes exhibit the expected behavior.

However, it has become evident that the sentiment analysis technique used is not capable of correctly classifying tweets. This is because the negative polarization does not directly imply a rejection of socio-sanitary measures. It has been observed that the technique used cannot differentiate between tweets expressing a negative opinion about the measures and tweets expressing a negative opinion about other questions. Therefore, it would be interesting to consider alternative techniques that allow us to address this problem and achieve accurate classification between these two types of tweets.

Acknowledgments

CITIC is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS).

Bibliography

A. Aleta and Y. Moreno. Evaluation of the potential incidence of covid-19 and effectiveness of containment measures in spain: a data-driven approach. *BMC medicine*, 18(1):1–12, 2020.

E. Amouroux, P. Taillandier, and A. Drogoul. Complex environment representation in epidemiology abm: application on h5n1 propagation. *Journal of science and technology*, pages 13–25, 2010.

P. Antosz, W. Jager, G. Polhill, D. Salt, A. Alonso-Betanzos, N. Sánchez-Maroño, B. Guijarro-Berdiñas, and A. Rodríguez. Simulation model implementing different relevant layers of social innovation, human choice behaviour and habitual structures. *SMARTEES Deliverable*, 2019.

A. D. Dubey. Twitter sentiment analysis during covid-19 outbreak. *Available at SSRN 3572023*, 2020.

E. R. Groff, S. D. Johnson, and A. Thornton. State of the art in agent-based modeling of urban crime: An overview. *Journal of Quantitative Criminology*, 35:155–193, 2019.

L. Hamill and G. Gilbert. Social circles: A simple structure for agent-based social network models. *Journal of Artificial Societies and Social Simulation*, 12(2), 2009.

S. Kaur, P. Kaul, and P. M. Zadeh. Monitoring the dynamics of emotions during covid-19 using twitter data. *Procedia Computer Science*, 177:423–430, 2020.

W. O. Kermack and A. G. McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, 115(772):700–721, 1927.

Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*, 2019.

C. E. Lopez and C. Gallemore. An augmented multilingual twitter dataset for studying the covid-19 infodemic. *Social Network Analysis and Mining*, 11(1):102, 2021.

J. M. Pérez, J. C. Giudici, and F. Luque. pysentimiento: A python toolkit for sentiment analysis and socialnlp tasks. *arXiv preprint arXiv:2106.09462*, 2021.

M. Saez, A. Tobias, D. Varga, and M. A. Barceló. Effectiveness of the measures to flatten the epidemic curve of covid-19. the case of spain. *Science of The Total Environment*, 727:138761, 2020. URL *https://www.sciencedirect.com/science/article/pii/S0048969720322786*.

Re-Identification of Rats with Transfer Learning

Andres Molares-Ulloa, Maria del Rocio, Julian Dorado, Nieves Pedreira, and Alvaro Rodriguez

Universidade da Coruña, Department of Computer Science and Information Technology, Faculty of Computer Science, 15071, A Coruña, Spain
Centro de Investigación CITIC, Universidade da Coruña, 15071 A Coruña, Spain
Department of Basic Psychology I, UNED, Ciudad Universitaria, Madrid, Spain.
Correspondence: andres.molares@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: The study of animal behavior in laboratory experiments is key in ethology, ecotoxicology, neuroscience and other fields. Although modern studies use computer imaging techniques, current solutions cannot preserve the identity of multiple individuals in social experiments. Thanks to the use of Transfer Learning we seek to overcome this limitations while maintaining the effectiveness of Deep Learning and reducing its computational times. With this technique we achieved promising results in the re-identification of rats after an occlusion process.

1 Introduction

The study of animal behavior is an important part of laboratory experiments conducted in fields such as neuroscience, ecotoxicology, ethology, drug testing and discovery, and many others. These experiments are needed to assess the well-being and health of animals both in nature, where their behavior is impacted by pollution and climate change, and in human facilities, like farms and fish farms, where healthier animals result in lower risks of zoonosis and other pathogens entering the human diet. To conduct this experiments, modern studies rely on computer-imaging techniques, but this doesn't work when studying multiple animals in complex environments is required. Researches still lack techniques to monitor animals and obtain complex behaviors.

Animal behavior experiments are based on the study of behavioral phenotypes that can be defined as evolutionary adaptive traits that emerge as complex patterns of behavior (Cote et al., 2010).

This experiments require large series of tests that run for long time periods of time on large cohorts. In contemporary animal research, laboratories employ video recording and tracking software to simplify analysis, as outlined in (Rodriguez et al., 2018). A tracking application is a computer program designed to assist scientists by extracting motion and location data from video recordings. This is achieved by identifying organisms within each frame of the video and connecting them to the respective animals, enabling the monitoring of their movements. Researchers then utilize this data to gain insights into specific behaviors.

One of the primary limitations of current methods becomes apparent when dealing with multiple individuals. In such scenarios, maintaining the identities of the animals throughout the experiment is crucial. Frequently, animals obscure each other, leading tracking applications to lose sight of the target animals. Consequently, researchers must reassign identities to each animal after they cross paths or become obscured.

The intricacy of this issue is highlighted in (Pérez-Escudero et al., 2014), where they examined a situation involving multiple interacting animals. Even when correctly solving 99% of all crossings, only 11% of the animals were accurately identified after 2 minutes of tracking, due to error propagation. In summary, preserving identity is a complex and computationally intensive task, and only a handful of applications make significant contributions to this field, which remains unsolved.

In prior research (Pérez-Escudero et al., 2014; Rodriguez et al., 2017, 2018), probabilistic texture analysis was utilized to assess the similarity between animals. However, this approach is not suitable for tracking numerous targets over extended periods of time. However, it proves valuable in short experiments involving small groups of animals.

With the rise in popularity of Deep Learning techniques, some authors (Romero-Ferrero et al., 2019; Xu and Cheng, 2017) have employed a deep learning approach known as Convolutional Neural Networks (CNNs). CNNs currently rank among the most potent image classification methods available. In their 2019 work, Romero-Ferrero and colleagues achieved the most robust identity preservation algorithm using this approach. Nevertheless, the computational time required to analyze a standard experiment using this method can be as long as one hour per frame, rendering it impractical for the majority of experiments.

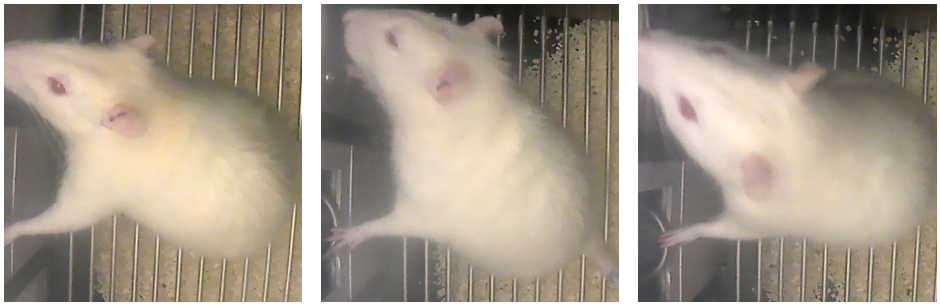
The main goal of this work is the development of a reliable and efficient identity preservation algorithm based on Deep Learning models of image classification, to track multiple identical organisms in social experiments, in scenarios where total occlusion occurs. To achieve this re-identification while maintaining the good results of CNNs but avoiding the long processing times usually required by this approach, we have developed a model based on Transfer Learning. This re-identification model is developed on the Xception architecture.

2 Materials and Methods

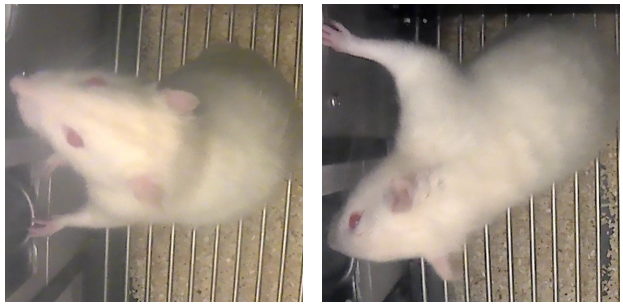
2.1 Dataset and its Construction

We have trained and validated our models with a dataset created in a learning trial with five adult laboratory rats of the species (*Rattus norvegicus domestica*) in conditioning cells (Operant conditioning chamber also known as a Skinner box) used for tasks such as teaching an animal to perform certain actions (like pressing a lever) in response to specific stimuli. We have had a dataset composed of color images from 5 different videos. Each video has been segmented to extract a total of 1500 images for each of the 5 individuals studied. These individuals are cataloged as AFH1, AFH2, AFH3, AFH4 and AFH5. In total we have a set of 1500 images of 5 individuals in 5 videos, totaling 37500 images. Each image has a size of 128 x 128 x 3. Figure 1 shows examples of each individual.

In addition, we have applied a normalization process to the images to verify that the models do not base their classification on technical characteristics of the image. This process consists of a standardization (zero mean and standard deviation were passed to each color channel separately from the whole image), and then the output was normalized to the 0-255 range. We have applied this normalization in two different ways. The first one is applied on the total image before performing the individual segmentation and the second one is applied on the already segmented image window. This has resulted in a total of 3 datasets named: Standard (without normalization), Norm and Norm_Window.



(a) AFH1. (b) AFH2. (c) AFH3.



(d) AFH4. (e) AFH5.

Figure 1: Images of unstandardized individuals.

2.2 Transfer learning and Xception

Transfer learning is a machine learning approach where an initially task-specific model serves as a foundational building block for constructing a new model customized for a different task. This strategy is widely embraced in the domain of deep learning, a subset of machine learning known for its frequent application of such techniques. As convolutional neural networks (CNNs) gained prominence in computer vision, various structured models leveraging CNNs were developed. One notable model is Inception, also referred to as Inception-v1 (Szegedy et al., 2015). The Inception architecture addresses challenges in image classification posed by objects of varying sizes by advocating the use of multiple filters of different sizes at the input stage. Furthermore, it recommends forwarding the output of this module to another Inception module.

The Xception network (Polat, 2021) used in this study can be called an interpretation of the Inception modules. The name Xception also comes from "extreme inception". In Xception architecture, differently from Inception architecture, a convolution operation almost same with depthwise separable convolution is used. This type of convolution contains a depthwise convolution and a pointwise convolution that follows it. In depthwise convolution each filter independently processes only one channel of the input image; and in pointwise convolution, 1x1 dimensional filter iterates every single point of the input. The flows and modules related to the Xception architecture are shown in the Figure 2.

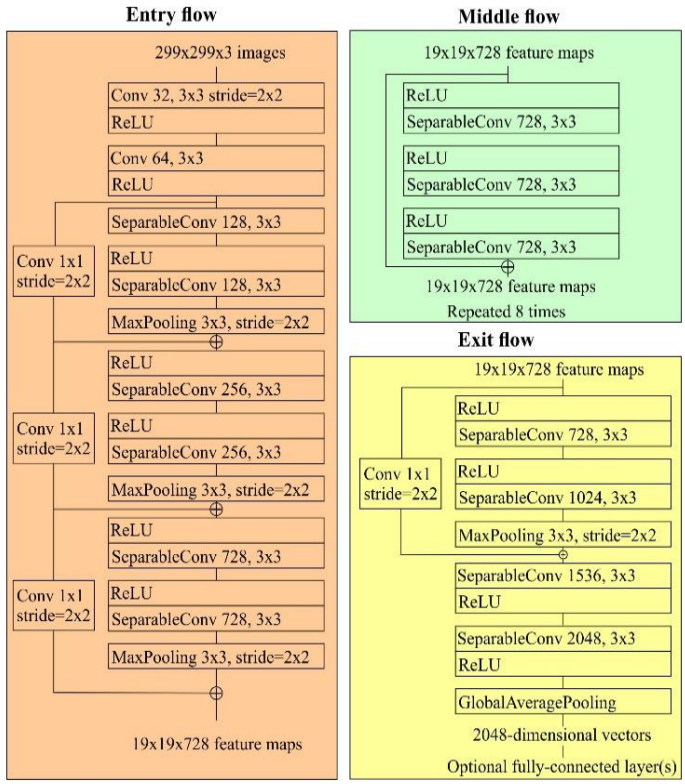


Figure 2: Xception architecture (Polat, 2021).

2.3 Experimentation Setup

To ensure the robustness of the models, we have applied the K-fold cross-validation strategy (Wong (2015)). In this study, the K-fold cross-validation strategy selected will be the 5-fold strategy, where k takes value 5. This is due to the use of 5 videos as a data set. When applying k-fold what we have done is to reserve the images from one video for testing while using the images from the other 4 videos for training. We perform this process until we obtain 5 metrics that robustly show the adequacy of the model. We have added a Data Augmentation layer to the model to increase the variability of the images during the training process. This layer randomly flips, rotates and zooms the training images to increase the generalisability of the system. To determine the values used in the tests, we have carried out an empirical experimentation. The specific configuration of this model is shown in the Table 1.

Table 1: Model configuration parameters

Module		Parameter	Value
Xception		weights	imagenet
		input shape	(128, 128)
Data augmentation		RandomFlip	horizontal and vertical
		RandomRotation	0.2
		RandomZoom	0.1
Output layer	Regular Fit	batch size	(32, 64 and 128)
		epochs	10
		optimizer	Adam(1e-3)
		loss	SparseCategoricalCrossentropy
	Fine tuning	batch size	(32, 64 and 128)
		epochs	5
		optimizer	Adam(1e-6)
		loss	SparseCategoricalCrossentropy

3 Results and Discussion

Table 2 shows the results of applying the model to the 3 datasets studied. We can see that there are no significant differences when applying the different normalization ways. Therefore, we decided to choose the Standard dataset as it has one of the highest means together with the N2 normalization but with a lower standard deviation. We can also observe that there are differences when applying the model depending on the video, with videos 2 and 4 being the most complicated.

Table 2: Model performance

Dataset	Accuracy						
	Video 1	Video 2	Video 3	Video 4	Video 5	Mean	σ
Standard	**53.57%**	**37.51%**	**50.61%**	**30.45%**	**62.07%**	**46.84%**	$\pm$ **12.72%**
Norm	54.07%	36.19%	52.93%	28.28%	62.85%	46.86%	$\pm$ 14.18%
Norm_Window	53.51%	29.44%	52.52%	35.92%	53.25%	44.93%	$\pm$ 11.42%

Figure 3 shows the confusion matrix showing the distribution of the results of the model applied to the unnormalised data set. In it we can see how there are differences in the classification of the different images, with those belonging to the AFH3 rat being the most complicated to distinguish and those belonging to the AFH5 rat the simplest. We can also observe how the model tends to confuse between individuals AFH1 and AFH2 and between AFH3 and AFH4 while AFH5 is the most recognizable individual. This may be due to the presence of characteristic features in some of the individuals. This difference in classification between individuals could also be due to behavioral elements reflected in the individual's habitual posture while others have similar routines.

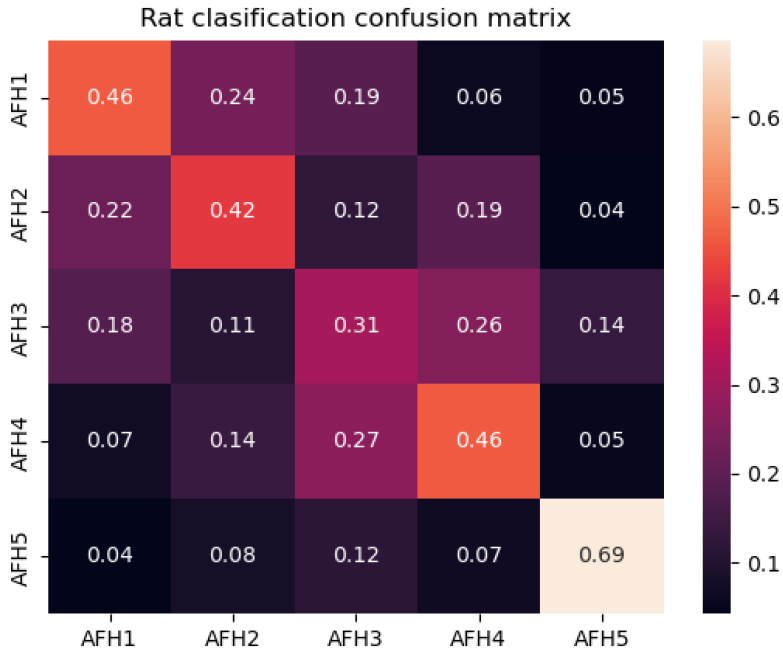


Figure 3: Confusion matrix of the model applied on the Standard dataset.

4 Conclusions

We have defined a performance framework when applying Deep Learning, and in particular Transfer Learning, to classify multiple individuals. We have achieved results of around 47% accuracy when re-identifying up to 5 individuals simultaneously. Being 20% accuracy the threshold from which we started (random classification) we can conclude that we have achieved promising results. We are therefore looking for further improvements in both the quality of the classification and the number of individuals that the system is able to re-identify simultaneously.

We have shown the importance of having a large data set that allows the development of a model with good generalization capacity and avoids biasing the system. Thus avoiding that the model bases its classification on artifacts present in the images and not on the individual's own characteristics. For this, it is important to have several sets of images obtained from different devices, as has been our case.

Acknowledgments

Grant PID2021-126289OA-I00 funded by MCIN/AEI/10.13039/501100011033 and by ERDF A way of making Europe. CITIC is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS).

Credit authorship contribution statement

Andres Molares-Ulloa wrote the manuscript, trained and validated the models. **Maria del Rocio** performed and recorded the experiments. **Alvaro Rodriguez** processed the data, created the datasets, and coordinated and supervised the work. All authors revised and edited the manuscript.

Bibliography

J. Cote, J. Clobert, T. Brodin, S. Fogarty, and A. Sih. Personality-dependent dispersal: characterization, ontogeny and consequences for spatially structured populations. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1560):4065–4076, 2010.

O. Polat. Detection of covid-19 from chest ct images using xception architecture: A deep transfer learning based approach. *Sakarya University Journal of Science*, 25(3):800 – 810, 2021.

A. Pérez-Escudero, J. Vicente-Page, R. C. Hinz, S. Arganda, and G. G. de Polavieja. idtracker: tracking individuals in a group by automatic identification of unmarked animals. *Nat Methods*, 2014.

A. Rodriguez, H. Zhang, J. Klaminder, T. Brodin, and M. Andersson. Toxid: an efficient algorithm to solve occlusions when tracking multiple animals. *Scientific reports*, 7(1):14774, 2017.

A. Rodriguez, H. Zhang, J. Klaminder, T. Brodin, P. L. Andersson, and M. Andersson. Toxtrac: A fast and robust software for tracking organisms. *Methods in Ecology and Evolution*, 9(3): 460–464, 2018.

F. Romero-Ferrero, M. G. Bergomi, R. C. Hinz, F. J. Heras, and G. G. De Polavieja. Idtracker. ai: tracking all individuals in small or large collectives of unmarked animals. *Nature methods*, 16 (2):179–182, 2019.

C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015.

