



UBICOMM 2024

The Eighteenth International Conference on Mobile Ubiquitous Computing,
Systems, Services and Technologies

ISBN: 978-1-68558-191-6

September 29 - October 03, 2024

Venice, Italy

UBICOMM 2024 Editors

Jaime Lloret Mauri, Polytechnic University of Valencia, Spain

Eugen Borcoci, National University of Science and Technology POLITEHNICA

Bucharest, Romania

UBICOMM 2024

Forward

The Eighteenth International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies (UBICOMM 2024), held between September 29th, 2024, to October 3rd, 2024, in Venice, Italy, continued a series of international events meant to bring together researchers from the academia and practitioners from the industry in order to address fundamentals of ubiquitous systems and the new applications related to them.

The rapid advances in ubiquitous technologies make fruition of more than 35 years of research in distributed computing systems, and more than two decades of mobile computing. The ubiquity vision is becoming a reality. Hardware and software components evolved to deliver functionality under failure-prone environments with limited resources. The advent of web services and the progress on wearable devices, ambient components, user-generated content, mobile communications, and new business models generated new applications and services. The conference makes a bridge between issues with software and hardware challenges through mobile communications.

Advances in web services technologies along with their integration into mobility, online and new business models provide a technical infrastructure that enables the progress of mobile services and applications. These include dynamic and on-demand service, context-aware services, and mobile web services. While driving new business models and new online services, particular techniques must be developed for web service composition, web service-driven system design methodology, creation of web services, and on-demand web services.

As mobile and ubiquitous computing becomes a reality, more formal and informal learning will take place out of the confines of the traditional classroom. Two trends converge to make this possible: increasingly powerful cell phones and PDAs, and improved access to wireless broadband. At the same time, due to the increasing complexity, modern learners will need tools that operate in an intuitive manner and are flexibly integrated in the surrounding learning environment.

Educational services will become more customized and personalized, and more frequently subjected to changes. Learning and teaching are now becoming less tied to physical locations, co-located members of a group, and co-presence in time. Learning and teaching increasingly take place in fluid combinations of virtual and "real" contexts, and fluid combinations of presence in time, space and participation in community. To the learner full access and abundance in communicative opportunities and information retrieval represents new challenges and affordances. Consequently, the educational challenges are numerous in the intersection of technology development, curriculum development, content development and educational infrastructure.

We take here the opportunity to warmly thank all the members of the UBICOMM 2024 technical program committee, as well as all the reviewers. The creation of such a high-quality conference program would not have been possible without their involvement. We also kindly thank all the authors who dedicated much of their time and effort to contribute to UBICOMM 2024. We truly believe that, thanks to all these efforts, the final conference program consisted of top-quality contributions. We also thank the members of the UBICOMM 2024 organizing committee for their help in handling the logistics of this event.

We hope that UBICOMM 2024 was a successful international forum for the exchange of ideas and results between academia and industry for the promotion of progress related to mobile ubiquitous computing, systems, services, and technologies.

UBICOMM 2024 Chairs

UBICOMM 2024 Steering Committee Chair

Jaime Lloret Mauri, Universitat Politecnica de Valencia, Spain

UBICOMM 2024 Steering Committee

Stéphane Galland, Belfort-Montbéliard University of Technology, France

Wladyslaw Homenda, Warsaw University of Technology, Poland

Chunguo Li, Southeast University, China

Dmitry Korzun, Petrozavodsk State University, Russia

UBICOMM 2024 Publicity Chairs

Laura Garcia, Universidad Politécnica de Cartagena, Spain

Lorena Parra Boronat, Universidad Politécnica de Madrid, Spain

UBICOMM 2024 Committee

UBICOMM 2024 Steering Committee Chair

Jaime Lloret Mauri, Universitat Politecnica de Valencia, Spain

UBICOMM 2024 Steering Committee

Stéphane Galland, Belfort-Montbéliard University of Technology, France

Wladyslaw Homenda, Warsaw University of Technology, Poland

Chunguo Li, Southeast University, China

Dmitry Korzun, Petrozavodsk State University, Russia

UBICOMM 2024 Publicity Chairs

Laura Garcia, Universidad Politécnica de Cartagena, Spain

Lorena Parra Boronat, Universidad Politécnica de Madrid, Spain

UBICOMM 2024 Technical Program Committee

Afrand Agah, West Chester University of Pennsylvania, USA

Wafaa Ait-Cheik-Bihi, Schneider Electric, France

Mehmet Akşit, TOBB ET University, Ankara, Turkey / University of Twente, The Netherlands

A. B. M. Alim Al Islam, Bangladesh University of Engineering and Technology, Bangladesh

Mrim Alnfiai, Dalhousie University, Canada

Tahssin Altabbaa, Istanbul Gelisim University / Huawei Istanbul, Turkey

Nafisa Anzum, University of Waterloo, Canada

Paramasiven Appavoo, University of Mauritius, Mauritius

Mehran Asadi, Lincoln University, USA

F. Mzee Awuor, Kisii University, Kenya

Muhammed Ali Aydin, Istanbul University - Cerrahpasa, Turkey

Nebojsa Bacanin, Singidunum University, Serbia

Chiara Bachechi, University of Modena and Reggio Emilia, Italy

Matthias Baldauf, Eastern Switzerland University of Applied Sciences, Switzerland