T. T. Wong. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recognition*, 48, 2015.

Z. Xu and X. E. Cheng. Zebrafish tracking using convolutional neural networks. *Scientific reports*, 7(1):42815, 2017.

Use of Deep Learning Techniques for Motor Events Detection in Polysomnographic Records

Inés Lloves García and Diego Álvarez Estévez

Faculty of Informatics, Universidade da Coruña, 15071 A Coruña, Spain
Centro de Investigación CITIC, Universidade da Coruña, 15071 A Coruña, Spain
Correspondence: ines.lloves@udc.es, diego.alvareze@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: Sleep medicine deals with the diagnosis and treatment of sleep-related disorders. The diagnosis is carried out through the manual analysis and labeling of polysomnographic studies, which record various electrophysiological and pneumological signals of the patient throughout the night. This process involves the analysis of long duration signals, is complex, and demands considerable resources and time on the part of the clinical expert. The purpose of this proyect is the construction of automatic analysis algorithms that considerably reduce the analysis duration, reducing the manual workload, and minimizing possible human errors, providing repeatability and robustness. In particular, the objective is to use machine learning algorithms, based on Deep Learning techniques, for the identification and location of physiological events in these polysomnographic records. Specifically, the goal is to locate physiological events associated with involuntary motor movements that occur in the limbs, known as Limb Movements.

1 Introduction

Polysomnography is a study utilized to evaluate and diagnose sleep-related disorders. The polysomnographic recording (PSG) obtained, records multiple physiological signals during the night. The patient sleeps in a controlled environment, usually in a sleep laboratory, where electrodes and sensors are placed on various parts of the body to record these signals.

PSG is considered the gold standard for diagnosing sleep-related breathing disorders. Each recording lasts at least 8 hours, so manually analyzing each signal is a slow, time-consuming process, that generates high costs in human resources. This causes delays in diagnosis and increases the waiting lists, and therefore, motivates the research and development of automatic analysis tools for PSGs.

In the context of automatic detection of involuntary motor events, the reference signal is the Electromyogram (EMG) of the tibialis anterior muscles. EMG records electrical signals generated by muscle activity. Involuntary movements, known as Limb Movements (LM), often occur during sleep, especially during deep sleep phases, and can be identified by observing the muscle activity through an EMG signal.

Brief and repetitive episodes of LM during sleep are known as Periodic Limb Movements (PLM), which are found between the 85% and 95% of patients with Restless Legs Syndrome (Rye and Trotti, 2012), a neurological disorder characterized by an uncomfortable sensation in the legs and the urgent need to move them to relieve the sensation. These movements can disrupt sleep and cause insomnia problems. They have also been observed in other sleep-related neurological disorders, such as narcolepsy (Baker et al., 1986), sleep-related breathing disorders (Ancoli-israel et al., 1985), Parkinson's disease (Wetter et al., 2000), multiple system atrophy (Vetrugno et al., 2004) and REM sleep behavior disorder (Fantini et al., 2002). Therefore the importance of the detection and analysis of LM.

2 Material

For the development of this project, we have used PSGs annotated by clinical experts. Only the EMG signals from the PSG records were used. This signal is stored in two channels, with the signal originated from the right tibialis anterior muscle (EMG RAT) and from the left tibialis anterior (EMG LAT), which will be used as independent signals. Three different data sets have been used:

HMCP: comes from the Álvarez Estévez and Rijsman (2022) study. The study consists of 20 PSGs individually selected for 4 tasks. In this project, we have used the 5 PSGs selected in this study for the LM annotation task, which are individually annotated by 12 clinical experts. The EMG signals are stored in two separate channels for each leg, and are sampled at a frequency of 128 or 256 Hz.

MrOS: comes from the Osteoporotic Fractures in Men Study (MrOS) (Blank et al., 2005). 3,135 of the participants were recruited to take the MrOS sleep study, and underwent the recording of a complete PSG. A subset of 10 randomly selected PSGs has been used in this project. Each PSG includes annotations made by one clinical expert. All EMG signals are stored in two separate channels for each leg, and are sampled at a frequency of 64 Hz.

WSC: comes from a follow-up of mortality in the population of the Wisconsin Sleep Cohort (WSC) (Young et al., 2008). Within the large sample, only a subset of patients had all PSG awakenings and LMs scored including the duration of the events. In this project, 16 PSG records were randomly selected from that subset. The EMG signals from these recordings are sampled at a frequency of 200 Hz, and only the right leg channel is recorded.

3 Methods

The task of detecting LM in EMG signals is approached as a supervised sequence-to-sequence classification problem, where the classifier is a deep learning model. The objective is to locate the start and the end point of each LM in the EMG signals, and to determine the specific channel in which the event takes place (EMG RAT/LAT).

3.1 Signal preprocessing

EMG signals can be contaminated by several types of noise, interferences and artifacts. To avoid possible misinterpretation of the data, a suitable filtering system is implemented to prepare the EMG signals. The filtering pipeline consists of a high-pass filter with a 15 Hz cut-off frequency, to remove low-frequency components not related to the muscle activity of interest; and a Notch filter centered on 50 Hz (Europe) or 60 Hz (North America) (according to each data set origin), to eliminate power line interference.

Signals can be recorded at different sampling frequencies. Among the three databases used, we have signals sampled at 64, 128, 200 and 256 Hz. In order to have all the data in a single homogeneous format, a resampling of the EMG LAT/RAT channels is performed at 128 Hz.

3.2 Sample processing

Once the EMG LAT and RAT signals have been preprocessed, the process of obtaining the samples, to train the models and subsequently to make predictions, is carried out. Training samples are pairs **(X, Y)**, where:

X: is a vector with the amplitude values of an interval of the input EMG signal. To obtain these vectors from the EMG signals, windowing is performed using a overlapping

sliding time window. A window length of 30 seconds is used (clinicians usually analyze these signals in 30 second intervals), and an overlap of the 50%.

Y: is a boolean vector that corresponds to the annotations of an expert indicating the presence or absence of LM in the signal interval. To obtain the label vectors, a resolution of 0.25 seconds is used, that is, every 0.25 seconds the signal has an associated label indicating the presence or absence of LM.

Training samples are obtained from the EMG signals of the PSGs of the HMCP dataset, while the MrOS and WSC datasets are utilized to evaluate the generalization of the models between different databases.

3.3 Training process

For the training process, samples are divided into three partitions:

- **Training partition**: 70% of the pairs (X, Y), are utilized for adjusting and learning the models.

- **Validation partition**: 15% of the pairs (X, Y), are utilized to monitor the learning of the models.

- **Test partition**: remaining 15% of the pairs (X, Y), are utilized to evaluate the performance of the trained models.

Different architecture models have been trained, they are presented in Section 4.

3.4 Prediction and annotation

To predict LM from the EMG LAT/RAT signals, the preprocessing and processing steps are carried out to obtain the X input vectors. Once the predictions (Y vectors) are obtained from the trained models, they are processed to transform predicted labels into annotations that follow the LM annotation criteria established by the World Association of Sleep Medicine (WASM) (Ferri et al., 2016).

3.5 Validation

Cohen kappa (Cohen, 1960) will be used as the main reference metric to measure and compare the quality of the algorithms. It is a useful metric because it considers the correct and the incorrect classification, and corrects the agreement due to chance. It is frequently used to test inter-rater agreement, and is the main reference metric for annotating events in PSG signals.

4 Deep learning approaches

Different architectures have been evaluated for the LM detection task. Table 1 shows a summary of the trained models and the Kappa values obtained by each of them in prediction of the test partition data.

Table 1: Trained models summary

Model	Arquitecture	Parameters	Training samples	Cohen Kappa
Method 1	LSTM	9,217,880	only pairs (X,Y) from EMG RAT	0.6769
Method 2	LSTM	9,217,880	all pairs (X,Y)	0.7422
Method 3	LSTM	69,117,300	all pairs (X,Y)	0.7001
Method 4	BiLSTM	46,119,940	all pairs (X,Y)	0.6830
Method 5	1D CNN - BiLSTM	4,610,680	all pairs (X,Y)	0.7612
Method 6	1D CNN - BiLSTM	4,610,680	only pairs (X,Y) containing some LM	0.7219
Method 7	1D CNN - BiLSTM	4,610,680	pairs (X,Y) containing some LM and 10% of pairs not containing any LM	0.7391
Method 8	LSTM - BiLSTM	9,220,360	all pairs (X,Y)	0.7547

5 Results

We evaluated the performance of the trained models. Additionally, we evaluated the performance of the automatic annotation method of Álvarez Estévez (2016) (Polyman method), based on conventional signal processing methods, so as to compare its performance to our models. This method was designed with PSGs coming from the same medical center as HMCP.

5.1 HMCP

An inter-expert validation was carried out among the different HMCP clinicians, to know the reference measure of performance between experts. We measured the Kappa agreement between each clinical expert individually and the consensus formed by the rest of the experts, in the annotation of LM in every EMG signal. The average Kappa of agreement between experts was **0.7867**, with a deviation of $\pm$ **0.0920**.

Equally, the evaluation between the automatic annotation methods and the expert consensus was carried out (see Table 2) for each EMG signal. Additionally, a multiple comparison test was carried out to analyze the statistical significance of the differences between methods (see Figure 1).

Table 2: HMCP validation results

Model	Cohen Kappa
Method 1	0.7739 ± 0.0648
Method 2	0.8100 ± 0.0677
Method 3	0.7964 ± 0.0691
Method 4	0.7941 ± 0.0616
Method 5	0.8217 ± 0.0450
Method 6	0.6889 ± 0.3148
Method 7	0.8285 ± 0.0350
Method 8	0.8088 ± 0.0487
Polyman	0.8415 ± 0.0818

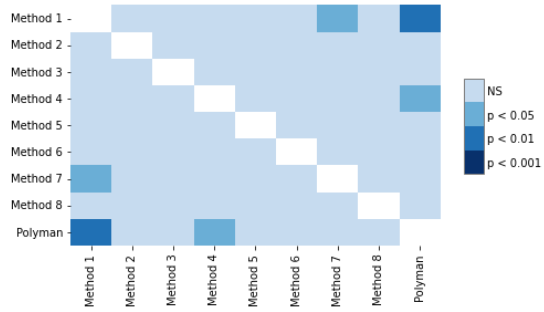


Figure 1: Multiple comparison test on HMCP

5.2 MrOS and WSC

Validation was carried out for the MrOS and WSC data sets, to evaluate the generalization capacity of the automatic annotation methods on data sets independent of the training one.

In this case, we only have the annotations of one clinical expert per EMG signal, therefore we measure the agreement between the methods and the expert (instead of a consensus). Tables 3 and 4 contain the corresponding Kappa means and standard deviations for MrOS and WSC respectively. Additionally, a multiple comparison test was carried out (see Figures 2 and 3).

Table 3: MrOS validation results

Model	Cohen Kappa
Method 1	0.4470 ± 0.1866
Method 2	0.2937 ± 0.1589
Method 3	0.2151 ± 0.1606
Method 4	0.2614 ± 0.1692
Method 5	0.2036 ± 0.1576
Method 6	0.4291 ± 0.1894
Method 7	0.3850 ± 0.1652
Method 8	0.2574 ± 0.1542
Polyman	0.1419 ± 0.2337

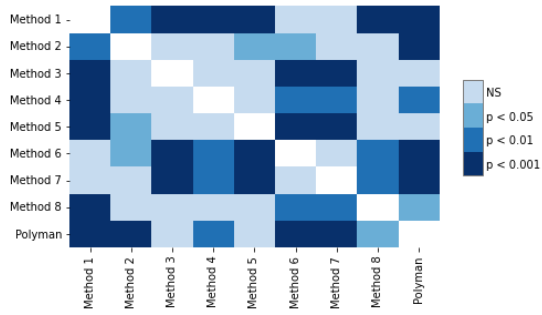


Figure 2: Multiple comparison test on MrOS

Table 4: WSC validation results

Model	Cohen Kappa
Method 1	0.7604 ± 0.0943
Method 2	0.7583 ± 0.0987
Method 3	0.7509 ± 0.1084
Method 4	0.7234 ± 0.1117
Method 5	0.7689 ± 0.0929
Method 6	0.7459 ± 0.1106
Method 7	0.7732 ± 0.0943
Method 8	0.7546 ± 0.0960
Polyman	0.4437 ± 0.3170

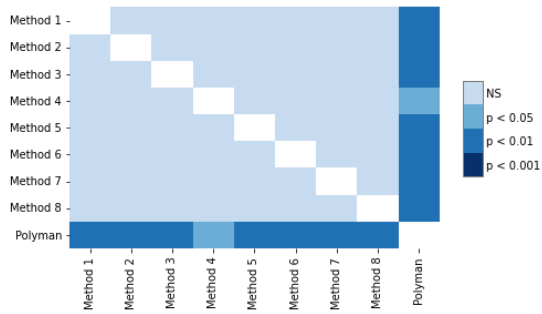


Figure 3: Multiple comparison test on WSC

6 Discussion

Firstly, analyzing the results of the HMCP data set (see Section 5.1), clinical experts obtained an average Kappa agreement of 0.7867 ± 0.0920, while six of our models (Method 2, Method 3, Method 4, Method 5, Method 7 and Method 8) and the Polyman method exceed this value and were more stable in deviation (see Table 2).

On the other hand, the Polyman method obtained the highest Kappa on average over HMCP, although our models have greater stability. However, in the multiple comparison test (see Figure 1), the Polyman method was only significantly different from the Method 1 and Method 4 models. Among our models, Method 7 obtained the highest Kappa and the lower deviation, although the multiple comparison test indicates that it is only significantly different from the Method 1 model, therefore, among the rest of the models, the computationally less expensive method (with fewer training parameters), which are Methods 5, 6 or 7, are preferred.

Analyzing Table 3 we can observe that in the MrOS set, that contains signals with worse quality than the training data (signals were sampled only at 64 Hz), all our models obtain better performance and are more stable than the Polyman method. The multiple comparison test (see Figure 2) indicates that the difference in performance between the Polyman method and our models is significant except for Method 3 and Method 5 models. Among our models, the one that obtained the highest Kappa was Method 1, which turned out to be significantly equivalent in performance to the Method 6 and Method 7 models, therefore, again, the computationally less expensive models are preferred, Methods 6 or 7.

Analyzing Table 4 we can observe that in the WSC set, that contains signals with comparable quality to the training data, all our models obtain better performance and are more stable than the Polyman method. Furthermore, the multiple comparison test (see Figure 3) indicates that the difference in performance between the Polyman method and each of our models is significant. Finally, the performance differences between our models are not statistically significant, so again, the computationally less expensive models are preferred, Methods 5, 6 or 7.

7 Conclusions

We implemented a process for automatic annotation of involuntary motor movements in the limbs, on EMG signals recorded in a PSG. The method includes the processes of preparing the EMG signals for prediction and the subsequent annotation of the detected events.

For the automatic detection of LM, the applicability and suitability of different deep learning architectures was investigated. Eight different deep learning models were trained: three with an LSTM architecture, one with a BiLSTM architecture, three 1D CNN-BiLSTM, and one LSTM-BiLSTM.

The models obtained solid performance, better or comparable to the performance of HMCP clinical experts, and above all, they were more stable, which is highly desirable to eliminate or minimize subjectivity and possible human errors.

The performance of our deep learning models was compared to the automatic annotation method of Álvarez Estévez (2016), based on conventional signal processing methods. The results indicated that methods based on deep learning are more robust and stable facing with signals from new data sources, with different qualities, configurations, recording processes, and in which clinical experts may have slight discrepancies in the annotation criteria.

Finally, models with more complex architectures, and therefore with higher computational cost, did not present significant advantages in performance compared to simpler and more efficient models.

The implemented process could be used to largely reduce the costly manual work of LM annotation, and thus, enable the diagnosis and treatment of many people that are not diagnosed due to limited human resources.

Acknowledgments

CITIC is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS).

Bibliography

S. Ancoli-israel, D. F. Kripke, W. Mason, and O. J. Kaplan. Sleep Apnea and Periodic Movements in an Aging Sample1. *Journal of Gerontology*, 40(4):419–425, 07 1985.

T. L. Baker, C. Guilleminault, G. Nino-Murcia, and W. C. Dement. Comparative Polysomnographic Study of Narcolepsy and Idiopathic Central Nervous System Hypersomnia. *Sleep*, 9 (1):232–242, 03 1986.

J. B. Blank, P. M. Cawthon, M. L. Carrion-Petersen, L. Harper, J. P. Johnson, E. Mitson, and R. R. Delay. Overview of recruitment for the osteoporotic fractures in men study (mros). *Contemporary Clinical Trials*, 26(5):557–568, 2005. URL *https://www.sciencedirect.com/science/article/pii/S1551714405001060*.

J. Cohen. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1):37–46, 1960.

M. Fantini, M. Michaud, N. Gosselin, G. Lavigne, and J. Montplaisir. Periodic leg movements in rem sleep behavior disorder and related autonomic and eeg activation. *Neurology*, 59(12): 1889–1894, 2002. URL *https://n.neurology.org/content/59/12/1889*.

R. Ferri, S. Fulda, R. Allen, M. Zucconi, O. Bruni, S. Chokroverty, L. Ferini-Strambi, B. Frauscher, D. Garcia-Borreguero, M. Hirshkowitz, B. Högl, Y. Inoue, A. Jahangir, M. Manconi, C. Marcus, D. Picchietti, G. Plazzi, J. Winkelman, and R. Zak. World association of sleep medicine (wasm) 2016 standards for recording and scoring leg movements in polysomnograms developed by a joint task force from the international and the european restless legs syndrome study groups (irlssg and eurlssg). *Sleep Medicine*, 26:86–95, 2016. URL *https://www.sciencedirect.com/science/article/pii/S1389945716302398*.

D. B. Rye and L. M. Trotti. Restless legs syndrome and periodic leg movements of sleep. *Neurologic Clinics*, 30(4):1137–1166, 2012. URL *https://www.sciencedirect.com/science/article/pii/S073386191200045X*. Sleep Disorders.

R. Vetrugno, F. Provini, P. Cortelli, G. Plazzi, E. M. Lotti, G. Pierangeli, C. Canali, and P. Montagna. Sleep disorders in multiple system atrophy: a correlative video-polysomnographic study. *Sleep Medicine*, 5(1):21–30, 2004. URL *https://www.sciencedirect.com/science/article/pii/S1389945703001953*.

T. Wetter, V. Collado-Seidel, T. Pollmacher, A. Yassouridis, and C. Trenkwalder. Sleep and periodic leg movement patterns in drug-free patients with parkinson's disease and multiple system atrophy. *Sleep*, 23:361–367, 05 2000.

T. Young, L. Finn, P. Peppard, M. Szklo-Coxe, D. Austin, F. Nieto, R. Stubbs, and K. Hla. Sleep disordered breathing and mortality: Eighteen-year follow-up of the wisconsin sleep cohort. *Sleep*, 31:1071–8, 09 2008. URL *https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2542952/*.

D. Álvarez Estévez. A new automatic method for the detection of limb movements and the analysis of their periodicity. *Biomedical Signal Processing and Control*, 26:117–125, 2016. URL *https://www.sciencedirect.com/science/article/pii/S1746809416300015*.

D. Álvarez Estévez and R. M. Rijsman. Computer-assisted analysis of polysomnographic recordings improves inter-scorer associated agreement and scoring times. *PLOS ONE*, 17 (9):1–19, 09 2022.

Development of a Virtual Sensor for COD Measurement in a Wastewater Treatment Plant

Antonio Díaz-Longueira, Míriam Timiraos, Álvaro Michelena, Óscar Ramos-Fontenla, and Jose Luis Calvo-Rolle

Ciencia e Técnica Cibernética (CTC), Department of Industrial Engineering, Universidade da Coruña, 15071 A Coruña, Spain
Fundación Instituto Tecnológico de Galicia, Department of Water Technologies, National Technological Center, Cantón Grande 9, Planta 3, C.P. 15003, A Coruña, Spain.
Centro de Investigación CITIC, Universidade da Coruña, 15071 A Coruña, Spain
Correspondence: a.diazl@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: The objective of the work is to develop a system that allows predicting, from a global perspective, the behavior of the process in a wastewater treatment plant. To do this, the chemical oxygen demand, a variable present in water, is estimated indirectly, avoiding difficult and complex measurements. This estimation is carried out in real time through the relationship between easily measured variables. This modeling will be done through the use of machine learning techniques. Different regression techniques are applied and compared. The dataset contains variables such as pH, conductivity, suspended solids and etc. In this way, a non-physical indirect sensor is implemented. Thresholds are established for the detection of deviations in the sensor parameters.

1 Introduction

Water is a scarce and irreplaceable good, with great importance in the field of health and production in our country and the world. The constant advance of climate change and the unstoppable growth of the world's population are affecting its availability, making it an increasingly scarce resource. In view of this situation, and as a possible measure in this situation, the possibility of reusing wastewater (Salgot and Folch, 2018) appears. This reuse is subject to eliminating harmful agents that may be present. With the aim of reducing wastewater pollution to values valid for reuse, wastewater treatment plants (WWTP) appear, capable of reducing the polluting load (EDAR, n.d.).

To ensure and make sure of the correct operation of these facilities, it is essential to know the state of the water, by means of certain markers that make it possible to establish the type and degree of contamination of the water, both in the inflow and outflow of the WWTP. These markers include physicochemical variables that are costly and/or technically complicated to measure. Instead of using physical sensors to measure these markers, it is possible to try to estimate, based on other variables with a simpler measurement, their value. An indicator of the degree of water contamination is the chemical oxygen demand, which provides an idea of the presence of both organic and inorganic agents (Clesceri et al., 1999).

This estimation is carried out in real-time based on the existing relationship with these third variables. In this way, a non-physical indirect sensor is implemented, which makes use of machine learning techniques to model this relationship with the rest of the variables and predict their value. For this reason, regression techniques will be used. With the indirect sensor already implemented, it is possible to establish thresholds for the detection of deviations in the parameters and create an alarm system. These thresholds may be established by different methods.

2 Materials and methods

This section describes the machine learning techniques and algorithms, the metrics and procedures used for the evaluation of the models, the graphs where the results will be plotted and the data set used in the analysis.

2.1 Machine learning techniques and algorithms

To try to carry out a study that is as heterogeneous and varied as possible, different supervised learning techniques and algorithms were evaluated to analyze and compare their performance in the sought estimates. The following techniques were used:

- Recursive Least Squares (RLS).
- K-Nearest Neighbors (KNN).
- Decision Tree (DT).
- Support Vector Regression (SVR).
- MultiLayer Perceptron (MLP).

2.2 Model evaluation

For the evaluation and subsequent selection of the models, different metrics were used to determine their performance. The metrics used were:

- Mean Absolute Error (MeanAE).
- Mean Squared Error (MSE).
- Root Mean Squared Error (RMSE).
- Symmetric Mean Absolute Percentage Error (SMAPE).
- Coefficient of determination (R^2).

2.3 Dataset

The data set used is a real set, taken from measurements from 3 wastewater treatment plants. These measurements were carried out over 3 months: June, July and August, with one measurement per day. In addition to the physicochemical variables related to water, there is also a variable that indicates the daily volume of water processed by the WWTP, two that establish the day and month of the measurement and a last one indicating the WWTP from which the data were obtained. These last four variables were not used in the development of the work.

The data set used for the development of the work is composed of 8 physical-chemical variables of water. These variables are pH, conductivity (Cond), biochemical oxygen demand (DBO), chemical oxygen demand (DQO), nitrates (N), phosphates (F), suspended solids and settleable solids (V60). Figure 1 shows the correlation matrix between the different variables. The case of DQO stands out.

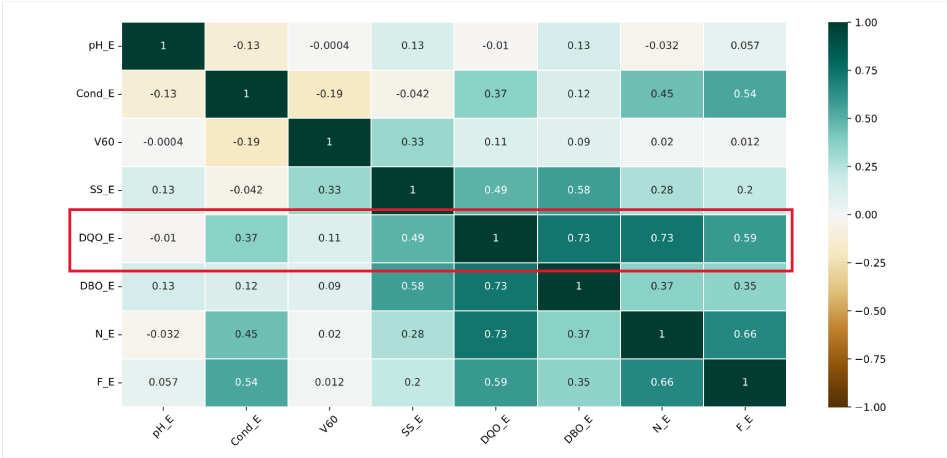


Figure 1: Correlation matrix

The measurements do not follow a specific periodicity; it is possible that no measurements were taken on a specific day, or that some of the physical-chemical variables were not measured. For the development of the work, since it is not known which variables will be used for a prediction, only those records in which all the variables are present will be used.

This implies that, of the 276 expected records, the data set will be smaller.

3 Experiments

This section details the experiments performed to design the best indirect sensor. For this purpose, and looking for the best model to implement, the performance of different regression techniques will be evaluated. They will be configured through their hyperparameters until the best prediction is reached.

The three variables with the highest correlation with COD were used. According to Figure 1, they are DBO, nitrates and phosphates.

The comparison between experiments will be performed based on the metrics obtained in 3-kfold cross-validation, selected due to the small size of the dataset. To know the performance of a model and to be able to compare it with that of the others, first of all, the metrics observed are the coefficient of determination and the SMAPE. Since the SMAPE does not handle over- and under-forecasting in the same way, as soon as these metrics do not allow us to select a clear winner, we will proceed to focus on the other metrics.

The following are the hyperparameters modified in each of the techniques, and the values tested.

3.1 Recursive Least Squares

The models generated were tested by forcing the coefficients to be positive (*positive* with possible values of True or False) and to calculate or not the independent term (*intercept* with possible values of True or False).

The configurations of this technique will receive their name following the following construc-

tion: RLS + intercept value + positive value. In this way, the configuration with *positive* set to True and *intercept* set to False will be named RLSFalseTrue.

3.2 K-Nearest Neighbors

It was obtained, experimentally, that the best results are obtained when using an odd number of neighbors between 3 and 9. Two functions were evaluated for assigning weights to each neighbor: *uniform* y *distance*.

The configurations of this technique will receive their name following the following construction: KNN + neighbor + weight function. In this way, the configuration with 5 neighbors and *uniform* weight distribution will be named KNN5uniform.

3.3 Decision Tree

Two methods for determining the best split at each node (criterion) were tested: *absolute_error* y *squared_error*. The maximum depth of the diagram was also modified, from 1 to 7. This range was obtained experimentally.

The configurations of this technique will receive their name following the following construction: DT + depth + criterion. In this way, the configuration with a maximum depth of 3 and squared_error as criterion will be named DT3squared_error.

3.4 Support Vector Regression

Experimentally, it was proven that the best results appeared for values of the regularization coefficient of 10 and 0.1. Also experimentally, it was found that the best *epsilon* values were 1. Finally, three kernels were tested: *linear*, *sigmoidal* and *tansig*.

The configurations of this technique will receive their name following the following construction: SVR + regularization coefficient + kernel. In this way, the configuration with a regularization coefficient of 10 and a linear kernel will be named SVR10linear.

3.5 MultiLayer Perceptron

Experimentally, it was proven that the best results were obtained when working with an intermediate layer of 8 to 9 neurons, so these were the tested values. Also, three activation functions were analyzed for the input and intermediate layers: *linear*, *sigmoidal* and *tangent-sigmoidal*.

The configurations of this technique will receive their name following the following construction: MLP + hidden neurons + activation function. In this way, the configuration with 10 hidden neurons and a linear function activation will be named MLP10linear.

4 Results

This section collects the results of the different experiments. A table is included with each ML method, with the configurations tested. The best result of each technique is highlighted in bold.

Table 1 shows the mean value of metrics obtained by the configurations in the Recursive Least Squares method.

Table 2 shows the mean value of metrics obtained by the configurations in the K-Nearest Neighbors.

Table 1: Mean value of metrics by the RLS

Configuration	MeanAE	SMAPE	MSE	RMSE	MaxError	R2
RLSTrueTrue	**69,554**	**5,584**	**8020,624**	**89,009**	**245,365**	**0,781**
RLSTrueFalse	**69,554**	**5,584**	**8020,624**	**89,009**	**245,365**	**0,781**
RLSFalseTrue	74,006	6,096	8888,193	93,335	257,115	0,761
RLSFalseFalse	74,006	6,096	8888,193	93,335	257,115	0,761

Table 2: Mean value of metrics by the KNN

Configuration	MeanAE	SMAPE	MSE	RMSE	MaxError	R2
KNN3distance	63,754	5,013	9627,722	95,873	385,558	0,742
KNN3uniform	64,074	5,039	9605,909	96,251	387,222	0,742
KNN5distance	63,515	4,929	9114,417	93,655	379,033	0,756
KNN5uniform	64,063	4,986	8885,404	92,860	376,933	0,762
KNN7distance	63,089	4,918	8857,006	92,510	367,355	0,763
KNN7uniform	**63,406**	**4,951**	**8663,092**	**91,774**	**367,048**	**0,768**
KNN9distance	62,373	4,846	8872,268	92,466	369,004	0,762
KNN9uniform	63,954	4,965	8809,346	92,514	368,889	0,764

Table 3 shows the mean value of metrics obtained by the configurations in the Decision Tree method.

Table 3: Mean value of metrics by the DT

Configuration	MeanAE	SMAPE	MSE	RMSE	MaxError	R2
DTabsolute_error2	85,552	16,183	14478,422	120,312	444,000	0,586
DTabsolute_error5	84,142	17,261	14283,554	119,385	414,833	0,584
DTabsolute_error6	88,583	17,201	15310,806	123,573	399,000	0,553
DTabsolute_error7	86,358	17,266	15681,057	124,625	420,833	0,537
DTsquared_error2	84,031	15,663	13666,092	116,779	433,070	0,614
DTsquared_error5	**72,930**	**17,033**	**10608,092**	**102,972**	**401,584**	**0,697**
DTsquared_error6	72,035	16,922	11176,430	105,543	416,976	0,679
DTsquared_error7	77,131	17,102	12385,229	111,223	423,194	0,648

Table 4 shows the mean value of metrics obtained by the configurations in the Support Vector Regression method.

Table 5 shows the mean value of metrics obtained by the configurations in the MultiLayer Perceptron method.

Table 4: Mean value of metrics by the SVR

Configuration	MeanAE	SMAPE	MSE	RMSE	MaxError	R2
SVR10linear	67,894	16,182	7891,394	88,251	265,448	0,766
SVR10rbf	124,656	13,063	25681,210	159,469	441,298	0,267
SVR10sigmoid	162,113	12,829	44650,823	209,667	498,596	-0,264
SVR0.1linear	**67,588**	**16,127**	**7796,304**	**87,756**	**260,655**	**0,768**
SVR0.1rbf	154,143	12,707	40419,934	199,541	483,021	-0,145
SVR0.1sigmoid	154,680	12,711	40700,431	200,229	483,797	-0,153

Table 5: Mean value of metrics by the MLP

Configuration	MeanAE	SMAPE	MSE	RMSE	MaxError	R2
MLPlinear8	86,098	6,979	12166,413	108,842	281,640	0,628
MLPlinear9	**74,370**	**6,138**	**9073,254**	**94,628**	**282,360**	**0,734**
MLPlinear10	78,792	6,586	9683,534	97,898	276,400	0,721
MLPtansig8	559,587	75,912	349629,048	590,234	1017,871	-9,447
MLPtansig9	550,588	73,704	339717,577	581,715	1008,872	-9,148
MLPtansig10	537,830	70,625	325936,970	569,668	996,115	-8,734
MLPsigmoid8	604,963	88,111	402400,503	633,426	1063,248	-11,039
MLPsigmoid9	584,359	82,296	377848,954	613,769	1042,644	-10,300
MLPsigmoid10	597,566	86,008	393466,340	626,357	1055,850	-10,769

5 Conclusions and future works

The objective of the work was to develop an indirect, non-physical sensor that would allow estimating the COD value through third variables, thus facilitating the measurement of this pollution marker. The sensor developed presents acceptable R^2 values. This is due to the strong correlation that is present between the variables used with respect to chemical oxygen demand. The highest value obtained is 0.781, achieved by using the RLS technique when the independent term is calculated, since forcing the coefficients does not affect performance. The value of the error metrics is also good, since it makes a relative error of $5,584\%$.The rest of the sensors present similar results, but because they use more complex techniques, it was decided to use the one mentioned above. In Figure 2 we check the performance of the indirect sensor.

The sensor obtained is relatively reliable, with a prediction that is closer to the ideal and does not make large errors. Once the sensor has been designed, but not yet implemented, the introduction of thresholds for detecting parameter deviations and generating early warnings can be established based on statistical methods, such as standard deviation or percentage margins; through expert knowledge indicating security thresholds and other methods, such as machine learning techniques for anomaly detection.

With the aim of improving indirect sensors, and increasing the accuracy of the prediction, these sensors could be developed specifically for a certain WWTP, instead of trying to generalize. For this, it would be necessary to increase the size of the dataset with new measurements. The method could be modified when selecting the variables to be used, opting for other requirements than the correlation between them. It would also be interesting to study the possibility

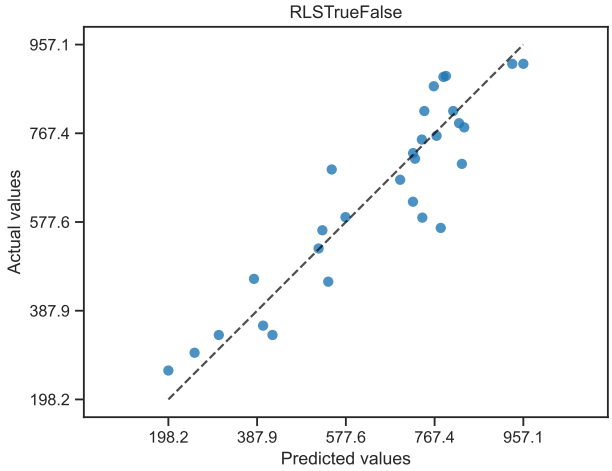


Figure 2: Predicted vs. actual values

that the physical-chemical variables could be part of a time series, for which it would be appropriate to apply other regression techniques such as Long-Short Term Memory (LSTM).

Acknowledgement

Míriam Timiraos's research was supported by the "Xunta de Galicia" (Regional Government of Galicia) through grants to industrial PhD (*http://gain.xunta.gal/*), under the "Doutoramento Industrial 2022" grant with reference: 04_IN606D_2022_ 2692965.

Álvaro Michelena's research was supported by the Spanish Ministry of Universities (*https://www.universidades.gob.es/*), under the "Formación de Profesorado Universitario" grant with reference: FPU21/00932.

CITIC is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS).

Bibliography

L. Clesceri, A. Greenberg, A. Eaton, and M. Franson. Standard methods for the examination of water and wastewater, ed. american public health association-american water works association-water environment federation, usa, 19 p., 1999.

EDAR. Estación depuradora de aguas residuales (edar), n.d. URL *https://www.miteco.gob.es/es/agua/temas/saneamiento-depuracion/sistemas/edar/*.

M. Salgot and M. Folch. Wastewater treatment and water reuse. *Current Opinion in Environmental Science & Health*, 2:64–74, 2018.

Understanding the Influence of Rendering Parameters in Synthetic Datasets for Neural Semantic Segmentation Tasks

Manuel Silva, Omar A. Mures, Antonio Seoane, and Jose A. Iglesias-Guitian

CITIC, Universidade da Coruña, 15071 A Coruña, Spain
Civil Engineering Department, Universidade da Coruña, 15071 A Coruña, Spain
Correspondence: j.iglesias.guitian@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: Deep neural networks are well known for demanding large amounts of training data, motivating the appearance of multiple synthetic datasets covering multiple domains. However, synthetic datasets have not yet outperformed real data for autonomous driving applications, particularly for semantic segmentation tasks. Thus, a deeper comprehension about how the parameters involved in synthetic data generation could help in creating better synthetic datasets. This work provides a summary review of prior research covering how image noise, camera noise and rendering photorealism could affect learning tasks. Furthermore, we presents novel experiments aimed at advancing our understanding around generating synthetic data for autonomous driving neural networks aimed at semantic segmentation.

1 Introduction

In recent years, deep neural networks have become the prevailing solution for addressing most of the computer vision challenges, like object detection, image classification, and semantic segmentation (Janai et al., 2020). However, the progress of deep neural networks is often hampered by the scarcity and quality of available data. This becomes even more drastic in the case of semantic segmentation, where per-pixel annotation of real-world data requires a large human annotation effort. To address these challenges, current research has focused on two orthogonal but complementary topics: data augmentation and techniques to reduce manual labeling efforts.

Specifically, semantic segmentation produces pixel-wise classification maps where each pixel is assigned a class label (e.g., car, sidewalk, road). Real-world datasets with pixel-wise labels are very expensive to annotate, primarily due to the human effort involved, in addition, human annotations can easily become inconsistent or less accurate after a certain period of time. Consequently, the community is increasingly leveraging the use of computer graphics and simulation to generate synthetic datasets as a practical alternative to simulate acquisitions in the real world, i.e. digital twins or digital reality (see Fig. 1). Synthetic datasets can, not only expedite the annotation process but provide potentially more precise and consistent training data.

A notable challenge in utilizing synthetic data is the presence of a performance gap between models trained exclusively on synthetic data and those trained on real-world data, called domain adaptation (Wilson and Cook, 2020). In this study, we aim to investigate the behavior of synthetic images generated for training semantic segmentation models in the autonomous driving domain (Iglesias-Guitian et al., 2019).

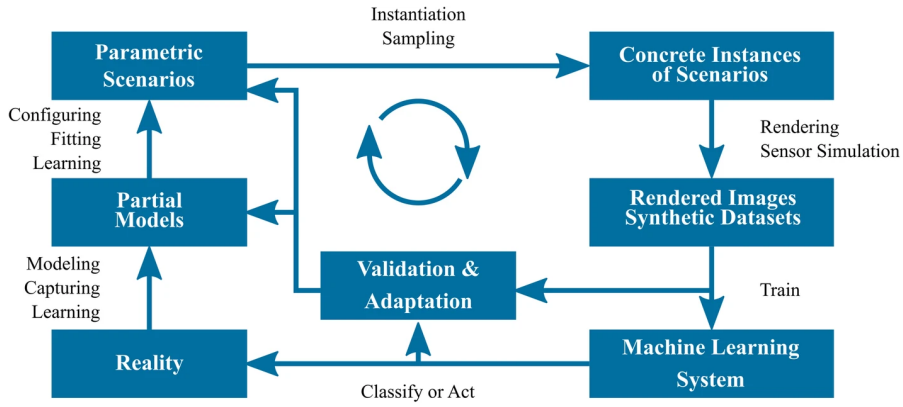


Figure 1: Digital reality approach (Dahmen et al., 2019) for supervised learning from synthetic data.