Anoud I Bani-hani, Zayed University, Dubai, UAE

Luca Bedogni, University of Modena and Reggio Emilia, Italy

Oladayo Bello, New Mexico State University, Las Cruces, USA

Imed Ben Dhaou, University of Turku, Finland

Imen Ben Lahmar, ISIM Sfax | ReDCAD laboratory | University of Sfax, Tunisia

Djamal Benslimane, Université Claude Bernard Lyon 1, France

Aurelio Bermúdez, Universidad de Castilla-La Mancha, Spain

Javier Berrocal, University of Extremadura, Spain

Nik Bessis, Edge Hill University, UK

Robert Bestak, Czech Technical University in Prague, Czech Republic

Sourav Kumar Bhoi, Parala Maharaja Engineering College, India

Lucas Botoni De Souza, Federal University of Technology - Paraná, Brazil

Nadia Bouassida, Higher Institute of computer science and Multimedia, Sfax, Tunisia
Chérifa Boucetta, University of Reims Champagne-Ardenne, France
Yassine Boujelben, National School of Electronics and Telecommunications of Sfax | University of Sfax, Tunisia
Azedine Boulmakoul, Université Hassan II de Casablanca, Morocco
Maurizio Bozzi, University of Pavia, Italy
Joseph Bugeja, Malmö University, Sweden
Christian Cabrera, Trinity College Dublin, Ireland
Diego Leoel Cadette Dutra, Federal University of Rio de Janeiro, Brazil
Juan Vicente Capella Hernández, Universitat Politècnica de València, Spain
Beenish Moalla Chaudhry, The University of Louisiana, Lafayette, USA
Chao Chen, Purdue University Fort Wayne, USA
Radu-Ioan Ciobanu, National University of Science and Technology POLITEHNICA Bucharest, Romania
Michael Collins, Technological University Dublin, Ireland
André Constantino da Silva, IFSP & NIED/UNICAMP, Brazil
Giuseppe D’Aniello, University of Salerno, Italy
Roland Dodd, Central Queensland University, Australia
Ivanna Dronyuk, Lviv Polytechnic National University, Ukraine
Jalel Dziri, National Engineering School of Tunis | University Tunis El Manar, Tunisia
Wael M. El-Medany, University of Bahrain, Bahrain
Francisco Falcone, ISC-UPNA, Spain
Przemyslaw Falkowski-Gilski, Gdansk University of Technology, Poland
Olga Fedevych, Lviv Polytechnic National University, Ukraine
Niroshinie Fernando, Deakin University, Australia
Renato Ferrero, Politecnico di Torino, Italy
Olivier Flauzac, University of Reims, France
Franco Frattolillo, University of Sannio, Benevento, Italy
Stéphane Galland, Belfort-Montbéliard University of Technology, France
Crescenzo Gallo, University of Foggia, Italy
Jose Garcia-Alonso, University of Extremadura, Spain
Vassilis C. Gerogiannis, University of Thessaly, Greece
Seyed Ali Ghorashi, University of East London, UK
Mikhail Gofman, California State University, Fullerton, USA
Javier Gozalvez, Universidad Miguel Hernandez de Elche, Spain
Clementine Gritti, University of Canterbury, New Zealand
Weixi Gu, UC Berkeley, USA
Zhichun Guo, University of Notre Dame, USA
Mesut Güneş, Institute for Intelligent Cooperating Systems | Otto-von-Guericke-University Magdeburg, Germany
Cornelia Aurora Győrödi, University of Oradea, Romania
Qiang (Nathan) He, Swinburne University of Technology, Australia
Songlin He, New Jersey Institute of Technology (NJIT), USA
Wladyslaw Homenda, Warsaw University of Technology, Poland
Tzung-Pei Hong, National University of Kaohsiung, Taiwan
Sergio Ilarri, University of Zaragoza, Spain
Yasser Ismail, Southern University and A&M College, USA
Jacek Izydorczyk, Instytut Elektroniki Politechniki Śląskiej, Gliwice, Poland
Bingbing Jiang, Purple Mountain Laboratories, China

Rim Jouini, ENSI-University of Manouba, Tunisia
Liuwang Kang, University of Virginia, USA
Attila Kertesz, University of Szeged, Hungary
Dmitry Korzun, Petrozavodsk State University, Russia
Konstantinos Kotis, University of the Aegean, Greece
Chandra Krintz, UC Santa Barbara, USA
Abhishek Kumar, University of Helsinki, Finland
Gyu Myoung Lee, Liverpool John Moores University, UK
Pierre Leone, University of Geneva, Switzerland
Clement Leung, Chinese University of Hong Kong, Shenzhen, China
Yiu-Wing Leung, Hong Kong Baptist University, Hong Kong
Chunguo Li, Southeast University, China
Wenjuan Li, The Hong Kong Polytechnic University, China
Mauro Henrique Lima de Boni, Federal Institute of Education, Science and Technology of Tocantins - IFTO, Brazil
Xiaodong Liu, Edinburgh Napier University, UK
Jaime Lloret Mauri, Universitat Politecnica de Valencia, Spain
Giuseppe Loseto, Polytechnic University of Bari, Italy
Aliane Loureiro Krassmann, Federal Institute Farroupilha, Brazil
Derdour Makhlof, University of Tebessa, Algeria
Meriem Mandar, National School of Applied Sciences Bd Béni Amir, Khouribga, Morocco
Jordi Mongay Batalla, Warsaw University of Technology, Poland
Sathiamoorthy Manoharan, University of Auckland, New Zealand
Stella Markantonatou, Institute for Language and Speech Processing, Greece
Francesca Martelli, Istituto di Informatica e Telematica (CNR), Italy
Márcio Mendonça, Universidade Tecnológica Federal do Parana (UTFPR), Brazil
Weizhi Meng, Technical University of Denmark, Denmark
Philippe Merle, Inria Lille - Nord Europe, France
Daniela Micucci, University of Milano - Bicocca, Italy
Mona Minakshi, Intel Corporation, USA
Habib Mostafaei, Universität Berlin, Germany
Mjumo Mzyece, University of the Witwatersrand, Johannesburg, South Africa
Tamer Nadeem, Virginia Commonwealth University, USA
Ryo Nishide, Shiga University, Japan
YoungTae Noh, Korea Institute of Energy Technology (KENTECH), Korea
Josef Noll, University of Oslo, Norway
Kouzou Ohara, Aoyama Gakuin University, Japan
Jorge Ortiz, Rutgers University, USA
Hamza Ouarnoughi, INSA Hauts-de-France, Valenciennes, France
Sungkyu Park, Kangwon National University, South Korea
K. K. Pattanaik, ABV-Indian Institute of Information Technology and Management, Gwalior, India
Giovanni Pau, Kore University of Enna, Italy
Isidoros Perikos, University of Patras, Greece
Ivan Pires, University of Beira Interior, Portugal
Laura Po, University of Modena and Reggio Emilia, Italy
Christian Prehofer, DENSO Automotive, Germany
Tomasz Rak, Rzeszow University of Technology, Poland
Ann Ramirez, University of Florida, USA