2 Related Work

Numerous studies have examined the relationship between image quality and the performance of neural network models. The great majority of studies are concerned with how artifacts in real images may affect the training of models for a specific task (Dodge and Karam, 2016; Kamann and Rother, 2020). In the context of real images, Pepik et al. (2015) showed that deep networks are not invariant to certain appearance changes, and demonstrated that the use of rendered data can serve to augment real image datasets. (Kamann and Rother, 2020) addressed the issue of image distortion impact over the semantic segmentation task, incorporating the autonomous driving dataset, Cityscapes (Cordts et al., 2016), alongside other datasets from diverse domains. After subjecting the data to various image distortions and noise, their study concluded that distortion artifacts preserving the fundamental characteristics of textures, such as various types of blur or brightness alterations, have a significantly lower impact than noise types that corrupt textures, like weathering effects or lossy compression. For a deeper and more general discussion about aspects affecting deep model's performance, we recommend the survey by Gawlikowski et al. (2023).

Studies on synthetic images . From the earliest applications of synthetic data for machine learning in autonomous driving (Pomerleau, 1991) to the widespread adoption seen today, synthetic datasets have emerged as a valuable alternative for data augmentation Paulin and Ivasic-Kos (2023).

One of the recurrent questions is how photorealism is really needed to achieve competitive model performance. For example (Movshovitz-Attias et al., 2016) addressed the question of how useful the photo-realistic rendering for synthetic data was for learning the viewpoint estimation task. The authors prepared three versions of a moderately sized synthetic dataset around cars using different levels of realism in terms of materials and lighting. The authors find that synthetic data could help achieve performances close to those of real data and that when mixing the two approaches the result could surpass that of using just real-world data. In addition, they also confirmed that increasing the training set size can help improve performance, but only up to a certain limit, probably due to the lack of variability in the source scenarios. Exploring the advantages of high-quality synthetic datasets, a study by McCormac et al. (2017) concluded that large-scale, high-quality synthetic datasets with task-specific labels could be more advantageous for pre-training models than generic real-world pre-training sources like ImageNet. This conclusion aligns well with previous findings (Tsirikoglou et al., 2017; Zhang et al., 2017) advocating the notion that pre-training models on physically based

rendered datasets with realistic lighting could substantially boost the performance for scene understanding tasks.

Previous works focused on how the photorealism of synthetic data could affect the neural networks, but they did not explore how other parameters used during the generation of the synthetic images could influence neural network's performance. In this sense, Liu et al. (2020) studied how different camera parameters generalize by using either real or synthetic datasets for object detection tasks, particularly cars. Their experiments evaluated pixel size, exposure control, color filters, bit depth, and post-acquisition processing. In the following, we summarized some of the preliminary conclusions drawn from their study. Notably, object size and distance within images showed limited correlation with generalization. Simulating diverse sensor pixel sizes was found to enhance generalization in specific contexts, suggesting its importance. Surprisingly, color filter generalization appeared symmetric, indicating potential cross-sensor applicability. Networks trained at distinct bit-depths displayed limited generalization, emphasizing bit-depth's significance. Exposure control algorithms played an important role, with good generalization observed, primarily affected by global exposure data quality. Finally, network performance exhibited robustness within certain gamma value ranges but deteriorated with significant deviations (e.g., > 0.1). Such preliminary conclusions shed light on potential factors influencing the generalization of the simulated camera parameters.

While previous research has been focused on what impact the synthetic image quality has in object detection tasks, a notable gap still remains in understanding how synthetic data could affect neural network performance in semantic segmentation for autonomous driving (Gómez et al., 2023; Khan et al., 2019).

3 Experiments

In order to address the aforementioned gap, we propose in this work a preliminary set of experiments focused on analyzing three key aspects while generating synthetic data for semantic segmentation: i) the influence of the training set size or number of images; ii) the camera sensor pixel size or image resolution; and iii) the number of samples per pixel(spp) used for path-traced physically-based rendering approaches. Our study aims to provide initial insights that can guide or inspire further research.

Figure 2: Various levels of Monte Carlo noise introduced by path tracing rendering. We compare how different samples per pixel could impact the semantic segmentation task.

To conduct our experiments, we adopt the widely recognized 19-class standard proposed in the Cityscapes dataset Cordts et al. (2016). We chose to have three different sources as synthetic

datasets: GTAV (Richter et al., 2016), Synscapes (Wrenninge and Unger, 2018) and a custom-tailored dataset for which we could control the samples per pixel for rendering. These three datasets will serve as sources, and two real-world datasets will serve as targets: Cityscapes and Mapillary.

Training set size. For this experiment our hypothesis is that, depending on the content and variation included in a synthetic dataset, the number of images required in the training set to achieve the best results for semantic segmentation may not follow the prevailing paradigm of *"more data leads to better performance"*. This phenomenon is particularly relevant when considering carefully crafted synthetic datasets, which may exhibit recurrent patterns or tend to overfit to specific layouts or scenarios. Our experiments tested the progressive reduction of the training set size for two source datasets, Synscapes and GTAV, while validating each of them on two different targets, Mapillary and Cityscapes. Please, check Fig. 3 and Sec. 4 for a detailed discussion.

Camera sensor size. In this experiment our hypothesis is that, depending on the simulated sensor resolution, certain details might be better represented at higher resolutions, while lower resolutions may better capture the general shape of larger objects without getting distracted by smaller details. Depending on the class of objects, different resolutions might obtain different results. For our experiments we tested the original (100%), half (50%), quarter (25%) and eighth (12, 5%) of the image resolution for both GTAV and Synscapes using Cityscapes as target (see Fig. 4).

Samples-per-pixel for path-tracing. With this experiment, we aim to test the performance of a neural network related to the path tracer's noise and the denoiser's smoothing effect. This type of noise is due to the variance of the Monte Carlo (MC) integration method used in path-tracer rendering engines, and it is not present in real images. Additionally, we aim to explore the influence of an AI denoiser, trained to reduce MC variance noise by adaptively smoothing the image and finding a trade-off to also preserve high-frequency details. For these experiments, we use a custom-developed dataset for autonomous driving based on an unbiased GPU path tracer, that offer us the possibility to control the aforementioned parameters. We use 425 images from this custom dataset, and we save renders at 8, 16, 64, 128, 256, and 512 spp, for both denoised and the original MC path traced samples. For a thorough discussion about these results, please see Fig. 4 and Sec. 4.

The framework utilized for all these experiments uses Detectron2's (Yuxin Wu et al., 2019) DeepLabV3+ Chen et al. (2018) semantic segmentation model and is based on the framework presented in (Gómez et al., 2023). We adopted the synth-to-real LAB space alignment defined in that framework to help reduce the synth-to-real domain gap.

4 Results

Training set size. In Figure 3 we illustrate the results about the training set size for both Synscapes and GTAV with two different targets. Basically, while both datasets caused great variance for lower number of images, the trend becomes more stable after 4k images with slight variations. These results seem to indicate that achieving equal or greater performance with a portion of the original dataset is possible.

Camera sensor size. In Figure 4 (left) we show the training set image resolution results for both Synscapes and GTAV using the Cityscapes validation set. Synscapes seems to find a trade-off while varyinf its image resolution, except for cases below one eighth of its original size. On

the contrary, GTAV presents a sort of sweet spot around using half the original resolution. Given the nature of GTAV, this result could support our notion that lower resolutions may help larger objects more visible in the image, resulting in a slightly better result.

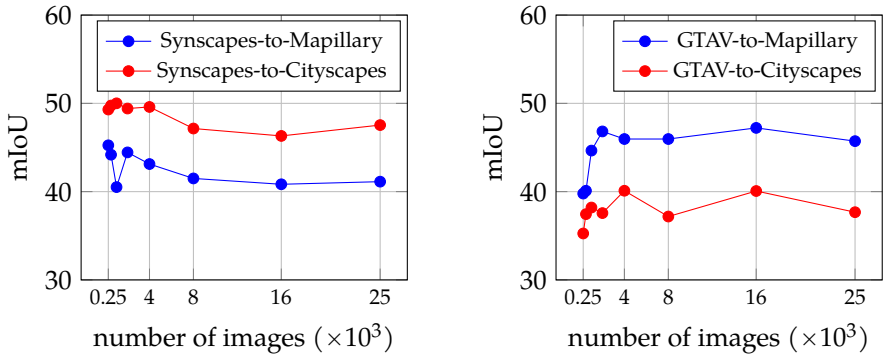


Figure 3: Semantic segmentation (mIoU) w.r.t. the number of images in Synscapes(left), and GTAV(right).

Samples-per-pixel for path-tracing. In Figure 4 (right) we show the performance while varying the number of samples per pixel when generating our custom urban dataset for both the original MC path-traced result and their respective denoised images. The differences between each run are very small and could be explained because of the standard deviation of the experiments. It's worth noting that the MC noise version appears to perform better with fewer samples, while the denoised version appears to function better with more samples. This could be explained by the network preferring the MC noise over the smoothed information supplied by the denoiser at lesser resolution. The denoiser works better at higher resolutions and does not remove as much texture detail as it does at lower levels, which may account for the chart's optimum performance at 256 spp.

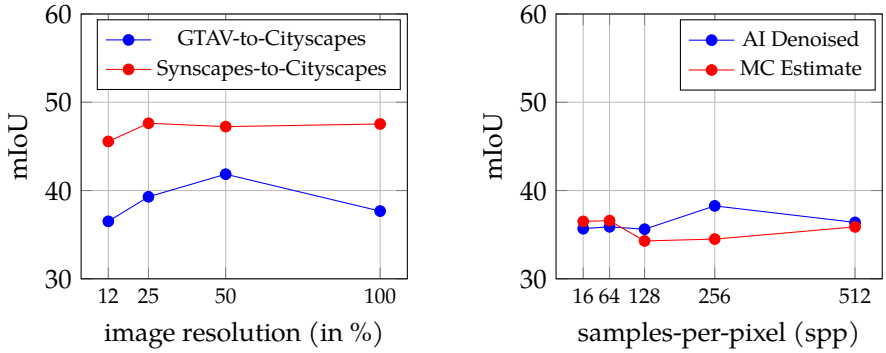


Figure 4: Semantic segmentation (mIoU) w.r.t. image resolution(left) and samples-per-pixel (right). The experiment on spp, used our custom generation process that allow us to control the spp. The two experiments shown here perform their validation using Cityscapes as target.

5 Conclusions

In this study, we reviewed the literature covering the impact that image quality has on neural network performance. We focus on works covering synthetic datasets and performed a preliminary set of experiments to gain new insights around the semantic segmentation task. Our initial findings suggest that similar results for different training set sizes can be obtained, probably due to the lack of variation in common autonomous driving synthetic datasets. Also, different image resolutions may have a positive impact on the performance. And finally, when using Monte Carlo path-tracing, it may be better to use denoising at higher sample rates, while MC noise might be preferable than using a denoiser at lower sampling rates for segmentation networks.

As future work we would like to confirm the initial findings of this work by using more rendering parameters and more experiments studying how the synthetic generation process may boost neural networks' performance.

Acknowledgements

This work has been supported by the Spanish Ministry of Science and Innovation (AEI/PID2020-115734RB-C22). We also want to acknowledge Side Effects Software Inc. for their support to this work. J.A. Iglesias-Guitian also acknowledges the UDC-Inditex InTalent programme, the Ministry of Science and Innovation (AEI/RYC2018-025385-I) and Xunta de Galicia (ED431F 2021/11). CITIC is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS).

Bibliography

L. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 40(04):834–848, apr 2018.

M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele. The cityscapes dataset for semantic urban scene understanding. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3213–3223, Los Alamitos, CA, USA, jun 2016. IEEE Computer Society.

T. Dahmen, P. Trampert, F. Boughorbel, J. Sprenger, M. Klusch, K. Fischer, C. Kübel, and P. Slusallek. Digital reality: a model-based approach to supervised learning from synthetic data. *AI Perspectives*, 1(1):1–12, 2019.

S. Dodge and L. Karam. Understanding how image quality affects deep neural networks. In *2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX)*, pages 1–6, 2016.

J. Gawlikowski, C. R. N. Tassi, M. Ali, J. Lee, M. Humt, J. Feng, A. Kruspe, R. Triebel, P. Jung, R. Roscher, et al. A survey of uncertainty in deep neural networks. *Artificial Intelligence Review*, pages 1–77, 2023.

J. L. Gómez, G. Villalonga, and A. M. López. Co-training for unsupervised domain adaptation of semantic segmentation models. *Sensors*, 23(2):621, Jan 2023.

J. Iglesias-Guitian, G. Ros, V. Kokkevis, J. Alvarez, Y. Lee, and P. Slusallek. Computer graphics for autonomous vehicles. In *ACM SIGGRAPH Frontiers Workshop*. ACM SIGGRAPH, 2019.

J. Janai, F. Güney, A. Behl, and A. Geiger. Computer vision for autonomous vehicles: Problems, datasets and state of the art. *Foundations and Trends® in Computer Graphics and Vision*, 12(1–3): 1–308, 2020.

C. Kamann and C. Rother. Benchmarking the robustness of semantic segmentation models. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8825–8835, Los Alamitos, CA, USA, jun 2020. IEEE Computer Society.

S. Khan, B. Phan, R. Salay, and K. Czarnecki. Procsy: Procedural synthetic dataset generation towards influence factor studies of semantic segmentation networks. In *CVPR workshops*, volume 3, page 4, 2019.

Z. Liu, T. Lian, J. Farrell, and B. A. Wandell. Neural network generalization: The impact of camera parameters. *IEEE Access*, 8:10443–10454, 2020.

J. McCormac, A. Handa, S. Leutenegger, and A. J. Davison. Scenenet rgb-d: Can 5m synthetic images beat generic imagenet pre-training on indoor segmentation? In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2678–2687, 2017.

Y. Movshovitz-Attias, T. Kanade, and Y. Sheikh. How useful is photo-realistic rendering for visual learning? *ArXiv*, abs/1603.08152, 2016. URL *https://api.semanticscholar.org/CorpusID:12684994*.

G. Paulin and M. Ivasic-Kos. Review and analysis of synthetic dataset generation methods and techniques for application in computer vision. *Artificial Intelligence Review*, pages 1–45, 2023.

B. Pepik, R. Benenson, T. Ritschel, and B. Schiele. What is holding back convnets for detection? In *Pattern Recognition: 37th German Conference, GCPR 2015, Aachen, Germany, October 7-10, 2015, Proceedings 37*, pages 517–528. Springer, 2015.

D. A. Pomerleau. Efficient training of artificial neural networks for autonomous navigation. *Neural computation*, 3(1):88–97, 1991.

S. R. Richter, V. Vineet, S. Roth, and V. Koltun. Playing for data: Ground truth from computer games. In B. Leibe, J. Matas, N. Sebe, and M. Welling, editors, *Computer Vision – ECCV 2016*, pages 102–118, Cham, 2016. Springer International Publishing. ISBN 978-3-319-46475-6.

A. Tsirikoglou, J. Kronander, M. Wrenninge, and J. Unger. Procedural modeling and physically based rendering for synthetic data generation in automotive applications. *arXiv preprint arXiv:1710.06270*, 2017.

G. Wilson and D. J. Cook. A survey of unsupervised deep domain adaptation. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 11(5):1–46, 2020.

M. Wrenninge and J. Unger. Synscapes: A photorealistic synthetic dataset for street scene parsing. *CoRR*, abs/1810.08705, 2018. URL *http://arxiv.org/abs/1810.08705*.

A. K. Yuxin Wu, F. Massa, W.-Y. Lo, and R. Girshick. Detectron2, 2019. URL *https://github.com/github-linguist/linguist*.

Y. Zhang, S. Song, E. Yumer, M. Savva, J. Lee, H. Jin, and T. Funkhouser. Physically-based rendering for indoor scene understanding using convolutional neural networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5057–5065, Los Alamitos, CA, USA, jul 2017. IEEE Computer Society.

Use of Machine Learning Algorithms for Network Traffic Classification

Adrián Nieto Antelo, Diego Fernández Iglesias, and Francisco J. Nóvoa

Facultad de Informática, Universidade da Coruña, 15071, A Coruña, España
Centro de Investigación CITIC, Departamento de Ciencias de la Computación y Tecnologías de la Información, Universidade da Coruña, 15071, A Coruña, España
Correspondence: adrian.nieto1@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: In recent years, the complexity of threats utilizing the network as an attack vector has significantly increased. Traditional attack prevention and detection systems (IPS/IDS) based on signatures do not provide an acceptable level of security for many organizations.
Furthermore, the volume of traffic on corporate networks has also grown exponentially, while quality of service requirements do not always allow for deep inspection (at the application layer) of packets.
The main objective of this work is to demonstrate that the application of machine learning techniques to the information of data flows circulating through the network allows for the satisfactory detection of malicious traffic. Specifically, this work is developed within an emerging network paradigm, such as software-defined networks.

1 Introduction

It is of vital importance to offer the highest possible protection to defend ourselves against all the current computer threats. Currently, our protection mainly relies on methods based on policies and rules. These methods have certain issues, such as human errors in rule configuration, response time to threats, and adaptability. That's why Artificial Intelligence, especially Machine Learning, is gaining more weight in cybersecurity.

This work aims to compare and understand different machine learning techniques used in the classification of network traffic on an SDN (Software-Defined Network). To achieve this, a methodology known as CRISP-DM (Cross-Industry Standard Process for Data Mining) has been employed Wirth and Hipp (2000). This article will be structured according to this methodology, with each section representing a different step in it. However, for this article, the business understanding section has been omitted, as it involves the theoretical explanation of the concepts that will be used throughout the work.

2 Understanding the data

This phase involves the collection, description, exploration, and verification of the data to be used.

The dataset used is one that already existed previously and was not generated specifically for this work, as creating a dataset is not one of the objectives of this project. In this case, the InSDN dataset Elsayed et al. (2020) is used.

This dataset presents various attack scenarios from different sources, both external and internal, that can affect an SDN network. It includes various attack scenarios, such as DoS attacks,

DDoS attacks, password guessing attacks, web application attacks, probe attacks, botnet attacks, and U2R attacks. The dataset is divided into three groups based on the type of traffic and target machines. The first group includes normal traffic. The second group includes attacks targeted at the Metasploitable2 server, which is a virtual machine server. The third group consists of attacks on the OVS machine, which is a switch.

After loading the dataset, exploration is conducted using a Pandas dataframe specifically created to work with this data. Among these explorations, the correlation matrix stands out, which will be used in the next step.

3 Data preparation

In this phase, the selection of features and samples to be used in the Machine Learning algorithm will be carried out. The quality of the data will be checked, formatted, modified, or new data will be created as necessary.

The data from the previous dataset are not in an appropriate format for training, so they need to be prepared for use in training.