Luca Reggiani, Politecnico di Milano, Italy
Elena Renda, IIT - CNR, Italy
André Restivo, University of Porto, Portugal
Amine Rghioui, EMI - Mohamed V University, Morocco
Ana Patrícia Rocha, University of Aveiro, Portugal
Federica Rollo, University of Modena and Reggio Emilia, Italy
Michele Ruta, Politecnico di Bari, Italy
Khair Eddin Sabri, The University of Jordan, Jordan
Prasan Kumar Sahoo, Chang Gung University, Taiwan
Zaineb Sakhravi, University of Sfax, Tunisia
Josep Maria Salanova Grau, Center for Research and Technology Hellas, Greece
Moid Sandhu, University of Queensland | Data61 - Commonwealth Scientific and Research Organization (CSIRO), Australia
José Santa, Technical University of Cartagena, Spain
Peter Schneider-Kamp, University of Southern Denmark, Denmark
Floriano Scioscia, Polytechnic University of Bari, Italy
Luca Sciuillo, University of Bologna, Italy
Hugo Sereno Ferreira, University of Porto, Portugal
Alireza Shahrabi, Glasgow Caledonian University, Scotland, UK
Jianchen Shan, Hofstra University, USA
Ahmed S. Shatnawi, Jordan University of Science and Technology, Jordan
Shih-Lung Shaw, University of Tennessee, Knoxville, USA
Haichen Shen, Amazon Web Services, USA
Michael Sheng, Macquarie University, Australia
Shouqian Shi, University of California, Santa Cruz, USA
Matteo Signorini, Nokia Bell Labs, France
Sandeep Singh Sandha, University of California-Los Angeles, USA
Rute C. Sofia, fortiss GmbH, Munich, Germany
Francesco Soldovieri, Institute for Electromagnetic Sensing of the Environment | CNR, Italy
Christoph Stach, University of Stuttgart, Germany
Álvaro Suárez Sarmiento, University of Las Palmas de Gran Canaria, Spain
K. Subramani, West Virginia University, USA
Apostolos Syropoulos, Greek Molecular Computing Group, Xanthi, Greece
Violet R. Syrotiuk, Arizona State University, USA
Yoshiaki Taniguchi, Kindai University, Japan
Sudeep Tanwar, Institute of Technology | Nirma University, India
Adrian Tarniceriu, Securecell, Switzerland
Angelo Trotta, University of Bologna, Italy
Takeshi Tsuchiya, Suwa University of Science, Japan
Sudhanshu Tyagi, Thapar Institute of Engineering & Technology, India / Jan Wyykowski University Polkowice, Poland
Hamed Vahdat-Nejad, University of Birjand, Iran
K. Vasudevan, IIT Kanpur, India
Miroslav N. Velez, Aries Design Automation, USA
Thierry Villemur, LAAS-CNRS | University of Toulouse, France
Halyna Vlasiuk, National University of Water and Environmental Engineering, Ukraine
Luping Wang, The Hong Kong University of Science and Technology (HKUST), Hong Kong
Xianzhi Wang, University of Technology Sydney, Australia

Hongyi "Michael" Wu, Professor, Old Dominion University, USA
Kesheng John Wu, Lawrence Berkeley National Laboratory, USA
Yuan Wu, University of Macau, Macau SAR, China
De-Nian Yang, Institute of Information Science - Academia Sinica, Taiwan
Fanghua Ye, University College London, UK
Jian Yu, Auckland University of Technology, New Zealand
Xiaojun (Jenny) Yuan, University at Albany, State University of New York, USA
Dong Zhang, Institute of Electrical Engineering - Chinese Academy of Sciences, China

Copyright Information

For your reference, this is the text governing the copyright release for material published by IARIA.

The copyright release is a transfer of publication rights, which allows IARIA and its partners to drive the dissemination of the published material. This allows IARIA to give articles increased visibility via distribution, inclusion in libraries, and arrangements for submission to indexes.

I, the undersigned, declare that the article is original, and that I represent the authors of this article in the copyright release matters. If this work has been done as work-for-hire, I have obtained all necessary clearances to execute a copyright release. I hereby irrevocably transfer exclusive copyright for this material to IARIA. I give IARIA permission to reproduce the work in any media format such as, but not limited to, print, digital, or electronic. I give IARIA permission to distribute the materials without restriction to any institutions or individuals. I give IARIA permission to submit the work for inclusion in article repositories as IARIA sees fit.

I, the undersigned, declare that to the best of my knowledge, the article does not contain libelous or otherwise unlawful contents or invading the right of privacy or infringing on a proprietary right.

Following the copyright release, any circulated version of the article must bear the copyright notice and any header and footer information that IARIA applies to the published article.

IARIA grants royalty-free permission to the authors to disseminate the work, under the above provisions, for any academic, commercial, or industrial use. IARIA grants royalty-free permission to any individuals or institutions to make the article available electronically, online, or in print.

IARIA acknowledges that rights to any algorithm, process, procedure, apparatus, or articles of manufacture remain with the authors and their employers.

I, the undersigned, understand that IARIA will not be liable, in contract, tort (including, without limitation, negligence), pre-contract or other representations (other than fraudulent misrepresentations) or otherwise in connection with the publication of my work.

Exception to the above is made for work-for-hire performed while employed by the government. In that case, copyright to the material remains with the said government. The rightful owners (authors and government entity) grant unlimited and unrestricted permission to IARIA, IARIA's contractors, and IARIA's partners to further distribute the work.

Table of Contents

Self-Navigating Mobile Robot for Assisting with Load Transportation for Physically Limited Users <i>Marcio Mendonca, Wagner Fontes Godoy, Luigi L. Saito, Rodrigo H. C. Palacios, Janaina F. S. Goncalves, and Marco Antonia M. Laia</i>	1
Improving Texts into Powerful Communication Tools with Classic Techniques and Deep Learning <i>Wagner F. Godoy, Jose A. Fabri, Rodrigo H. C. Palacios, Marcio Mendonca, Janaina F. S. Goncalves, and Luiz O. M. Moraes</i>	5
Attraction-Based Reinforcement Learning: A Real-Time Approach Using Techniques Based on the Animal Behavior <i>Marcos Laia, Marcio Mendonca, Edimilson Santos, Rodrigo Palacios, Wesley Guimaraes, and Janaina Goncalves</i>	10
Using Fuzzy Cognitive Maps and Chatbots to Evaluate Student Satisfaction in a University: A Comparison between Strong and Weak AI <i>Wagner F. Godoy, Jose A. Fabri, Rodrigo H. C. Palacios, Marcio Mendonca, Janaina F. S. Goncalves, and Luiz O. M. Moraes</i>	16
Leveraging Large Language Models for Enhanced Personalised User Experience in Smart Homes <i>Jordan Rey-Jouanchicot, Andre Bottaro, Eric Campo, Jean-Leon Bouraoui, Nadine Vigouroux, and Frederic Vella</i>	21
One Cash: A New Generation of E-wallet for Mobile Money <i>Macaire Ngomo and Clavy Viannet Mbou Mouyele</i>	28
Real-Time Emotion Assessment System in Smart Classrooms Using Wearable Bracelets <i>Edgar Batista, Laia Cot, Valeria Perez, and Antoni Martinez-Balleste</i>	36

Self-Navigating Mobile Robot for Assisting with Load Transportation for Physically Limited Users