The first step is to remove outliers that may be present in one of the dataset's features. The Isolation Forest algorithm, which is based on decision trees, was used for this task. After applying the algorithm, it was decided to remove values from four features in the dataset, namely Tot Fwd Pkts, TotLen Fwd Pkts, TotLen Bwd Pkts, and Bwd Header Len.

Next, modifications to existing features were made to provide more information. This includes the creation of the target class, which divides the data flows into normal and attack traffic. The Timestamp variable was also modified to obtain only the hour, and the source and destination ports of the flows were grouped into categories defined by IANA IANA (2023). In the case of the hour, as it contains a cyclical character, it is necessary to transform this variable in a way that algorithms can understand this cyclical nature. Therefore, a circular representation of the hour is used, where the value 0 appears immediately after 23. To achieve this, sine and cosine functions are used.

Afterward, the data was formatted because two of the algorithms used (SVM and Logistic Regression) require their features to be numeric values, both integers and floating-point values. Specifically, five object-type attributes were formatted. The source and destination IP addresses were encoded using Label Encoding, similar to what was done with the ports.

Following this, standardization was performed using the scikit-learn StandardScaler library. This process is necessary because algorithms like SVM are sensitive to the scale and magnitude of variance in the data, which can influence the model. This is because they cannot interpret the meaning of the data; they only see numbers. If the predictors are not scaled equally, features with a larger scale or more variance can have a greater influence on the final prediction.

The last step is data selection, as the database is extensive and optimization of performance and accuracy is needed. First, features that do not provide information were removed, whether they take only one value or have already been used to create other more relevant features. For this, the values from the correlation matrix, calculated in the previous step, were used to eliminate features that are correlated both directly and indirectly with others, removing the less important one after calculating its importance using Random Forest. Finally, from the resulting features, only a subset was chosen for model creation, which was done using the RFECV algorithm. To use this algorithm, a decision tree classifier was chosen as the estimator, Stratified K-fold as the cross-validation strategy, accuracy as the scoring metric, and a step size of one.

In the end, a dataset with 31 features was obtained.

4 Modeling

This phase involves selecting the algorithms to use, generating the design test, building the models, and evaluating them.

First, the methods to assess the quality and validity of the model will be established. In this case, two methods will be used: cross-validation and certain evaluation metrics, such as those obtained from the confusion matrix.

For designing the models, an approach based on Stratified K-fold Cross Validation was used. This was implemented through two nested cross-validation loops. The value of k was set to 5 for both the outer and inner loops. These values are standard and were chosen because they yield good results within a reasonable time frame.

The outer loop performs the initial partition, dedicating 20% of the dataset to the test set in each iteration. The inner loop takes the remaining 80% of the original dataset and divides it into five parts. One of these partitions is used for validation, while the rest are used for training. This means that 16% of the dataset is used for validation, and 64% for training.

For the inner loop, a scikit-learn class called GridSearchCV was used, allowing for hyperparameter tuning. GridSearchCV enables the comparison of different parameters to find the optimal ones. However, when using an outer loop, it provides 5 parameter combinations, which can be the same or different. At the end of all outer loop executions, the set of parameters that yielded the best results is obtained, but this value corresponds to a single execution of the outer loop. In other words, one combination may perform best in one run, while performing poorly in the other four. For this reason, a measurement has been devised to calculate the best parameters using information from all 5 iterations of the outer loop. This measurement takes into account both the mean and the standard deviation. A lower standard deviation indicates less variability, so the results tend to be more consistent. To use these two values and obtain the best possible metric, the decision was made to divide the mean by the standard deviation. This causes high means and low standard deviations to yield higher results. A higher final result indicates that the parameters generally yield better and more consistent results.

The following three algorithms were used: SVM, Random Forest, and Logistic Regression.

4.1 SVM

For this algorithm, the parameters C and Tol have been configured. C is the parameter that adjusts the hyperplane separating the two classes, and Tol specifies the tolerance for stopping criteria. In the end, the values chosen are 1 for the C parameter and 0.0001 for Tol.

4.2 Random Forest

For this algorithm, the parameters "n_estimators," which indicates the number of decision trees to be created in the forest, and "max_depth," which is the maximum depth a tree can have, have been configured. In this case, the chosen values are "max_depth" equal to 9 and "n_estimators" equal to 50.

4.3 Logistic Regression

For this algorithm, the parameters "Penalty," which specifies the penalty norm, "C," which is the inverse of the regularization strength, and "Tol," which specifies the tolerance for stopping criteria, have been configured. The chosen values are C=1, penalty=l2, and tol=0.0001.

5 Evaluation

In this phase, the results are evaluated, the process is reviewed, and the next steps to be taken are determined. To evaluate the results, cross-validation using a test set will be employed, and metrics derived from the obtained confusion matrices will be utilized. The results can be seen in Table 1.

As observed, Random Forest is the algorithm that yields the best results, and it will be the one we deploy.

Table 1: Métricas obtenidas a partir de las matrices de confusión.

	Precisión	Exactitud	Sensibilidad	F1-score
RL	0.893330	0.657411	0.968729	0.783270
RF	0.999952	0.999951	0.999805	0.999878
SVM	0.908813	0.707398	0.924062	0.801343

6 Deployment

Finally, in this phase, the model is implemented, maintenance and monitoring are conducted, a final report is produced, and the entire project is reviewed.

In this phase, the obtained model will be tested in an environment where a controller changes the behavior of a switch to block a flow if it has been classified as an attack by the model. Only the algorithm with the best results, which is Random Forest, will be used for this purpose.

For the deployment, two machines have been used, one running the network topology, while the other runs the network traffic controller.

The first step in this section is to create the environment in which the models will be tested. This will be an SDN that uses the OpenFlow protocol for communication between the controller and the switch. The tool used to create it is known as Mininet. Afterward, the controller to be used is chosen, which will be Ryu, and the application to be used within it. In this case, we have chosen the "simple_switch_13" template, which uses OpenFlow 1.3. This template allows defining the switch's behavior so that it can rewrite packet addresses, transfer packets from a specified port, transfer received packets to the controller, and transfer packets forwarded by the controller from a specified port Ryubook (2014).

When implementing the model, the first thing to consider is how to obtain the information corresponding to the attributes used by the model for classification. This is where our first problem arises. Obtaining this information in our test environment is not straightforward and would involve excessive complexity. Therefore, an alternative solution has been proposed. This solution involves retrieving this information from the dataset itself. This is possible because we are going to simulate the data flows that make up that database, as the .pcap files that allow simulation are available. The available traffic in a .pcap file will be injected into the Mininet network. After this, the FlowID of the flow will be obtained, which is the characteristic that identifies the flow, and it will be searched for in the dataset by this ID. Once found, the information corresponding to the attributes required for the model will be saved.

After implementing the template, it is monitored and its operation is displayed. To achieve this, the "simple_monitor_13" template has been modified to display information about the flow table on the screen.

However, this monitor has the problem that adjusting the time between messages is difficult, which can make the controller's screen unreadable at times. Additionally, it was not possible to display the output of the model because this information is not stored in the flow table. For this reason, the table will be printed after the processing of a flow, which leaves a cleaner interface and simplifies the code.

To observe the controller's behavior on a flow, two .pcap files belonging to the InSDN dataset will be injected, one corresponding to an attack flow and another containing normal traffic. To inject both files, tcpreplay has been used.

7 Conclusions

The main conclusions obtained is that it has been possible to deploy an application in the controller that is capable of classifying traffic and modifying the operation of an SDN switch depending on the response obtained, using machine learning. This machine learning-based security implementation would be possible to implement by varying certain steps performed in

this work, such as performing the deployment phase in a real environment instead of a test environment. In this environment, it would be possible to capture the traffic flowing over the network over a period of time, and then build the model using that more specific network data, rather than searching the dataset.

Acknowledgements

This research has been funded by the Spanish Ministry of Economy and Competitiveness and European Union ERDF funds (Project PID2019-111388GB-I00). CITIC is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS).

Bibliography

M. S. Elsayed, N.-A. Le-Khac, and A. D.Jurcut. Insdn: A novel sdn intrusion dataset. *https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9187858*, 2020. [Online; accessed 14-September-2023].

IANA. Service name and transport protocol port number registry. *https://www.iana.org/assignments/service-names-port-numbers/service-names-port-numbers.xhtml*, 2023. [Online; accessed 14-September-2023].

Ryubook. Switching hub. *https://osrg.github.io/ryu-book/en/html/switching_hub.html*, 2014. [Online; accessed 14-September-2023].

R. Wirth and J. Hipp. Crisp-dm: Towards a standard process model for data mining. *Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining*, 2000.

Solution for Capturing Data from Wearable Devices

Eva Pardo-Otero, Iago Fernández-Garrido, Santiago Fernández-Noriega-Balseiro, María Martínez-Pérez, and Laura Nieto-Riveiro

Universidade da Coruña, CITIC, LIA[2] Group, Faculty of Informatics, Campus de Elviña, 15071, A Coruña, Spain.
Universidade da Coruña, CITIC, TALIONIS Group, Faculty of Health Sciences, Campus de Oza, 15071, A Coruña, Spain.
Correspondence: eva.pardo@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: Thanks to the rise of wearable devices, people have more direct access to a variety of health data, such as physical activity, sleep and heart rate. For the research field, these devices represent a powerful tool for monitoring and evaluating different parameters. However, the procedure of capturing data for storage in an independent and self-managed database is not standardised. In this project we analysed two methods of data capture for the Xiaomi Mi Bands. One uses the official application together with Google Fit and the other uses the open source application GadgetBridge. The advantages and disadvantages of each system were studied, concluding that both could be very beneficial as data capture solutions for wearable devices in research, although with different target projects due to their particularities. Future work will explore these systems in more depth, addressing limitations, automation and optimising for specific research needs.

1 Introduction

Wearable devices have significantly transformed the field of healthcare, becoming essential tools for personal health monitoring and, increasingly, for biomedical research (Huhn et al., 2022). In the case of activity wristbands, they offer a reliable solution to automatically track health data throughout the day and store it in the cloud. This introduces a novel avenue for monitoring patients' health status within the healthcare sector, potencially allowing, for instance, the automatic detection of health problems (Beniczky et al., 2021; Miranda-duro et al., 2021), the continuous tracking of a patient's progress during treatment (MacEira-Elvira et al., 2019) or the evaluation of sleep patterns to assess sleep quality (Concheiro-Moscoso et al., 2023). One of the key issues when using wearables in research is access to data (Huhn et al., 2022). This aspect is one of the main challenges of this technology (Muzny et al., 2020). Most applications designed to syncing with these devices use cloud-based storage services and do not facilitate data sharing. They also export processed information, making it difficult to perform in-depth analysis with raw data. In addition, providers often employ their own data schemas and use non-standardized, closed transfer protocols (Muzny et al., 2020). All in all, this limits the study and integration of health data and prevents the full potential of these devices from being exploited.

In this context, we present two approaches to retrieve data automatically from fitness trackers, process it and store it in an independent and self-managed database.

2 Objectives

The main objective of this work is to develop an automatic system for capturing and managing data from Xiaomi Mi Band wristbands for its application in research projects. To this end, we have defined the following specific objectives: 1) to investigate the different forms of access to the data storage location, 2) to convert the captured data into usable information, and 3) to assess the advantages and disadvantages of the analysed systems.

3 Methodology

In this section we first introduce the wristband model we use and its official application, as well as its limitations regarding data access. Subsequently, we describe two alternative methodologies for data capture: accessing cloud data via Google Fit (Google LLC, 2014) and accessing local data via a separate open-source application, GadgetBridge (Gadgetbridge contributors, 2021).

3.1 Xiaomi

Both methods focus on the Xiaomi Mi Band devices (Xiaomi, 2023), one of the most popular and affordable fitness trackers on the market. These devices use sensors to track information such as activity intensity, heart rate, blood oxygen saturation, sleep, training sessions, etc. The data is then stored in ZeppLife (Zepp Health, 2014), one of the official apps provided by the manufacturer for visualisation.

The data is stored in the cloud, on the servers of the provider Huami (Huami, 2023), every time it is synchronised with ZeppLife. It is also stored locally on the mobile device in a restricted access folder, which could only be accessed with a rooted phone. ZeppLife allows exporting the data, but manually and slowly, and already processed in separate CSV files.

3.2 Google Fit

The first method involves the Google's Fit REST API (Google Developers contributors, 2023). This API, however, is not supposed to connect to any sensor or wearable. Thus, it is intended to access user data in the fitness store (Nobakht et al., 2020). To use it, it is necessary for each user to have the Google Fit application and a Gmail account. In this case, when using the Xiaomi Mi Smart Band 7, it is also necessary to have the ZeppLife application installed, as the entity responsible for providing the data to Google.

The user synchronizes both applications, enabling Google Fit to read the data stored in ZeppLife and manage it independently. To make this possible, the use of OAuth 2.0 is required, an authorization protocol that enables a website or application to access resources hosted by other web applications on behalf of a user. In our case, we utilize a temporary access token generated by this tool and associated with the provided Gmail account. As a result, the information recorded by the wristband is accessible through a series of REST queries launched against Google's databases. More specifically, we access recorded data related to:

- Physical activity based on the number of steps.
- The duration of each sleep phase (REM, light sleep, deep sleep, or awake) for every sleep session.
- Heart rate in beats per minute.
- The body measurements (weight and height) that the user records.

This requests return the raw data in JSON format, according to the grouping conditions we specify, for then to be processed in our server and stored in our own database in the desired format. A thorough view of the system is displayed in Figure 1.

The reason of using this Google API is its compatibility with almost any type of wearable available on the market, the reliability of having data replicated in an external and trusted database, the API's compatibility with Google's data model and its user-friendly simplicity both for end-users and programming.

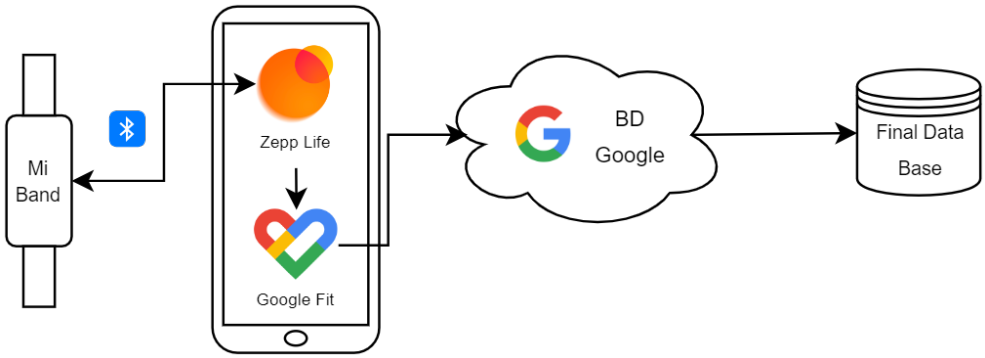


Figure 1: Diagram of the Google Fit-Based Capture System

3.3 Gadgetbridge

The second method is to use the open-source application GadgetBridge as an alternative to ZeppLife. It does not require account registration and data storage on the vendor's servers, instead the data is stored locally on the device in an SQLite database. The steps necessary to implement this system are depicted in Figure 2.

Firstly, the wristbands must be linked to Gadgetbridge. Wristbands communicate with apps via Bluetooth Low Energy (BLE), which uses key-based pairing protocols to establish secure connection sessions. Depending on the model of the wristband there are two different options for pairing. For older models such as the Mi Band 2/3, when GadgetBridge scans the device and orders the pairing, the key is sent without protection from the app to the wristband, and the user confirms the pairing (Casagrande et al., 2022). However, for new models (tested on Mi Band 5/7) the pairing is mediated by Huami's servers. The server signs a pairing key generated from the wristband, based on a random number and the Bluetooth address. The wristband then verifies this signature to establish the pairing, which is confirmed by the user as above (Casagrande et al., 2022). This implies that it is first required to link the wristband to ZeppLife, then obtain this pairing key and use it to link to GadgetBridge, one or more wristbands, as it supports multi-device connection.

The key is stored in a database located in the restricted folder of the ZeppLife application. Hence, an own application was developed (Android Studio Giraffe 2022.3.1) to automatically access this location by activating superuser permissions on a rooted smartphone, and then through an SQL query retrieve the key. Then, the key is inserted manually into GadgetBridge on a different phone and the wristband is permanently paired.

Secondly, once the wristbands are paired, the data can be synchronised and exported from the GadgetBridge application itself. It allows the option to auto-export the data at least every 1 hour and to select the desired destination directory. It is relevant to mention that GadgetBridge has integrated an API to control the Bluetooth connection, synchronisation, and data export via intents (messages for communication between application components).

Regarding the data, with this system an SQLite database is obtained, with multiple tables created for every type of wearable supported by GadgetBridge. Each row represents one record per minute. The registered variables are:

- TIMESTAMP: time in Unix format (number of seconds since 00:00:00 UTC 1 January 1970).
- DEVICE_ID: identifier of the connected device.
- HEART_RATE: beats per minute; when not measured records the value 255.
- RAW_INTENSITY: related to the amount of movement, probably data from the accelerometer in the wristband.
- STEPS.
- RAW_KIND: discrete, device-specific values that are sent processed from the wristband and collected by GadgetBridge. They indicate different situations, for example for the Mi Band 7, the value 155 appears when the wristband is not worn, or the value 120 when there is sleep.
- SLEEP, DEEP_SLEEP, REM_SLEEP: coded sleep data, as in the previous case. GadgetBridge developers set thresholds roughly so that the sleep phases correspond as closely as possible to those shown in ZeppLife (personal communication, GadgetBridge developers).

At the moment, for the Mi Band 5, the registration of sleep variables in the database has not been implemented, although the wristband does record them (personal communication, GadgetBridge developers). Therefore, its data is stored, together with the Mi Band 3 data, in a separate table than the Mi Band 7.

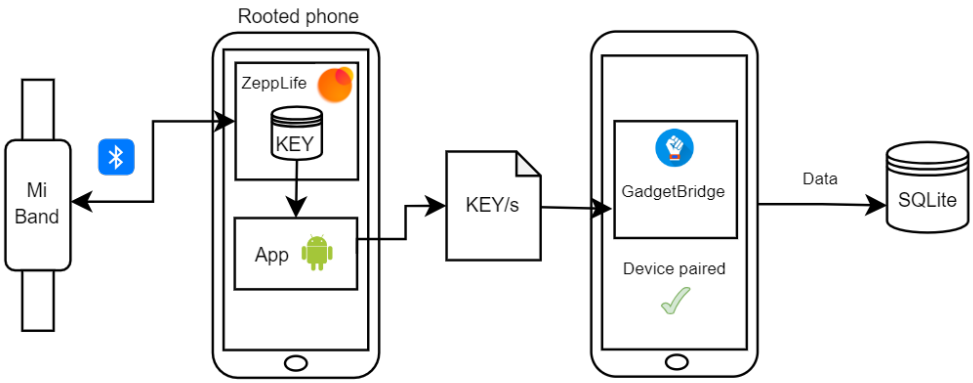


Figure 2: Diagram of the GadgetBridge-Based Capture System

4 Discussion

Wearable devices are increasingly common among the population offering significant advantages, since the interpretation of this data can be beneficial for the user's health (Beniczky et al., 2021; Concheiro-Moscoso et al., 2023; MacEira-Elvira et al., 2019; Miranda-duro et al., 2021). It is also an interesting topic in research as it allows access to a larger amount of data at relatively low costs (Huhn et al., 2022). For this reason, we proposed two different alternatives for data extraction for research use. Both technologies have been studied, since each one has a way of operating that can be adapted to different requirements. Table 1 provides a summary of the advantages and disadvantages of each of the approaches.

Implementing through the Google's Fit API makes data extraction independent of the wearable device manufacturer and, consequently, of the application they offer for data storage and visualization. If Google Fit can be downloaded on the mobile device and connected to the wearable application, the recorded data is accesible. Furthermore, there is a copy of the information stored in our database hosted on Google servers, and in case any issues arise, the data could be recovered .

On the other hand, using Google Fit as a data source implies the dependence on both device compatibility and its speed and reliability. Although the list of devices compatible with Google Fit is extensive, it is possible that some devices with older software versions may not be compatible with the application. Additionally, the data is only accessible once it has been uploaded to Google's servers, and there is no control over that operation.

The system based on GadgetBridge offers an independent alternative to the official app for data capture. The main advantage of this lies in the improved privacy, as the data remains exclusively on the device, without being stored in the cloud. However, it should be noted that it is not 100% independent from the official app ZeppLife, as it requires obtaining pairing keys from its database, at least with recent models of Mi Band. Although we have solved the problem of obtaining the key, it would be recommendable to test it on more types of wearables. In addition, the server-based type of pairing implies that the two applications are not compatible on the same smartphone. These limitations could be minimised by analysing the vulnerabilities of the official applications and the connection protocols they use. Recently, Casagrande et al. (2022) succeeded in impersonating any Xiaomi fitness tracker and companion app such as ZeppLife.

Regarding the independence of wearable devices, unlike the Google Fit method, GadgetBridge only works with the wearables models it has implemented, although they currently cover most of the most popular ones. In relation to this, it should be noted that the fact that it is an open source application allows for its modification, which has already been done in other works (Grossi et al., 2021; Ramachandran et al., 2020).

Table 1: Comparison between methodologies based on Google Fit and Gadgetbridge

	Google Fit	Gadgetbridge
Independence of the Wearable	✓	✗
Independence from External Application	✗	✓
Cloud Backup	✓	✗
Account Registration Required	✓	✗
Open Source	✗	✓
No Third-party Data Readings	✗	✓

5 Conclusions

Despite having identified certain limitations in each of the systems analysed, we consider that the benefits outweigh the associated disadvantages. Therefore, both methodologies can be potentially highly beneficial as data capture systems for wearable devices in the context of research projects, especially by improving automation and minimising potential failures. In addtion, we believe that each can be targeted at different types of projects due to their particularities.

Regarding the Google Fit system, it could be suitable for projects looking for broad compatibility with different wearable devices. Its simplicity of use and familiarity to users also make it an attractive option for projects focused on user convenience. Furthermore, it allows for greater scalability, making it an attractive option for projects anticipating significant growth. In the case of GadgetBridge, it is best suited for projects that place a high priority on data privacy or those

that require the simultaneous management and synchronization of multiple wearable devices. Furthermore, in order to fully exploit the potential of this data capture system, it would be advisable to conduct a dedicated study aimed at decoding the encoded variables, both "raw_kind" and those related to sleep. Such an investigation could enhance the usability and understanding of the collected data.

In the future, we intend to explore these systems more deeply, focusing on addressing identified limitations, improving automation, and optimizing their usage for specific research needs.

Acknowledgments

Funding: CITIC is financed by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS). The research carried out by TALIONIS group is financed by the Xunta de Galicia (Aid from the Consellería de Cultura, Educación, Formación Profesional e Universidades for the consolidation and structuring of competitive research units) ED431B 2022/39. The publication is part of the project TED2021-130127A-I00, funded by MCIN/AEI/10.13039/501100011033 and by the European Union "NextGenerationEU"/PRTR.

Bibliography

S. Beniczky, S. Wiebe, J. Jeppesen, W. O. Tatum, M. Brazdil, Y. Wang, S. T. Herman, and P. Ryvlin. Automated seizure detection using wearable devices: A clinical practice guideline of the international league against epilepsy and the international federation of clinical neurophysiology. *Clinical Neurophysiology*, 132:1173–1184, 2021.

M. Casagrande, E. Losiouk, M. Conti, M. Payer, and D. Antonioli. Breakmi: Reversing, exploiting and fixing xiaomi fitness tracking ecosystem. *IACR Transactions on Cryptographic Hardware and Embedded Systems*, 2022:330–366, 2022.

P. Concheiro-Moscoso, B. Groba, D. Alvarez-Estevez, M. d. C. Miranda-Duro, T. Pousada, L. Nieto-Riveiro, F. J. Mejuto-Muiño, and J. Pereira. Quality of sleep data validation from the xiaomi mi band 5 against polysomnography: Comparison study. *Journal of Medical Internet Research*, 25:e42073, 2023.

Gadgetbridge contributors. Gadgetbridge, 2021. URL *https://github.com/Freeyourgadget/Gadgetbridge*. Mobile application software.

Google Developers contributors. Google fit api — Google Developers, inc. *https://developers.google.com/fit/rest?hl=es-419*, 2023. [Online; accessed 06-09-2023].

Google LLC. Google fit, 2014. URL *https://play.google.com/store/apps/details?id=com.google.android.apps.fitness&hl=en&gl=US*. Mobile application software.

N. R. Grossi, B. Batinic, and S. Moharitsch. Sleep and health: examining the relation of sleep to burnout and well-being using a consumer fitness tracker. *Health and Technology*, 11:1247–1257, 2021.

Huami. Huami. *https://www.huami.com/*, 2023. [Online; accessed 06-09-2023].

S. Huhn, M. Axt, H. C. Gunga, M. Anna Maggioni, S. Munga, D. Obor, A. Sié, V. Boudo, A. Bunker, R. Sauerborn, T. Bärnighausen, and S. Barteit. The impact of wearable technologies in health research: Scoping review. *JMIR Mhealth Uhealth*, 10:e34384, 2022.

P. MacEira-Elvira, T. Popa, A. C. Schmid, and F. C. Hummel. Wearable technology in stroke rehabilitation: Towards improved diagnosis and treatment of upper-limb motor impairment. *Journal of NeuroEngineering and Rehabilitation*, 16:1–18, 2019.

M. D. C. Miranda-duro, L. Nieto-Riveiro, P. Concheiro-Moscoso, B. Groba, T. Pousada, N. Canosa, and J. Pereira. Analysis of older adults in spanish care facilities, risk of falling and daily activity using xiaomi mi band 2. *Sensors*, 21:3341, 2021.

M. Muzny, A. Henriksen, A. Giordanengo, J. Muzik, A. Grøttland, H. Blixgård, G. Hartvigsen, and E. Årsand. Wearable sensors with possibilities for data exchange: Analyzing status and needs of different actors in mobile health monitoring systems. *International Journal of Medical Informatics*, 133:104017, 2020.

M. Nobakht, Y. Sui, A. Seneviratne, and W. Hu. Pgfit: Static permission analysis of health and fitness apps in iot programming frameworks. *Journal of Network and Computer Applications*, 152:102509, 2020.

M. Ramachandran, N. Chowdhury, A. Third, Z. Jan, C. Valentine, and J. Domingue. A framework for handling internet of things data with confidentiality and blockchain support. In *IoT4Safe, in conj. with Extended Semantic Web Conference (ESWC) 2020*, 2020.

Xiaomi. Xiaomi Mi Band. *https://www.mi.com/es/search/Mi%20Band?tab=product*, 2023. [Online; accessed 06-09-2023].

Zepp Health. Zepp life, 2014. URL *https://play.google.com/store/apps/details?id=com.huami.watch.hmwatchmanager&hl=en&gl=US*. Mobile application software.

Improving Authentication in the Amazon Alexa Virtual Assistant by Using a Geofence

Jorge Fernández-García, Martiño Rivera-Dourado, Rubén Pérez-Jove, Cristian R. Munteanu, and Jose Vázquez-Naya

Grupo RNASA-IMEDIR, Facultade de Informática, Universidade da Coruña, Campus de Elviña, A Coruña, 15071, Spain
Centro de Investigación CITIC, Universidade da Coruña, 15071 A Coruña, Spain
IKERDATA S.L., ZITEK, University of Basque Country UPVEHU, Rectorate Building, Leioa, 48940, Spain
Correspondence: jorge.fgarcia@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: Amazon Alexa processes voice commands as input to help users perform tasks. For protecting this commands, Amazon Alexa implements some security measures. These security measures, such as voice recognition and user's PIN, do not have the ability to mitigate replay attacks. In order to mitigate replay attacks, in this paper, we propose an authentication method based on Geofencing, consisting of (1) an Android application and (2) an Alexa Skill. By using the Android application, the user is able to configure a geofence near the Amazon Echo smart speaker. The developed Alexa Skill only accepts requests when the user is within the established geofence. This method mitigates replay attacks: an attacker could only try to use a replay attack when the legitimate user is close to the speaker, making it unfeasible.

1 Introduction

Amazon Alexa is a virtual assistant that processes voice commands quickly and efficiently. It provides basic functionalities and allows the extension of its capabilities through the creation of third-party functionalities, known as Alexa Skills (Amazon Alexa Skills, 2016). This is made possible by the development environment provided by Amazon (Amazon, 2023).

This virtual assistant has security measures such as voice recognition and user's PIN (Alattar et al., 2023), (Amazon, 2020). These security measures are used to process some sensitive requests, such as an online purchase. In particular, the user's PIN must be spoken aloud so that Amazon Alexa can authenticate the user and accept or reject the sensitive requests.

Despite these security measures, Amazon Alexa has a number of vulnerabilities published by INCIBE (Incibe, 2023). Some vulnerabilities to highlight are: the processing of voice commands at high frequency spectrum, the existence of backdoors and impersonation. Consequently, due to these vulnerabilities, differents attacks arise such as: dolphin attacks (Zhang et al., 2017), voice squatting (Zhang et al., 2018), voice masquerading (Zhang et al., 2018) and replay attacks (Malik et al., 2019).

As a result, these attacks can bypass the security measures, specifically replay attacks. Replay attacks consist of recording and replaying a victim's voice command to impersonate his identity. For example, attackers could record a user's spoken PIN and then replay it for Amazon Alexa to process the sensitive request.

In this paper, we developed an authentication method based on Geofencing to protect requests in Amazon Alexa, mitigating replay attacks that were previously possible. We propose

a Proof of Concept (PoC) where Geofencing technology is used to improve the authentication on requests. Geofencing uses GPS, Wi-Fi or Bluetooth to create virtual geographic barriers, named geofence, that monitor the location of a physical device or user, with the aim of detecting when the user enters or exits the geofence (Rahate and Shaikh, 2016). In our PoC, the geofence is configured at the location of the smart speaker using a developed Android application. When the user is inside the geofence, Amazon Alexa, through a developed Alexa Skill, processes the user's requests. Otherwise, it rejects them. With this approach, if an attacker tries to perform a replay attack, he needs the legitimate user to be inside the geofence, making this attack unworkable.

2 Material and methods

This paper seeks to create an authentication method that protect requests in Amazon Alexa using Geofencing technology. To achieve this, we employed a personal computer running the Windows 10 operating system with internet connection. Subsequently, we used the Android Studio IDE to develop the Android mobile application and tested it on a Samsung J7 2016 Android mobile device. Finally, we created accounts on Amazon Developer and Amazon Web Services (AWS, 2020) to develop the Alexa Skill and utilizing Amazon services.

We started this research by analyzing the security measures and vulnerabilities of Amazon Alexa. Next, we looked at what types of Skills exist in the market and how to develop them. Then, we searched for information to communicate Amazon Alexa and Android technology through Amazon services (AWS, 2020).

For software development, we adopted the agile Scrum methodology. This allowed us to work in Sprints and have different functional versions of the software. We began with a basic version of the application, which was refined during each Sprint until we reached the final product.

3 Development

To demonstrate how to improve authentication in Amazon Alexa, the development of a PoC was carried out. It is important to highlight three main blocks of this PoC: (1) the analysis and design of the authentication method based on Geofencing, (2) the development of an Android application for configuring the Geofencing technology and (3) the development of the Alexa Skill responsible for accepting or denying the requests. The following is a brief explanation of the three main blocks.

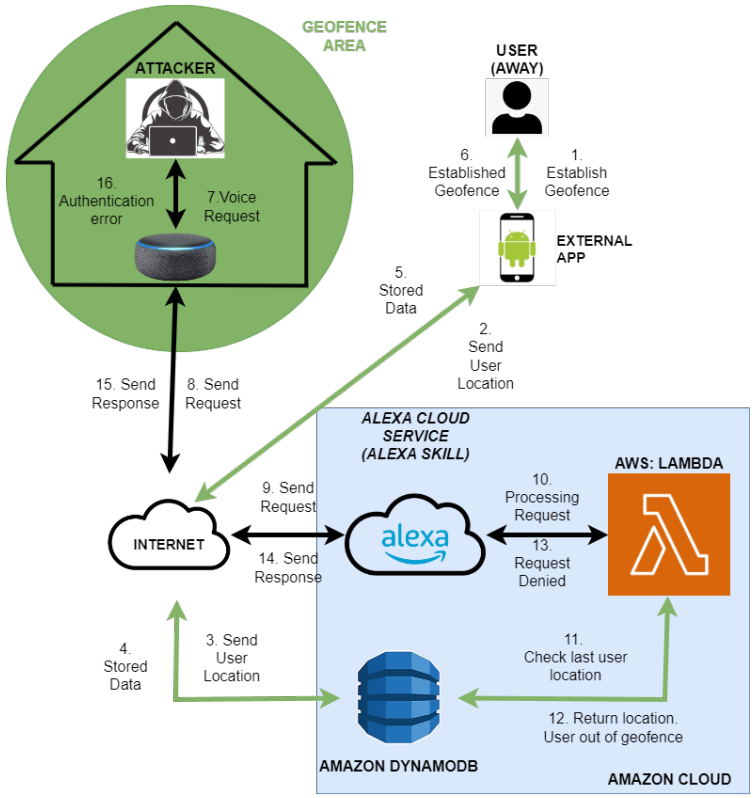


Figure 1: Request flow in PoC with the authentication method based on geofence. Black arrows represent the standard flow. In green, the additional flow added with our authentication method.

3.1 Analysis and design of the authentication method based on Geofencing

To design the authentication method, we first analyzed the flow of a request to any Alexa Skill. When a user sends a voice command to an Amazon Echo, it is forwarded to the Alexa cloud service, where the Skill is hosted. The request is then processed on the Skill's backend in AWS Lambda (Docs.Amazon, 2020) and a response is sent to the user via the Amazon Echo. This flow is represented by the black arrows in the figure 1. Without improved authentication, an attacker could impersonate the victim and use the Amazon Echo to send requests to the Alexa Skill, successfully performing a replay attack.

With this in mind, Geofencing technology was used to add two additional steps when processing a request. These additional steps are explained below. They can be seen as green arrows in figure 1.

1. The user starts the developed Android application and configures a geofence at the position of Amazon Echo. Additionally, this application continuously monitors the user's location with regard to the geofence and sends it to the DynamoDB service database (DynamoDB, 2020). Steps 1 to 6 in figure 1.
2. The developed Alexa Skill checks the user's last location in the DynamoDB service database before processing the request. If the user is within the geofence, the request is processed. Otherwise, the Alexa Skill returns an authentication error message and the request is not processed. Steps 7 to 16 in figure 1.

These two steps mitigate replay attacks. When the legitimate user is located outside the geofence, the attacker will receive an authentication error message. Therefore, if the attacker wants to perform a replay attack, he needs the legitimate user to be inside the geofence and therefore close to the Amazon Echo. This requirement makes the replay attack unfeasible.

3.2 Mobile application for geofence configuration

The development of the application, called "AlexaGeoApp", was carried out in the Android Studio IDE using the Java programming language, and Google SDKs, such as Geofencing and Google Maps (Geocoding, 2021),(GoogleMaps, 2021). The application renders a map and sets a geofence near the Amazon Echo smart speaker. It also continuously monitors the user's position with respect to the geofence and sends it to the AWS DynamoDB database. Finally, the application has an interface that informs the user and guides him through the whole setup process to establish the geofence.

3.3 Alexa Skill that uses our authentication method based of Geofencing

The developed Alexa Skill consists of two parts: the frontend and the backend. The frontend was developed in Alexa Developer Console (Console, 2020) and is responsible for detecting user requests and sending them to the back end. The backend, written in JavaScript and hosted in AWS Lambda, handle requests processing and provides responses to the frontend. Furthermore, a communication with AWS DynamoDB was implemented through the backend to verify the user's last location in relation to the geofence. If the user is within the geofence, the Alexa Skill accepts requests, otherwise it denies them.

4 Results

In this paper, we developed an authentication method based on Geofencing to protect requests in Amazon Alexa, mitigating replay attacks. Geofencing technology was used as an additional authentication factor when processing voice requests. To achieve this, a PoC was developed, consisting of the creation of an Android application and an Alexa Skill.

On the one hand, the user can configure a geofence at the position of the Amazon Echo smart speaker through the developed Android application. On the other hand, the developed Alexa Skill can use the configured geofence to accept or deny the user's requests. If the user is within the geofence the requests are accepted. Otherwise, they are denied. Therefore, if an attacker wants to perform a replay attack, he needs the legitimate user to be inside the geofence, making the replay attacks unfeasible.

Some important parts of the developed Android app and Alexa Skill are highlighted below.

4.1 AlexaGeoApp: an Android application

The developed Android application, called "AlexaGeoApp", features a user-friendly interface that allows the user to configure the geofence near Amazon Echo smart speaker. Once the geofence is configured, "AlexaGeoApp" renders a map representing this geofence as a red zone. The application monitors the user's location with respect to the geofence and sends it to the DynamoDB service. In addition, when the user enters or exits the geofence, the user is informed via a notification on his mobile device. This can be seen in the figure 2.

4.2 Secure Authentication: an Alexa Skill

The Alexa Skill, called "Secure Authentication", uses our authentication method based on Geofencing to process requests. "Secure Authentication" Skill connects to the DynamoDB database to make decisions based on the latest data stored by "AlexaGeoApp" to process user

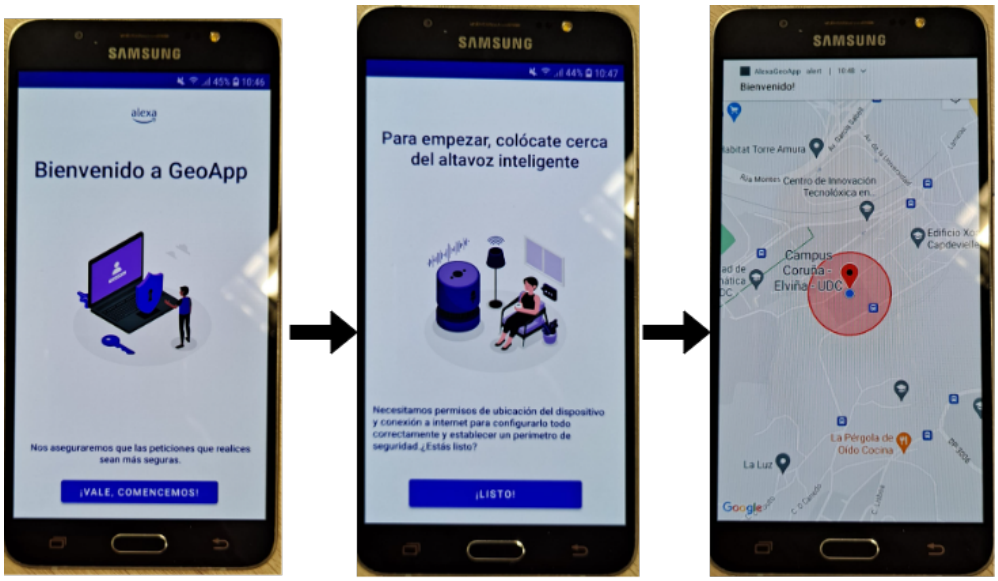


Figure 2: "AlexaGeoApp" flow. From left to right: (1) Welcome message to the user. (2) We inform the user to stand near the smart speaker. (3) We retrieve the user's location, set the geofence and inform the user with a notification.

requests. The user can verify his authentication status through the "Secure authentication" Skill. If the user is successfully authenticated by this method, he can make requests to the "Secure authentication" Skill. However, if authentication fails, the "Secure Authentication" Skill denies the request.