Márcio Mendonça, Wagner F. Godoy, Luigi L. Saito
 Academic Department of Electrical Engineering
 Technological Federal University of Paraná (UTFPR)
 Cornélio Procópio, Paraná - Brazil
 e-mail: mendonca@utfpr.edu.br,
 wagner.godoy@utfpr.edu.br, luigisaito@alunos.utfpr.edu.br

Rodrigo H. C. Palácios, Janaína F. S. Gonçalves,
 Marcos A M. Laia
 PPG of Mechanical Engineering - CP/PG
 Technological Federal University of Paraná (UTFPR)
 Cornélio Procópio, Paraná - Brazil
 e-mail: rodrigopalacios@utfpr.edu.br,
 janainaf@utfpr.edu.br, marcoslaia@gmail.com

Abstract—Robotics is an expanding field that aims to automate tasks traditionally performed by humans using machines. Mobile robotics involves using mechanical system dynamics to execute tasks, with sensors gathering environmental data to navigate and perform these tasks. Assistive robotics uses autonomous robots to aid individuals with disabilities or limitations, such as transporting loads for pregnant women and the elderly. This paper details the development of a low-cost autonomous mobile robot capable of following its user and avoiding obstacles, analogous to a leader-following robot. The system employs a fuzzy controller and a state machine. A simulation and a prototype were created, demonstrating the ability to track a target and avoid obstacles in real time. The study concludes with recommendations for future research.

Keywords—*Technological applications in social environments; Service robotics; Autonomous robotics.*

I. INTRODUCTION

For many researchers, robotics can be divided into two main areas: programmed and artificial intelligence. Programmed robotics involves systems that require reprogramming whenever the task changes [9]. In contrast, intelligent robotics integrates aspects of Artificial Intelligence (AI), enabling robots to learn and adapt autonomously. This integration allows machines to emulate human reasoning, analyze situations, and make decisions, ultimately aiming to mimic or even surpass human intelligence [5].

The application of robotics in medicine has evolved significantly from performing early repetitive tasks to providing substantial patient benefits. Surgical robots enhance precision during operations, reducing trauma to patients and extending the capabilities of surgeons [1]. Beyond surgical applications, mobile nursing robots have been developed to assist elderly patients with minor physical and cognitive limitations, thereby supporting patients and the nursing staff [3].

The work in [11] introduces a robust and flexible control system for mobile robots, leveraging motion description language in behaviours-based robotics. This approach

effectively manages the control of mobile robots by describing the interactions between discrete and continuous dynamics within a robotic system. Additionally, it quantitatively measures the complexity of the steering a robot. The study focuses on the problem of object-following by nonholonomic wheeled mobile robots using this method. The effectiveness of the proposed approach is validated through MATLAB simulations, demonstrating its capability to handle the object-following task successfully.

In [10], clinical and healthcare Knowledge Management (KM) has garnered significant global attention as an emerging discipline in recent years. This approach integrates a broad spectrum of interconnected themes, including clinical informatics, clinical governance, artificial intelligence, privacy and security, data mining, genomic mining, information management, and organizational behaviour. This paper presents critical manuscripts exploring case studies and KM applications in healthcare and clinical settings.

By employing motion description language, the system captures the intricate dynamics of robot behaviours and provides a structured framework for implementing and assessing control strategies. This method offers a significant advancement in the control of mobile robots, particularly in environments requiring precise navigation and object tracking. The MATLAB simulation results confirm the practicality and efficiency of this control approach, showcasing its potential for broader applications in mobile robotics.

The integration of motion description language in behaviours-based robotics presents a comprehensive solution for controlling nonholonomic wheeled mobile robots, with proven effectiveness in object-following tasks, as demonstrated by simulation results. Future work could explore the application of this method to more complex robotic systems and dynamic environments.

Initially, mobile robotics focused on guided object transportation within production lines. However, contemporary advancements have enabled autonomous mobile robots to operate effectively in both factory settings

and open environments. These robots employ sophisticated processes such as perception, localization, trajectory planning, and movement execution [2]. To navigate and track targets within obstacle-rich environments, they utilize various sensors, including ultrasonic sensors and cameras, to gather positional data essential for trajectory planning [8].

The primary aim of this study is to develop an autonomous robot capable of transporting small loads, such as luggage, for individuals with physical limitations, such as the elderly and pregnant women. This goal is aligned with applications in service and assistive robotics, which aim to alleviate physical strain on these individuals. Additionally, the study underscores the ergonomic benefits of adjusting working conditions to enhance human efficiency and well-being. By mitigating ergonomic risks, such as physical strain, improper posture, and repetitive tasks, the study highlights the potential for robotics to improve overall human health and productivity in various settings [10].

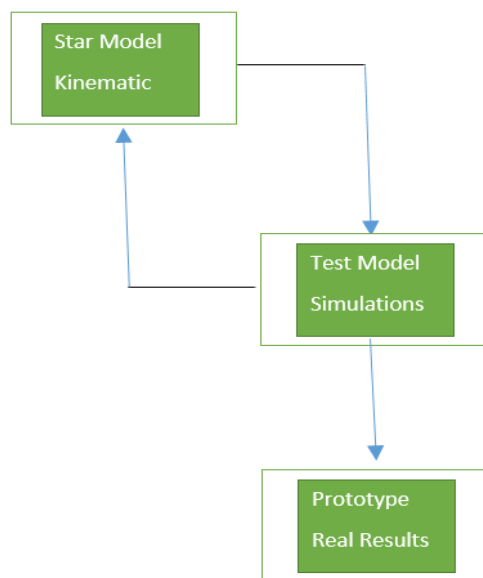


Figure 1. Development Prototype.

The diagram in Figure 1 outlines a process beginning with the "Star Model Kinematic" as the initial phase. This is followed by the "Test Model Simulations," where simulations will be executed exhaustively. The next step is the "Prototype Real Results" stage, which focuses on implementing and validating the prototype. It is important to note that the prototype testing should include a maturation period to ensure that no unexpected behavior occurs. Additionally, there is a feedback loop between "Test Model Simulations" and "Star Model Kinematic," suggesting iterative refinement based on the simulation results.

The results are preliminary, obtained through a simulated experiment, and a second phase will likely be implemented, incorporating another knowledge base with fuzzy logic in a smaller search space. The aim at the end of the first search is

to obtain initial results from this research and then proceed to the second phase, at which point the development of the prototype will begin.

II. DEVELOPMENT

This work aims to create a low-cost mobile robot capable of transporting loads up to 10 kg for individuals with physical limitations. The robot must follow the user, avoid obstacles, and prevent collisions in real time. The development involves building the robot's structure, including chassis, wheels, motors, and its electronics—comprising H-bridge, microcontroller or computer, and sensors (ultrasonic sensor or camera). Programming will interpret sensor data and make decisions.

The system will determine the distance between the robot and the target using ultrasonic sensors or a camera. This distance will inform the pulses sent to the robot's motors, adjusting speed based on proximity to the target. Fuzzy or multivalued logic will control the output, simplifying the modeling compared to Proportional-Integral-Derivative (PID) controllers, especially for multivariable systems [7].

Target identification may use markers based on color and shape, like in the approach in [4]. A safety system will be developed to avoid obstacles, either by manoeuvring around static obstacles or waiting for dynamic obstacles to clear. The embedded system must handle the computational demands, with options like Arduino for simplicity and cost or Raspberry Pi for higher computational capacity.

III. RESULTS AND DISCUSSIONS

A. Simulation

Control model validation involved MATLAB simulations, comparing fuzzy and multivalued logic controllers—the fuzzy controller determined wheel pulses based on the Euclidean distance between the robot and the target. Simulation results showed satisfactory performance for target tracking with acceptable computational complexity. Table I presents the functions used as the basis for the rules of the developed fuzzy controller: the Euclidean distance between the robot and the target and the increment given to the pulse sent to the robot's wheels.

TABLE I. RULE BASE OF THE FUZZY CONTROLLER

<ol style="list-style-type: none"> 1. If (hi is close) then (const is close) 2. If (hi is medium) then (const is medium) 3. If (hi is medium far) then (const is medium plus) 4. If (hi is far) then (const is medium)
--

Figure 2 shows the structure of the fuzzy system. Figure 3, in turn, presents the surface generated by the fuzzy controller. In Figure 2, the inputs are, a priori, the distance of the person in the x and y-axis, and the outputs are the pulses given to both wheels of the kinematic model.

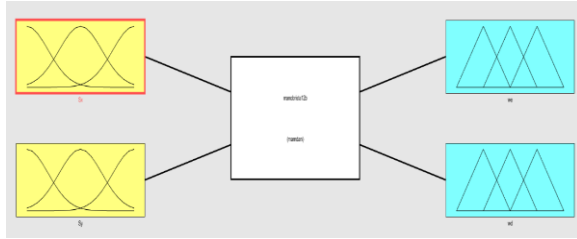


Figure 2. Fuzzy Controller Structure.

The diagram illustrates a process where two input signals (labeled 1 and 3, represented by sine waves) are fed into a central module, labeled "network", which then outputs two results (labeled 2 and 4, represented by triangular waves). The inputs and outputs are connected to the network through lines, indicating data flow or signal processing.

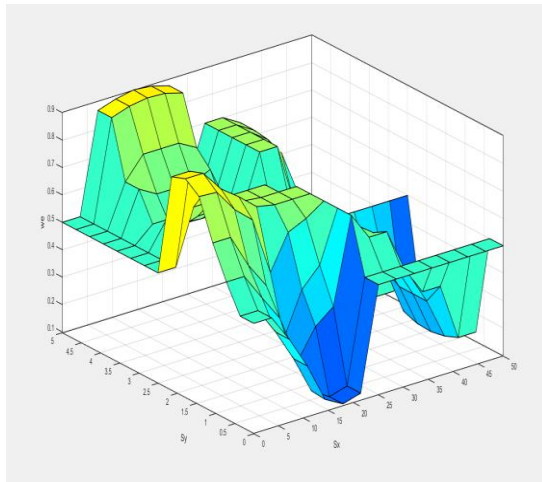


Figure 3. Fuzzy Controller Surface.

The preliminary outcomes from the first simulation conducted on the mobile robot are illustrated in Figure 4. This figure visually represents the robot's movement within the designated environment. Figure 5, correspondingly, presents the results derived from the second simulation, offering a comparative analysis of the robot's performance under modified conditions. In both figures, the black point represents the mobile robot's current position, while the orange triangle denotes the target or goal location. The blue line traces the robot's trajectory as it navigates the environment. The graph's axes indicate distances measured in centimeters; however, the scale can be adjusted to meters when applied to larger environments. This flexibility in scale ensures the applicability of the simulation results across a range of spatial domains, enabling robust analysis in both confined and extensive operational settings.

These graphical depictions illustrate not only the robot's efficiency in path planning and execution, but also provide a foundation for further adjustments in navigation algorithms, should the need arise based on environmental complexity or task demands.

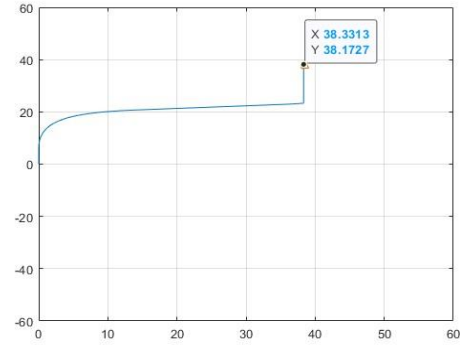


Figure 4. First robot simulation.

The results obtained, although preliminary, were satisfactory for target pursuit. For small, stipulated error values, the computational complexity was not compromised. Subsequently, this control logic will be implemented in the prototype mentioned in introduction section.

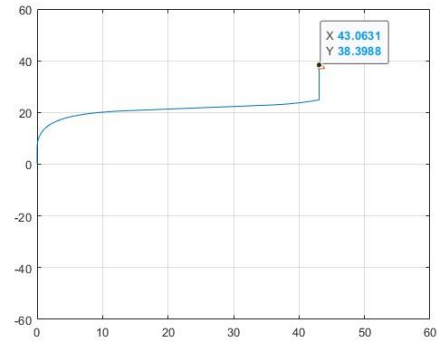


Figure 5. Second robot simulation.

B. Prototyping

A prototype was developed using Arduino UNO R3 and an ESP-32 camera. Initial tests demonstrated satisfactory target detection capabilities; however, software stability issues were identified, indicating a need for further improvements to enhance system robustness. Figure 6 showcases the initial prototype.

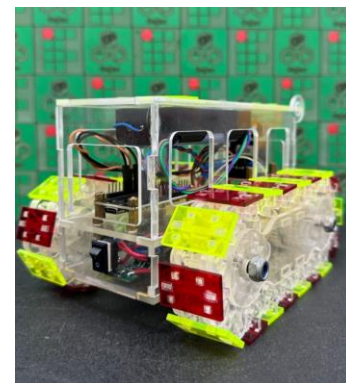


Figure 6. Mobile robot prototype.

The marker detection tests are illustrated in Figure 7 and Figure 8. These tests involved assessing the system's ability to accurately identify and track predefined markers. While the results were promising, the observed software instabilities highlight the necessity for additional refinement and optimization.