Finally, "Secure Authentication" has a validity window for the data stored in DynamoDB. This sets a maximum time for the validity of the data. If the data is very old and there have been no updates to the user's location, "Secure authentication" rejects all the requests, following the fail-safe principle (Leedeo, 2020). With this principle we avoid any denial of service issues on the mobile device that containing the developed "AlexaGeoApp", such as running out of battery or losing internet connectivity. "AlexaGeoApp" needs internet connectivity to detect user entries and exits in the geofence and update the DynamoDB database. If these state changes are not correctly detected, the authentication based on Geofencing would not work correctly. Thanks to the validity window we avoid this problem.

The resulting code of this work has been published as an open source project on the following Github repository: https://github.com/jorgefgarcia/AlexaGeoApp

5 Conclusions

The lack of security measures in the Amazon Alexa virtual assistant underscores the importance of addressing request security. Solutions that tackle the lack of authentication in this kind of settings, as the one proposed in this paper, may encourage manufacturers to develop more robust security measures for a safer user experience. Amazon environment could integrate our solution by default to protect sensitive requests, such as an online purchase through the Amazon Echo smart speaker, against replay attacks. In addition, it would be interesting to consider future vulnerabilities and adjust the design accordingly.

6 Future work

This paper has introduced an authentication method based on Geofencing that improves the authentication of requests directed at Amazon Alexa. However, it's important to consider future enhancements:

- **Utilization of other technologies:** Expanding the authentication enhancement to other virtual assistants such as Google Assistant, Bixby, Cortana or Siri. Additionally, developing a cross-platform application to configure security from multiple devices.
- **Responsive design:** Achieve a user interface design that adapts to different device screens.
- **Multi-user mode:** Add the capability of setting multiple geofences on a map with a group of users.
- **Wi-Fi usage:** Implementing a feature to process requests only when the user's mobile is connected to the same Wi-Fi network as the smart speaker.

Acknowledgements

This work was supported by the grant ED431C 2022/46 – Competitive Reference Groups GRC – funded by: EU and "Xunta de Galicia" (Spain). This work was also supported by CITIC, funded by "Xunta de Galicia" through the collaboration agreement between the "Consellería de Cultura, Educación, Formación Profesional e Universidades" and the Galician universities to strengthen the research centres of the "Sistema Universitario de Galicia" (CIGUS). Also, the work is founded by the "Formación de Profesorado Universitario" (FPU) grant from the Spanish Ministry of Universities to Martiño Rivera Dourado (Grant FPU21/04519).

Bibliography

Z. S. Alattar, T. Abbes, and F. Zerai. Smartphone-key: Hands-free two-factor authentication for voice-controlled devices using wi-fi location. *IEEE Transactions on Network and Service Management*, pages 1–1, 2023. Conference Name: IEEE Transactions on Network and Service Management.

Amazon. What is alexa voice ID? - amazon customer service. *https://www.amazon.com/gp/help/customer/display.html?nodeId=GYCXKY2AB2QWZT2X*, 2020. [Online; accessed 7-June-2023].

D. Amazon. Amazon developers. *https://developer.amazon.com/es-ES/home.html*, 2023. [Online; accessed 7-June-2023].

Amazon Alexa Skills. Alexa skills | amazon.com. *https://www.amazon.com/alexa-skills/b?ie=UTF8&node=13727921011*, 2016. [Online; accessed 6-September-2023].

AWS. ¿qué es AWS? *https://aws.amazon.com/es/what-is-aws/*, 2020. [Online; accessed 6-September-2023].

A. D. Console. Amazon alexa voice AI | alexa developer official site. *https://developer.amazon.com/en-US/alexa.html*, 2020. [Online; accessed 7-June-2023].

Docs.Amazon. ¿qué es AWS lambda? - AWS lambda. *https://docs.aws.amazon.com/es_es/lambda/latest/dg/welcome.html*, 2020. [Online; accessed 22-April-2023].

A. DynamoDB. ¿qué es amazon DynamoDB? - amazon DynamoDB. *https://docs.aws.amazon.com/es_es/amazondynamodb/latest/developerguide/Introduction.html*, 2020. [Online; accessed 7-June-2023].

G. D. A. Geocoding. Comenzar | geocoding API. *https://developers.google.com/maps/documentation/geocoding/start?hl=es-419*, 2021. [Online; accessed 5-September-2023].

G. D. A. GoogleMaps. Guía de inicio rápido del SDK de maps para android | maps SDK for android. *https://developers.google.com/maps/documentation/android-sdk/start?hl=es-419*, 2021. [Online; accessed 5-September-2023].

Incibe. El uso de altavoces inteligentes que riesgos de seguridad y privacidad nos exponen | ciudadanía | INCIBE. *https://www.incibe.es/ciudadania/formacion/infografias/el-uso-de-altavoces-inteligentes-que-riesgos-de-seguridad-y-privacidad-nos-exponen*, 2023. [Online; accessed 7-June-2023].

Leedeo. ¿qué significa que un sistema es fail safe? *https://www.leedeo.es/l/fail-safe/*, 2020. [Online; accessed 22-April-2023].

K. M. Malik, H. Malik, and R. Baumann. Towards vulnerability analysis of voice-driven interfaces and countermeasures for replay attacks. In *2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*, pages 523–528, 2019.

S. W. Rahate and D. M. Z. Shaikh. Geo-fencing infrastructure: Location based service. *IRJET*, 03(11), 2016.

G. Zhang, C. Yan, X. Ji, T. Zhang, T. Zhang, and W. Xu. Dolphinattack: Inaudible voice commands. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, CCS '17, page 103–117, New York, NY, USA, 2017. Association for Computing Machinery.

N. Zhang, X. Mi, X. Feng, X. Wang, Y. Tian, and F. Qian. Understanding and mitigating the security risks of voice-controlled third-party skills on amazon alexa and google home. *ArXiv*, abs/1805.01525, 2018.

Ethical and Legal Risks of Algorithmic and AI Tools Developed to Fight Against Trafficking in Cultural Property in the RITHMS Project

Amara García Adán and María Ángeles Fuentes Loureiro

ECRIM, Faculty of Law, Universidade da Coruña, 15071 A Coruña, Spain
Correspondence: amara.garcia.adan@udc.es, maria.fuentes.loureiro@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: Algorithmic technologies, big data and artificial intelligence have also disrupted the legal field. AI tools used in police investigations and in the judicial process stand out. The tools developed in the RITHMS project are aimed at combating the illegal trade in cultural goods. These tools facilitate the identification of criminal networks and their members. They are also useful for monitoring art markets, online auction sites and social networks to detect suspicious transactions. This poses ethical and legal challenges, requiring risk analysis and ensuring compliance with data protection, procedural and fundamental rights legislation. The aim is to address these challenges to ensure responsible use by competent authorities.

1 Introduction

The RITHMS project (Research and Innovation in Tackling the Illicit Trade of Cultural Goods through Multidisciplinary Science and Technology)[1] [2] was born out of the need, on the one hand, to understand the phenomenon of illicit trade in cultural goods and, on the other hand, to increase technological preparedness against the crime. RITHMS arises from the need to understand the functioning of markets for cultural goods. It is one of the least understood and supervised markets; there is still uncertainty in the provenance of objects, lack of traceability and poorly tracked transactions that contribute to an optimal scenario for crime. Thus, the project is aimed at solving the problems faced by Law Enforcement Agencies (LEAs) in dealing with this type of crime.

Experts widely acknowledge that the illicit trade in cultural property is no longer a localised phenomenon driven by a handful of individuals. On the contrary, it has become a lucrative source of income for criminal organisations that take advantage of the opportunities provided by online auctions and the visibility offered by social networks to further expand the already flourishing "grey market". This involves smuggling, theft and cross-border trade in valuable

[1] This paper has been prepared within the framework of the research project RITHMS – Research, Intelligence and Technology for Heritage and Market Security (GA 101073932) [HORIZON-CL3-2021-FCT-01-08]. Funded by the European Union. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

[2] Much of the information used in this article has been gathered from work on the above project, which will be publicly available on the RITHMS website. In particular, reference is made to Deliverable 1.1 on the Initial Legal Requirements, which is included in Work Package (WP) 1. On the other hand, Deliverable 7.1 on the Legal Framework is included in WP 7.

cultural and heritage property. This boom in transactions of illicitly obtained cultural property has been fuelled by the general lack of sound regulations governing this market, inconsistencies in national legislation and the inherent difficulties in tracing the origin of objects.

In response to this context, the RITHMS project advocates an interdisciplinary approach, driven by the international nature of this criminal activity and its intricate connections with other illicit networks and its connections with organised crime. The project aims to formulate a replicable strategy to effectively counter the challenges posed by the illicit trafficking of stolen or looted cultural property.

The RITHMS project aims to strengthen the operational capacity of police, customs and border authorities to deal with the increasingly organised and polycriminal nature of trafficking in cultural property through research, technological innovation, outreach and training. It is a collaborative effort aimed at enhancing the operational capabilities of law enforcement agencies to address the complex and evolving challenges posed by illicit trafficking in cultural property.

To do so, we must first come to understand the dynamics of the criminal phenomenon of illicit trafficking in cultural heritage. This research work is essential, as it will underpin the developing of digital technologies in the collection and processing of evidence to be used in court. On this basis, the consortium will develop an innovative digital platform to be used by LEAs. RITHMS platform will stand out for its interoperability and multi-functionality, allowing the identification, assessment and analysis of relationships between criminal and non-criminal actors. It will be based on Social Network Analysis (SNA) which allows mapping and analysis of social connections between individuals and groups of individuals by exploiting graph theory (RITHMS, n.d.). By outlining the networks involved in the illicit trafficking of cultural property and their possible evolution, the platform will improve the accessibility and accuracy of information available to LEAs. This tool will empower authorities in their efforts to combat illegal trade and better understand the dynamics driving these criminal groups.

Thus, RITHMS will equip LEAs with technological tools to increase their capacity to trace trafficking in cultural property and to prevent the emergence of organised crime networks through a technological platform developed and validated according to the needs and requirements of the professionals. To this end, it is necessary to the consortium will create a knowledge-base (in the form of multiple, structured datasets) related to the illicit heritage market and its actors, including: existing open datasets from LEAs, open sources (especially social media platforms), mobile network traffic datasets, satellite imagery, financial forensics and a database fed through custom-developed tracking software capable of detecting objects auctioned or sold directly online. Thus, based on this database, an algorithmic application capable of identifying the dynamics and actors of criminal networks involved in the trafficking of illicit goods will be created.

The development of this tool raises important legal questions. In order to use such a tool in the field of criminal law, whether at the investigative stage or in court, it is necessary to analyse the ethical and legal risks involved in the development and final use of the RITHMS platform. Data protection legislation, procedural law and, ultimately, fundamental rights, must be taken into account when developing such applications in order to avoid irregularities and infringements of rights and to ensure their legal validity.

2 Methodology

Within this project there are nineteen partners from eleven different countries including: Bosnia Herzegovina, Bulgaria, Croatia, Finland, Germany, Italy, Moldova, Romania, Spain, Switzerland and The Netherlands. The project coordinator is Dr. Arianna Traviglia from the Istituto Italiano di Tecnologia. The RITHMS consortium has been strategically assembled to ensure that it can effectively support exploitation and is aligned with an end-user approach. In particular, the consortium includes six police and border agencies, providing a deep understanding of end-user requirements. This will ensure that the technology developed is fit for purpose.

The Law Enforcement Agencies (LEAs) working with RITHMS will continue to be involved in the post-project phases, actively testing and providing real-time feedback to improve the platform. At the same time, industry partners will play a key role in the technical development of the platform, including data collection, satellite data integration, graph generation and AI advancements. In addition, research and technology organisations (RTOs) with expertise in technology transfer will provide strategic support by assessing market potential and identifying secondary market opportunities beyond crime-fighting applications.

The involvement of legal, ethical and policy partners (UDC, EEMA, HföD, CPT, EIM) will ensure that RITHMS complies with and contributes to European and international regulations, principles and standards. This involvement will also enable future collaboration, sharing of best practice and policy contributions., ii) concepts and iii) models that will guide the actions to be taken and the development of the technological deliverables, in particular the RITHMS Platform.

The core concept of RITHMS is based on a series of assumptions that include the widespread diffusion of intelligence-led policing as an organisational model for law enforcement. Such organisational models can greatly benefit from an automated intelligence tool based on SNA that provides actionable knowledge and information on complex and often transnational criminal networks. The information generated by the analysis of social networks involved in cultural heritage trafficking greatly benefits the effectiveness of investigations and prevention, allowing law enforcement agencies to direct their resources (human and technological) towards specific targets, where the reward is more relevant (such as effectively dismantling the criminal network or recovering assets).

The concepts that RITHMS will use during its lifetime revolve around SNA: an analytical tool that studies the links within social entities in a rigorous and quantitative way. It models social interactions through a graphical representation and can shed light on the relationships between actors (nodes) and on the flow of information, financial resources and goods through the network. Several modules will be able to feed the Platform with various data sources, merged into a correlation engine. AI models will also be used to implement the predictive capabilities of RITHMS, with a graph representation. Based on the above, the RITHMS methodology will be organised according to this structure: needs assessment and user-centred design; research on the nature of cultural heritage crime and the intersections between organised crime; development of an SNA-based platform and validation.

Regarding the work of the legal ethics team of the University of A Coruña, it begins with the completion of several studies on the initial legal requirements, the legal framework and some etichal issues. The study on the legal framework follows a methodology consisting of comparative and legal analysis techniques to analyse the qualitative data collected through the following means: - National reports on the six countries of the industry partners, five of which are EU Member States and one is not, and the six countries of the Consortium's law enforcement agency partners (LEAs), three of which are EU Member States and three of which are not. - Desk research to assess information published in the EU, internationally and in the countries under study.

As for the second study, the initial legal requirements, its methodology follows a requirements definition process. The requirements determination process usually consists of three stages (Pitts and Browne, 2007): information gathering, representation and verification. It is necessary to search for and analyse all useful sources to identify the elements that should guide the development of the system from the design phase. For this process, two perspectives can be used: on the one hand, the localising perspective, applicable to legal issues, which assumes that requirements are something that actually exist and simply need to be found. This perspective implies that the requirements are stable and recognisable. On the other hand, the constructionist perspective aims to create something new by combining identified elements in new ways.Compliance with these requirements will be monitored throughout the duration of the project and, in this case, will consist of respect for both rights and freedoms and legal provisions.

The legislative acts relevant to the project need to be identified and specific requirements extracted, in particular those stemming from the General Data Protection Regulation (hereinafter GDPR[3]) and the Law Enforcement Directive (LED [4] The detailed methods and approaches for analysing the EU and national legal frameworks were as follows:

A) Approach to analysing the EU legal framework: in this phase of the analysis we studied, through desk research, the relevant EU legislation. The analysis followed the approach described above as applied to international law, but also took into account the distinctive features of the EU legal system.

B) Approaching national legal frameworks: this research also analysed the legislations of the six EU and non-EU countries relevant to the use of the RITHMS platform in the end-user phase. The LEAs participating in the consortium were identified in order to balance the following criteria: i) participation of the countries (Bulgaria, Bosnia and Herzegovina, Spain, Italy, Moldova and the Netherlands); ii) presence of both police forces and border authorities; iii) relevance of the countries and their role in the illicit trafficking of cultural property; iv) presence of experienced partners with specialised units and others in the start-up phase.

The final step was to gather input and expertise from the consortium partners and project stakeholders to develop the initial version of the system requirements. An online questionnaire was used to collect information from the LEAs that will be involved in the use of the system. The legal requirements were then extracted from the information submitted by the Consortium's legal experts.

Since the specific characteristics of the multifaceted RITHMS platform and its components cannot be fully known in advance at the time of drafting the requirements, two corrective measures have been put in place to prevent the risk of impertinent requirements. The first measure is to ensure close interaction between partners with different expertise within the Consortium, in order to ensure that legal experts are aware of technology developments within the project, and that industrial partners are fully briefed on the practical consequences of changes in the legal framework. The second corrective measure is to involve end-users in the development of the system. End-users' views on their expectations regarding the project's results and the functioning of the platform will be taken into account throughout the project. During the co-creation phase, the views of the LEAs involved in the project and of the companies' ethics committees will be gathered.

3 Discussion

As we have already mentioned, developing a tool of these characteristics entails a series of ethical and legal risks. After analysing the project proposal, a series of ethical risks were identified that can be classified into three groups. Firstly, Internal AI risks, which are those inherent to AI technology. In the second place, Operational AI risks, which correspond to risks that emanate from the interaction of the technology with the real world and, in the third place, User risks, which are characterised by the risks that appear when the user applies the technology.

Regarding internal AI risks, i.e. inherent AI risks, opacity and lack of accountability emerge. This constitutes a risk in terms of the formation of algorithm black boxes. The lack of trans-

[3] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016R0679

[4] Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA

parency about the functioning of the algorithm may lead to the potential incomprehensibility of automated decisions to human reasoning (Danaher, 2016) and, thus, to a lack of explainability and justification of the decisions taken, which is a fundamental requirement of European legal systems. Therefore, the use of these systems around police or judicial decisions should not undermine the ability to be able to explain decisions taken by public officials as their legitimacy depends on this.

In terms of operational AI risks, one of the first risks that appears is privacy interference. The very nature of the RITHMS platform, or similar platforms, has the risk of violating data protection law. This poses a very high risk in the development and use of the platform and will require privacy and data governance.

Data protection rights have been recognised for some time, with different levels of protection depending on whether the data is public or private. As mentioned, at the European Level, GDPR must be taken into account. In the specific context of criminal investigations within the European Union, the LED has established specific guidelines for the collection and analysis of data for the purpose of preventing, investigating, detecting or prosecuting criminal offences and ensuring public security. As a tool for crime investigation, all these provisions should be taken into account in the development of the RITHMS platform. However, it is important to remember that those rules constitute a minimum framework and that national rules may vary from country to country. It is therefore crucial to take into account the specificities of the countries in which the RITHMS platform is developed, tested and deployed.