Future efforts will focus on addressing these stability issues and improving the overall robustness of the system. This includes optimizing the software algorithms and potentially upgrading the hardware components to ensure more reliable performance in real-world applications. Enhancing the system's stability and reliability is crucial for its effective deployment in various practical scenarios.



Figure 7. Marker filtering.



Figure 8. Marker detection.

IV. CONCLUSION

This study proposed the development of an autonomous robot designed to transport loads. The system integrated a camera and ultrasonic sensors to identify targets effectively, ensuring the prototype responded accurately to control commands. This combination of sensors facilitated precise navigation and obstacle avoidance, demonstrating the robot's potential for real-world applications.

Future work aims to enhance the system's robustness by implementing embedded systems with higher processing power. This upgrade will allow for more complex

computations and faster data processing, improving the robot's efficiency and reliability. Additionally, further research will explore advanced algorithms for better decision-making and adaptive learning capabilities, enabling the robot to operate in more dynamic and unpredictable environments.

By leveraging state-of-the-art technologies and continuously improving the system, this autonomous load-transporting robot holds promise for various applications, including logistics, healthcare, and personal assistance, significantly enhancing productivity and quality of life.

V. ACKNOWLEDGMENT

The authors would like to thank the Federal University of Technology – CP/PR for the financial support provided for this research

REFERENCES

- [1] A. M. Okamura, M. J. Mataric, and H. I. Christensen, "Medical and Health-Care Robotics," *IEEE Robotics & Automation Magazine*, vol. 17, no. 3, pp. 26-37, Sept. 2010.
- [2] E. Garcia, M. A. Jimenez, P. G. De Santos, and M. Armada, "The evolution of robotics research," *IEEE Robotics & Automation Magazine*, vol. 14, no. 1, pp. 90-103, March 2007.
- [3] J. Pineau, M. Montemerlo, M. Pollack, N. Roy, and S. Thrun, "Towards robotic assistants in nursing homes: challenges and results," *Robotics and Autonomous Systems*, vol. 42, pp. 271-281, 2003.
- [4] M. N. A. Bakar, R. Nagarajan, and A. R. M. Saad, "Development of a doctor following mobile robot with Monovision based marker detection," *2011 IEEE Applied Power Electronics Colloquium (IAPEC)*, 2011, pp. 86-91, doi: 10.1109/IAPEC.2011.5779843.
- [5] M. J. Mataric, "The robotics primer," *Choice Reviews Online*, vol. 45, no. 06, pp. 45-3222-45-3222, Feb. 2008.
- [6] L. Oda, S. Ávila et al., *Biossegurança em Laboratórios de Saúde Pública - Biosafety in Public Health Laboratories*, Brasília, Ministério da Saúde, 1998.
- [7] K. M. Passino and S. Yurkovich, *Fuzzy Control*, 1st ed., Addison-Wesley, 1998, pp. 1-500.
- [8] Y. Meng, X. Liu, and Z. Liang, "A Robust Method for Mobile Robot Tracking Object on Behaviour-Based Robotics," *2007 First IEEE International Symposium on Information Technologies and Applications in Education*, 2007, pp. 468-472.
- [9] M. Mendonça et al., "Multi-robot exploration using Dynamic Fuzzy Cognitive Maps and Ant Colony Optimization," *IEEE International Conference on Fuzzy Systems*, Glasgow, UK, 2020, pp. 1-8, doi: 10.1109/FUZZ48607.2020.9177814.
- [10] R. K. Bali, D. D. Feng, F. Burstein, and A. N. Dwivedi, "Guest Editorial Introduction to the Special Issue on Advances in Clinical and Health-Care Knowledge Management," *IEEE Transactions on Information Technology in Biomedicine*, vol. 9, no. 2, pp. 157-161, June 2005, doi: 10.1109/TITB.2005.849395.
- [11] Y. Meng, X. Liu, and Z. Liang, "A Robust Method for Mobile Robot Tracking Object on Behavior-Based Robotics," *First IEEE International Symposium on Information Technologies and Applications in Education*, Kunming, China, 2007, pp. 468-472, 2007.

Improving Texts into Powerful Communication Tools with Classic Techniques and Deep Learning

Wagner F. Godoy, José A. Fabri, Rodrigo H. C. Palácios

Academic Department of Electrical Engineering
Technological Federal University of Paraná (UTFPR)
Cornélio Procópio, Paraná - Brazil
email:wagner.godoy@utfpr.edu.br,fabri@utfpr.edu.br
rodrigopalacios@utfpr.edu.br

Márcio Mendonça, Janaína F. S. Gonçalves, Luiz O. M. Moraes

PPG of Mechanical Engineering - CP/PG
Technological Federal University of Paraná (UTFPR)
Cornélio Procópio, Paraná – Brazil
e-mail: mendonca@utfpr.edu.br, janainaf@utfpr.edu.br,
luizotaviomoraes@alunos.utfpr.edu.br

Abstract— Investing in transforming written content into powerful communication tools involves utilizing both classical techniques and advancements in deep learning. Classical methods offer a wealth of time-tested strategies, while deep learning technologies provide modern insights for crafting messages with more significant impact. Combining these approaches can significantly enhance the effectiveness and reach of any text. The intricacies of language and communication are explored through classical methodologies, while deep learning algorithms bring predictive capabilities and adaptability. Elevating texts into compelling communication instruments requires integrating traditional knowledge with modern technology, resulting in messages that captivate, persuade, and inspire across various platforms. The synergy between classical approaches and deep learning allows texts to transcend mere words, becoming dynamic tools of connection and influence in today’s communication landscape. This article concludes with a summary of key findings and suggestions for future research.

Keywords-Textual Analysis, Machine Learning, Text Summarization.

I. INTRODUCTION

With the growing volume of online information, text summarization has become an essential field for distilling the key messages of textual content. Manual summarization of large documents is challenging, making automatic text summarization increasingly important. Text summarization refers to the process of automatically condensing a document into a shorter version while preserving its core meaning and content. It has become a significant area of research in Natural Language Processing (NLP), attracting considerable attention from researchers. There is a close relationship between text mining and summarization, and summarization systems must be tailored to the type of input text. This study explores the connection between text mining and summarization, reviews summarization approaches and key parameters for extracting relevant sentences and presents the main stages and criteria for summarization. Lastly, it highlights fundamental evaluation methods for assessing summarization quality [7].

Text summarization is summarizing text documents while preserving their context and overall content. A good summary

should highlight the fundamental concepts of any textual document. Text summarization is a fundamental area of Natural Language Processing (NLP), which uses various tools to extract relevant information from the given text [5].