Regarding data protection and privacy rights, we can highlight the risk that arises from the processing of data for investigative purposes that were not originally collected for that purpose. In the use a scraping technique, on which the RITHMS data collection is based, it is likely that the data were collected in the first place by private companies in the course of their business and are not related to crime. So there is a risk of using data for purposes other than those for which it was originally produced and stored. As the regulation of data collection is different depending on the purpose of data collection and processing, this entails the legal risk of the violation of the purpose limitation principle of Article 5(1)(b) GDPR. We should bear in mind that the misuse of such investigative technologies could disrupt law enforcement proceedings and affect the admissibility of evidence in court. For example, an abusive use of the platform as it can be used to identify any person of interest raises questions about the privacy of individuals and, in addition, this tool could start to be used for wider and wider purposes. Discimination and bias are also one of the risks identified. This risk is found in all three groups mentioned above. Discrimination can be found in how the target variable and the class labels are defined, in discrimination based on feature selection or in proxy discrimination which, in the appearance of neutrality, emerges as discrimination. Therefore, special attention should be paid to how data are collected and what kind of data are involved. If data are poorly labelled, inaccurate, incomplete or if it reflects human prejudices, then the AI model will reproduce those same biases (Surden, 2020)

Closely linked to the above is the timeless sequencing risk. AI needs to be trained with inputs, but those inputs belong to the past. The use of information in a sequenced and timeless way can lead AI to reiterate paradigms that are culturally, economically or legally outdated. This is a risk that raises the need for the information used as input to be subjected to a process of timing, i.e. adaptation to the present. Thus, we cannot fall into mathwashing, i.e. we cannot cling toto the false idea that an algorithm is free of bias, as it is assumed to be neutral and objective, but must seek control measures to minimise this risk. The use of such platforms may also lead to a possible violation of due process and fair trial guarantees (Quezada-Tavárez et al., 2021). This may result in the invalidity of certain evidence and lead to impunity for criminal actions for lack of lawful evidence.

First, it may lead to the inadmissibility of evidence obtained in violation of fundamental rights or without due proportionality. This is important because there are a number of legal standards that must be met for evidence to be admissible at trial, including procedural standards for obtaining evidence without violating rights. Otherwise, the evidence will not be

admitted at trial or will be declared invalid. And even if lawful evidence comes to light after the trial, the principle of ne bis in idem prevents the same individuals from being tried again for the same facts, even if lawful evidence subsequently comes to light.

On the other hand, the defendant's access to the information used for sentencing may be reduced with the use of technologies such RITHMS platform due to the algorithm opacity, and this infringes the principle of contradiction that governs criminal proceedings and thus the right of defence (Marquenie, 2019). A person who is accused has the right to know on what information his or her accusation is based, so if he or she is accused or convicted on the basis of a partially secret or incomprehensible algorithm, his or her right to defence would be violated (Virginia Foggo, 2020).

There may also be a violation of the right to the presumption of innocence. The principle of presumption of innocence requires that persons be presumed innocent until proven guilty according to law. Consequently, this principle is violated if a conviction is based on illegal or unlawful evidence, as it would result in a verdict without valid incriminating evidence. It is therefore essential to prevent evidence obtained during the police investigation phase from being obtained without due respect for the fundamental rights of the accused person during the collection of such evidence.

These risks of rights violations, among others, in the investigative process are exacerbated by the fact that many EU member states do not have specific regulations on intelligence-led policing and, consequently, the limits or requirements for such activity are not established. It should be borne in mind that the development of technological tools similar to RITHMS is relatively recent, so we should not be surprised by the lack of regulation in this regard. The EU and non-EU Member States that do have specific legislative provisions have mostly enacted them recently, but without considering the notion of tools using Big Data or Artificial Intelligence.

Therefore, for the time being, its use must be must be guided by the general rules of criminal procedure of each country. Finally, it is necessary to mention the Proposal for Regulation on Artificial Intelligence (the AI Act[5]) which is expected to enter into force next year. This Act is structured in categories according to the risk presented by the system. There are four levels of risk in relation to AI practices: unacceptable risks; high risks; limited risks; minimal risks. The RITHMS platform would fit into the high risk level, which are not prohibited (unlike unacceptable risks) but are subject to a detailed certification regime.

The IA Act requires suppliers of high-risk IA systems to conduct a prior conformity assessment before placing them on the market. Suppliers must ensure that their systems comply with the "essential requirements" set out in Chapter 2 of Title III of the IA Act. They can then affix a CE marking on compliant systems, which can be freely imported and distributed throughout the EU. This includes data governance, i.e. rules on how input data sets should be designed and used, rules on the preparation of the data and the assessment of the formulation of relevant assumptions [about] the information that the data are supposed to measure and represent. In addition, these systems must be designed and developed in such a way that they can be 'effectively monitored by natural persons during the period that the AI system is in use, allowing the human supervisor to detect anomalies and to correctly interpret the results. If a high-risk system is operated by a "user" rather than the original provider - for example, a LEA buys and installs the RITHMS Consortium's platform - the allocation of responsibilities is very different in the Act than in the GDPR.

[5] European Commission. (2021). Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts. COM/2021/206 final. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206

4 Conclusion

For all these reasons, the tool developed in the framework of the RITHMS Project is a very interesting tool for combating trafficking in cultural goods. However, it is necessary to pay attention to all the requirements that these tools must meet in order not to violate the fundamental rights of citizens. This will require all partners to work together to provide the necessary expertise to develop this tool.

Bibliography

J. Danaher. The threat of algocracy: Reality, resistance and accommodation. *Philosophy & Technology*, 3(29):245–268, 2016. An optional note.

T. Marquenie. The impact of predictive policing and law enforcement ai on human rights: The right to fair trial under pressure, 2019.

M. G. Pitts and G. J. Browne. Improving requirements elicitation: an empirical investigation of procedural prompts. *Information Systems Journal*, 17(1):89–110, 2007.

K. Quezada-Tavárez, P. Vogiatzoglou, and S. Royer. Legal challenges in bringing ai evidence to the criminal courtroom. *New Journal of European Criminal Law*, 12(4):531–551, 2021.

RITHMS. Rithms. https://rithms.eu/communication/press-kit/rithms-press-release/file, n.d. [Online; accessed 1-September-2023].

H. Surden. 719Ethics of AI in Law: Basic Questions. In *The Oxford Handbook of Ethics of AI*. Oxford University Press, 07 2020. ISBN 9780190067397.

J. V. Virginia Foggo. Artificial intelligence, due process, and criminal sentencing. *Michigan State Law Review*, pages 295–354, 2020. URL *https://core.ac.uk/download/pdf/327102171.pdf*.

BDWatchdogFaaS: A Tool for Monitoring and Analysis of Functions-as-a-Service in Cloud Environment

Manuel Framil, Mario Carpente, David Fraga, Jonatan Enes, Roberto R. Expósito, Guillermo L. Taboada, and Juan Touriño

Centro de Investigación CITIC, Universidade da Coruña, 15071 A Coruña, Spain
Correspondence: m.framil.deamorin@udc.es

DOI: *https://doi.org/543210/xxxxx1234567890*

Abstract: BDWatchdog is a framework to assist in the in-depth and real-time analysis of the execution of Big Data frameworks and applications. BDWatchdog was originally developed to monitor Hadoop ecosystems deployed on serverless containers, in order to detect bottlenecks and spot certain patterns that frameworks or applications may have. In this work we shift the focus to monitoring serverless functions in the public cloud, by proposing an extension of BDWatchdog which captures, transforms and analyzes logs from both AWS Cloudwatch and Azure Application Insights, which store the logs from AWS Lambda and Azure Functions (respectively), the FaaS (Function-as-a-Service) solutions of the two main public cloud providers, AWS and Azure. The extension, called BDWatchdogFaaS, builds and stores a common model to both providers, allowing to consult, analyze and monitor function logs from AWS and Azure indistinctly. The transformation of logs into the common model is done by a FaaS of the corresponding provider, which in near real-time ingests, processes and sends the data to a common storage. In addition, the data is forwarded to a Power BI dashboard so that the serverless functions can be monitored easily.

1 Introduction

Cloud computing refers to a paradigm of computing that has undergone a enormous growth in the last decade, and in which the computing resources and services are consumed on demand through the internet. This has resulted in a fundamental change in every stage of the software applications life cycle, from it's conception and development until the way they are used by the final user. In cloud computing, the cloud provider is responsible for allocating, managing and scaling the underlying resources of the application, releasing the developers from the responsibility of administrating the infrastructure, thus allowing them to focus only in the functionality.

In recent years a new model of cloud computing, called Function-as-a-Service (FaaS) has been presented. This new paradigm allows to run small fragments of code (functions) without the need for allocating virtual machines or deploy any kind of server. The code is uploaded to the cloud provider, which automatically allocates the resources needed, run the function and returns the results. Faas are commonly employed to build microservices applications, because they are easily integrated between them and with other cloud services, and their code can be written en different programming languages depending of the needs of the application. Besides, FaaS are event-driven, which means they can be activated in response to specific events, such as HTTP requests, database changes or messages in queues, among others.

This work is part of the project PICSA (Productivity Increase by Cloud Serverless Automation, for further details see the acknowledgement section), which main objective is to increase

the productivity of developers and organizations when working with serverless technologies, particularly with FaaS.

2 BDWatchdogFaaS Architecture

The main contribution of this work is the design, implementation and publication of BDWatchdogFaaS, an application performance monitoring (APM) tool for FaaS in a multicloud environments. Previous relevant work is BDWatchdog (Enes et al., 2018), a monitoring and profiling tool for Big Data applications in Hadoop environments; and Serverless Containers (Enes et al., 2020), a framework for real-time container auto-scaling. BDWatchdogFaaS has been built as an cloud-native extension of BDWatchdog (Enes et al., 2018) to capture logs from AWS Lambda and Azure Functions, the FaaS solutions of Amazon Web Services and Azure, respectively. The main issue that BDWatchdogFaaS solves is the dependency of a particular provider due to the incompatibility between FaaS logs formats of different cloud providers. To overcome this problem, BDWatchdogFaaS proposes a common data model that enables to store the data in a single database, extract overall metrics and visualize the executions of FaaS from different hyperscalars simultaneously.

The architecture of BDWatchdogFaaS, presented in Figure 1 is based on microservices , which are independently deployable services, very loosely coupled and integrated through APIs.

BDWatchdogFaaS is composed of four modules, each of them being responsible for a specific task:

- **Azure monitoring module:** This module deals with the monitoring of Azure Functions and it has been implemented completely using Azure services. When an Azure Function is executed, Azure automatically generate a series of logs which are stored in Application Insights, the APM tool provided by Azure. Once in Application Insights, logs are immediately sent to a queue of EvenHub, which work as trigger to a Azure Function, called *AzureLogsForwarder*. This FaaS handles the task of processing logs to extract the relevant information, and then forwards it to the statistics and visualization modules.
- **AWS monitoring module:** This module deals with the monitoring of AWS Lambda. Analogous to the Azure monitoring module, this module has been built on top of AWS services. First, when a Lambda is executed, AWS registers the logs in a log group of Cloudwatch, the monitoring service of AWS. Using a subscription filter logs are automatically sent to a Lambda function, called *AWSLogsForwarder*, which is trigger as response to the incoming data. As for *AzureLogsForwarder*, this function parses logs to capture the most relevant information, and then forwards it to the statistics and visualization modules.
- **Statistic module:** This module has been developed to store the logs processed by the previous modules in an AWS DynamoDB Serverless table. DynamoDB is a fully managed, key-value NoSQL database designed to run high-performance applications provided by AWS.

  The table keeps all processed logs from both Azure Functions and AWS Lambda in a single point, so that another Lambda function, *LogStatsForwarder*, can read the data and obtain overall statistics about the FaaS being monitored in both cloud providers. These statistics are forwarded to the visualization module as well.
- **Visualization module:** Finally, the visualization module is built making use of the capabilities for real-time visualization of streaming datasets in Power BI Service. Up to four datasets have been defined: one for Azure logs, another for Lambda, the third dataset collects logs from both, and the last one receives the stats data. From these streaming datasets it has been created several visualizations that are updated in real-time. All of them have been gathered in a single dashboard panel in Power BI Service, available to users to review the behaviour of the running FaaS.

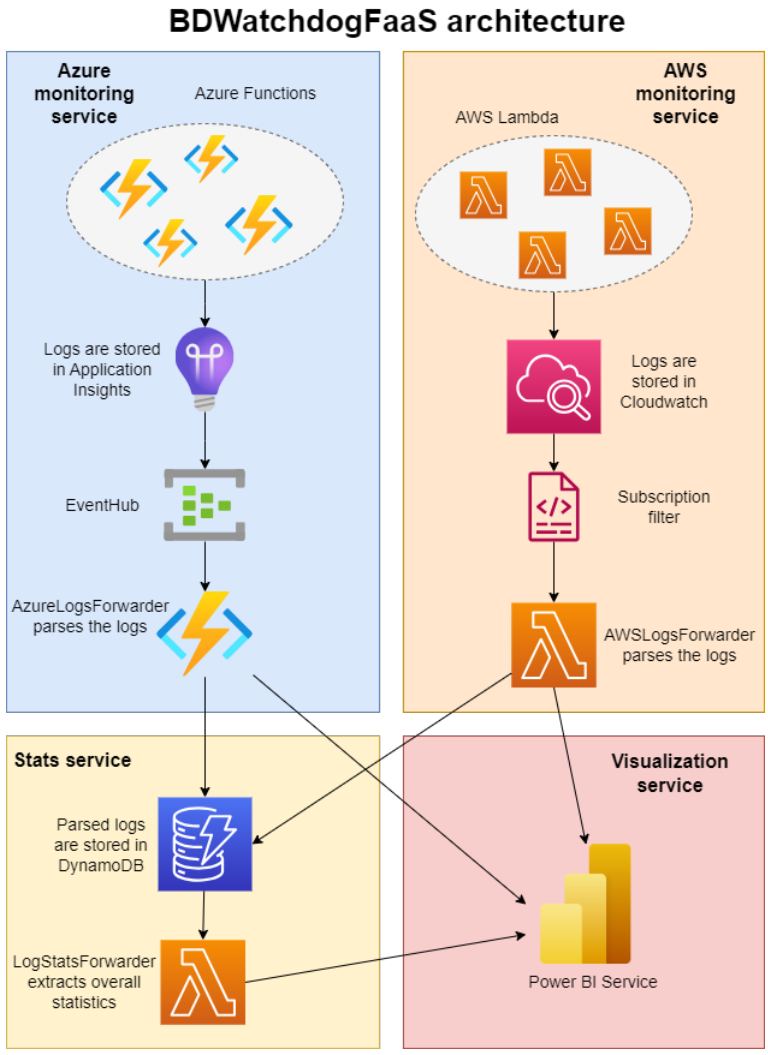


Figure 1: Architecture of BDWatchdogFaaS

3 BDWatchdogFaaS Deployment

BDWatchdogFaaS has been publish in the Serverless Application Repository (SAR) of AWS, a platform where developers can find, share, and deploy serverless applications and components. SAR allows users to discover pre-built serverless applications or functions created by the AWS community or AWS partners. The objective is to ease its deployment and broad its application. The applications are described using Cloudformation templates. They are key to define the architecture of serverless applications in the AWS ecosystem, enabling its automatic deployment. Cloudformation templates have been written using the AWS Serverless Application Model (SAM), an open-source framework for building serverless applications. It provides shorthand syntax to express functions, APIs, databases, and event source mappings. During deployment, SAM transforms and expands the SAM syntax into AWS CloudFormation syntax, enabling developers to build serverless applications faster.

Therefore, a SAM template defining the tool's architecture and configuration parameters (variables that must be set by the final user, such as the Power BI Service URLs to the streaming datasets) has been developed in order to publish BDWatchdogFaaS.

4 BDWatchdogFaaS Evaluation

BDWatchdogFaaS has been developed employing a TDD (Test Driven Development) approach to test its individual components and module integration. By doing so, it is ensured that the final product is well-tested when it comes to functionality and integration, leaving to this stage load and stress tests. Another aspect that needs to be tested is the fit of BDWatchdogFaaS in the market of FaaS monitoring tools.

The tests carried on are aimed to measure the reliability of BDWatchdogFaaS, putting the tool through several experiments that simulate real world environments. Furthermore, there are other tests that has been designed to verify the usability and fit of BDWatchdogFaaS in the market, checking out how easily can it be discovered inside the SAR and how quickly can it be deployed by a potential user. The tests were the following:

1. **Publication and discovery of BDWatchdogFaaS:** The objective is to verify that an user can find BDWatchdogFaaS inside the SAR. To do so, we designed a set of queries that potential users would use to look for applications in the SAR, and measured in which position is BDWatchdogFaaS placed on each one. These queries include a combination of relevant terms that describe BDWatchdogFaaS, such as "monitoring", "FaaS" or "multicloud".

   The results show that for 8 of the 12 queries tested, BDWatchdogFaaS placed itself between the top 3 applications retrieved by the search engine.

2. **Deployment of BDWatchdogFaaS:** We tested that, once the user has found the application in the SAR, it can be deployed easily. We tried, successfully, to deploy BDWatchdogFaaS directly from the SAR interface, and were able to start running the tool in less than 5 minutes, proving that it can be effortlessly set up.

3. **Reliability during peaks of demand:** In order to validate that BDWatchdogFaaS can scale properly to respond peaks of demand, we have monitored a thousand executions of a Lambda in a period of 5 seconds. This test generated a large amount of logs. Rather than end up collapsing BDWatchdogFaaS, the tools was able to process all the logs without delays or errors.

4. **Long-running execution:** A selected set of four functions, has been executed in AWS and Azure, once per minute, during 24 hours straight so we can test whether BDWatchdogFaaS monitors continuously and seamless for such intervals of time. During the experiment, the operation of BDWatchdogFaaS was uninterrupted and no errors were yield, confirming that the tool can run for that, and probably longer, periods of time.

5 Conclusions

This paper has presented BDWatchdogFaaS, an extension of our previous work, BDWatchdog, which is able to watch over FaaS applications in a multi-cloud environment, allowing organizations to monitor simultaneously AWS Lambdas and Azure Functions, the main FaaS implementations. This tool, publicly available in AWS Serverless Repository (SAR), captures, transforms and analyzes logs from both AWS Cloudwatch and Azure Application Insights, which store the logs from AWS Lambda and Azure Functions. A common model for FaaS logs has been designed to unify data process and analysis. Moreover, data has been processed in near real-time and visualized in Power BI. This way, serverless functions can be monitored straightforwardly using a standard software stack, providing value to developers and stakeholders by increasing the productivity when working with FaaS.

As future work, it is planned to further develop BDWatchdogFaaS in the search of product market fit in the sector of FaaS monitoring tools. Moving from SAR to BDWatchdogFaaS as SaaS will increase its adoption and benefit from the lessons learnt analysing logs of a higher number of FaaS.

Acknowledgement

This work was funded by the Ministry of Science and Innovation of Spain (ref. PDC2021-121309-I00/MCIN/AEI/10.13039/501100011033) and by the European Union "NextGenerationEU/PRTR". CITIC is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS).

Bibliography

J. Enes, R. R. Expósito, and J. Touriño. Bdwatchdog: Real-time monitoring and profiling of big data applications and frameworks. *Future Generation Computer Systems*, 87:420–437, 2018. URL *https://www.sciencedirect.com/science/article/pii/S0167739X17316096*.

J. Enes, , R. R. Expósito, and J. Touriño. Real-time resource scaling platform for big data workloads on serverless environments. *Future Generation Computer Systems*, 105:361–379, 2020. URL *https://www.sciencedirect.com/science/article/pii/S0167739X19310015*.