Automatic text summarization is essential to Natural Language Processing (NLP), a subpart of the Artificial Intelligence domain. The widespread usage of text summarization is due to the massive use of the Internet in every aspect of life. In this research article, we perform a statistical text summarization technique.

Uniting classic methodologies with the revolutionary advances of deep learning opens the doors to communication with significant impact. The mix of traditional strategies and cutting-edge algorithms can be used to give each word power and meaning [2]. This holistic approach, which fuses the wisdom of the past with cutting-edge technology, is the foundation for creating messages that captivate, persuade, and inspire.

The rest of the paper is structured as follows. Section 2 compares two of the largest deep neural networks, the Gemini and ChatGPT 3.5 and 4.0. Section 3 discusses classical techniques and those based on deep neural networks. Section 4 contains concepts and fundamentals of summarization with deep learning and examples of applying the tools in two texts. Finally, Section 5 closes the work with a conclusion and addresses future work.

II. GEMINI AND CHATGPT 3.5 AND 4.0

Textual analysis has been a central focus of machine learning studies, as noted in [2]. It can be exemplified by platforms like Bard's Gemini Pro and ChatGPT's free version, which primarily differ in the data sources used during their training [3]. Gemini can process multiple data formats simultaneously, including text, code, audio, static images, and videos, allowing for a more comprehensive manipulation of diverse information types before generating feedback. Additionally, the accessible version of Gemini offers real-time Internet access.

In contrast, GPT-3.5, the public version of ChatGPT, is trained on a static dataset that has not been updated since 2021. A detailed comparison between Gemini and its competitors remains challenging, as Google has not yet fully

released Gemini's capabilities for public testing. Given that Gemini is still in development, its features are subject to change. However, according to tests published by Google, the Ultra version of Gemini surpassed OpenAI's ChatGPT-4 in various tasks, achieving a 90% success rate in a test involving 57 subjects, compared to 86.4% for ChatGPT-4. ChatGPT and Gemini are Large Language Models (LLMs) with remarkable capabilities, yet they differ significantly in focus, training data, and functionality.

The ability of these tools to emulate summarization highlights the potential of deep learning in handling large volumes of data and extracting relevant information efficiently. Gemini and ChatGPT utilize deep neural networks to identify patterns and relationships across different data sources, enabling them to generate concise and coherent summaries. These deep learning networks are trained to understand the meaning of words and the context in which they appear, making them powerful tools for performing summarizations with a high degree of accuracy and adaptability.

In short, deep learning offers significant advantages in summarization tasks due to its ability to capture context and nuances in language, making it more adaptable to diverse data formats and large volumes of information. Unlike classical techniques, which often rely on predefined rules and require manual feature engineering, deep learning models can automatically learn relevant features, streamlining the process. However, deep learning comes with its challenges. It demands substantial computational resources and training data, and its models are often less interpretable, functioning as "black boxes." Additionally, there is a risk of overfitting, where models may perform well on training data but need help with new inputs.

On the other hand, classical summarization techniques are generally more efficient, faster to implement, and easier to understand due to their more uncomplicated, rule-based nature. They also require fewer computational resources. However, these methods must be revised to handle complex linguistic contexts and be more scalable, particularly for large or diverse datasets. Ultimately, the choice between deep learning and classical techniques depends on the task's specific needs, balancing accuracy, efficiency, and resource availability. In the next section, more details will be provided.

Other Differences:

While typical sequence-to-sequence models with attention mechanisms have shown promising results in abstractive text summarization, they often need help with several things that could be improved. These include Out-Of-Vocabulary (OOV) issues, occasional repetition, and a tendency towards lower overall quality [6].

By applying these techniques, researchers and analysts can extract valuable information and insights from large volumes of text. Automatic text summarization is a rapidly expanding area in Artificial Intelligence, driven by the transformative power of Deep Learning. This innovative technique allows the generation of accurate and concise summaries of extensive

texts, capturing their key points and essential information autonomously.

A. What is Text Summarization with Deep Learning?

Text summarization using deep learning surpasses traditional rule-based and heuristic approaches, which typically yield shallow, context-insensitive summaries. By leveraging machine learning, deep learning architectures such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) are trained on extensive textual datasets, enabling them to capture complex relationships between words, sentences, and paragraphs. This advanced linguistic comprehension allows these models to generate summaries that are not only more informative and coherent, but also more faithful to the original meaning of the text.

B. How Summarization Works with Deep Learning?

The summarization process using deep learning consists of several essential and structured steps that ensure the generation of accurate and coherent summaries. First, in preprocessing, the input text is cleaned to remove noise such as punctuation, special characters, and unnecessary formatting, followed by tokenization into individual words or sub-words for subsequent analysis. In the encoding phase, these tokenized words are transformed into numerical representations using techniques like word embeddings (for example, Word2Vec, GloVe) or contextual embeddings (for example, BERT, GPT), enabling the model to process the text computationally. Next, during modeling, the deep learning model—trained on large corpora of text and their corresponding summaries—processes the encoded sequence, identifying the most relevant sentences, phrases, or paragraphs by learning patterns and relationships within the data. Finally, in the decoding phase, the selected numerical representations are converted back into text, forming the final summary. Depending on the approach, this step can involve extractive summarization, where key sentences are retained verbatim, or abstractive summarization, where the model generates new sentences that capture the meaning of the original content.

C. Advantages of Summarization with Deep Learning

Accuracy: Deep learning models produce highly precise and informative summaries, capturing the essential elements of the source text with exceptional fidelity.

Relevance: These models can select the most contextually pertinent information, aligning with the specific goals and purpose of the summarization task.

Cohesion: Summaries generated by deep learning exhibit a high degree of linguistic cohesion and grammatical correctness, ensuring fluidity and readability.

Flexibility: The adaptability of deep learning allows for the generation of summaries in various lengths and styles, tailored to the specific needs of users or applications.

Efficiency: The automation of the summarization process via deep learning models significantly reduces manual effort and time, enhancing overall operational efficiency.

D. Core Concepts and Applications of Summarization with Deep Learning

Text summarization powered by deep learning has broad applicability across numerous fields:

News and Media: Enables the rapid generation of concise and informative summaries of news articles, reports, and media content.

Academic Research: Assists in synthesizing literature reviews by identifying key articles and producing comprehensive research summaries.

Legal Documents: Facilitates the creation of accurate and concise summaries of complex legal texts, such as contracts, regulations, and case laws.

Customer Service: Automates the summarization of support tickets, FAQs, and other service-related documents to enhance customer interaction efficiency.

Education: Supports the development of personalized learning materials, including book summaries, study guides, and educational articles, promoting more efficient knowledge dissemination.

III. RESULTS

Two texts will be used to instantiate the application of a classical and machine learning technique.

A. Text 1 - Robotics is the Science

“Robotics is the science that studies the technologies associated with the design and construction of robots. Robots are automatic mechanisms that use integrated circuits to perform simple or complex human activities and movements. Robotics has wide application in several areas, from industrial production, medicine to domestic activities. The concept of robotics emerged at the beginning of the 20th century, in the work “The Liar” by science fiction author Isaac Asimov. It was the author who created the word ‘robotics’ and it was also he who presented, in the field of science fiction, the Laws of Robotics.”

Text summarization by frequency, order, and context of words in the text will be performed.

Textual Analysis

Robotics: 4 mentions (title, 1st sentence, 2nd sentence, last sentence) - Keyword that defines the main theme of the text.

Science: 1 mention (1st sentence) - Describes the nature of robotics as a field of study.

Technologies: 1 mention (1st sentence) - Covers the tools and knowledge used in robotics.

Conception: 1 mention (1st sentence) - Refers to the robot creation and planning stage.

Construction: 1 mention (1st sentence) - Indicates the manufacturing and assembly phase of the robots. Robots: 3 mentions (1st sentence, 2nd sentence, last sentence) - Central term of the text, defines the autonomous machines in question.

Mechanisms: 1 mention (2nd sentence) - Emphasizes the physical and functional nature of robots. Automatics: 1 mention (2nd sentence) - Highlights the ability of robots to operate without constant manual intervention.

Integrated circuits: 1 mention (2nd sentence) - Mentions the electronic components that allow the control of robots.

Activities: 1 mention (2nd sentence) - Covers the tasks that robots can perform.

Movements: 1 mention (2nd sentence) - Refers to the physical actions performed by the robots.

Humans: 1 mention (2nd sentence) - Establishes a comparison between the capabilities of robots and humans

Application: 1 mention (3rd sentence) - Indicates the practical use of robotics in various sectors.

Areas: 1 mention (3rd sentence) - Expands the scope of application of robotics, from industry to the home.

Industrial production: 1 mention (3rd sentence) - Exemplifies the application of robotics in manufacturing.

Medicine: 1 mention (3rd sentence) - Highlights the use of robotics in medical procedures.

Domestic activity: 1 mention (3rd sentence) - Indicates the presence of robotics in the residential environment.

Concept: 1 mention (4th sentence) - Refers to the idea of robotics as a field of study. Early 20th century: 1 mention (4th sentence) - Places the origin of the concept in time.

Work: 1 mention (4th sentence) - Indicates a specific work that addressed the topic.

"The Liar": 1 mention (4th sentence) - Title of the work that explored the concept of robotics. Isaac Asimov: 1 mention (4th sentence) - Name of the author who contributed to the development of the concept.

Author: 1 mention (5th sentence) - Returns to the figure of Isaac Asimov as responsible for something. Word: 1 mention (5th sentence) - Indicates the term "robotics" as Asimov's creation.

Laws of Robotics: 1 mention (5th sentence) - Highlights a set of principles related to robotics, devised by Asimov.

With this classic technique, the text would look like:

"Robotics, the science of technologies for designing and building robots, has wide applications in industrial, medical, and domestic activities. The concept emerged in the 20th century in the work "The Liar" by Isaac Asimov, author of science fiction, who also introduced the Laws of Robotics."

Resume by deep learning 1 Microsoft (ChatGPT 3.5 free):

"Robotics studies technologies to create robots, automatic mechanisms that perform human tasks. It has applications in production, medicine, and domestic activities. The concept emerged at the beginning of the 20th century in work "The Liar" by Isaac Asimov, the author who also created the word "robotics" and the Laws of Robotics in science fiction".

Resume by deep learning 2 (Gemini free - Google):

“Uncovering the World of Robotics: Robotics presents itself as a fascinating science that explores the creation and development of autonomous machines capable of performing everything from simple tasks to complex movements that imitate humans. These mechanisms come to life through integrated circuits and become valuable tools in several areas”.

Resume by deep learning 3 Microsoft (ChatGPT4):

“Robotics is the study of technologies for designing and building robots, automatic mechanisms that employ integrated circuits to perform everything from simple tasks to complex movements without direct human intervention. With extensive applications in sectors such as industrial production, medicine, and domestic tasks, robotics is a discipline that emerged at the beginning of the 20th century, as highlighted by the work "The Liar" by Isaac Asimov, a pioneer of science fiction.”

*B. Text 2: The Myth of the Cave “Inspirational text”
(Plato)*

The myth tells about prisoners (since birth) who live trapped in chains in a cave and who spend all their time looking at the back wall, which is illuminated by the light generated by a fire. Shadows of statues representing people, animals, plants, and objects are projected onto this wall, showing everyday scenes and situations. The prisoners keep giving names to the images (shadows), analysing and judging the situations. Let us imagine that one of the prisoners was forced to escape his chains to explore the cave and the outside world [4]. He would get in touch with reality and realize that he spent his entire life analysing and judging only images projected by statues. When leaving the cave and meeting the real world, he would be enchanted by real beings, nature, animals, etc. He would return to the cave to pass on all the knowledge acquired outside the cave to his colleagues who were still trapped. However, he would be ridiculed when he told everything he saw and felt, as his colleagues can only believe in the reality they see on the illuminated wall of the cave. The prisoners will call him crazy, threatening him with death if he does not stop talking about those ideas considered absurd.

Prisoners in the cave since birth, chained and limited to the perception of shadows on the wall (symbolizes ignorance and limitation of sensory knowledge). Awakening and Search for Truth: A freed prisoner explores the cave and the outside world, discovering actual reality (represents the search for knowledge and philosophy). The prisoner returns to the cave to share his knowledge but is ridiculed and misunderstood by the other prisoners (evidence resistance to change and attachment to preconceived beliefs). The myth serves as a metaphor for the human condition, trapped in one's own

beliefs and perceptions limited, with the search for truth being an arduous and often lonely challenge.

*Textual Analysis**Frequent Words:*

Cave (4x): Place where prisoners are imprisoned.
Prisoners (3x): People who live chained in the cave.
Shadows (3x): Images projected on the cave wall.
Reality (2x): True world outside the cave. Truth (2x): Real knowledge of the world.
Ignorance (2x): Lack of knowledge of the truth.
Light (2x): Source of lighting for the cave and the outside world.
World (2x): Refers to both the cave and the outside.
Belief (2x): Conviction based on limited perceptions.
Freedom (1x): Possibility of escaping the cave and knowing the truth.
Knowledge (1x): Acquisition of information and understanding of reality. Plato proposes that most people live like prisoners, trapped in their limited beliefs and perceptions, without access to the truth.

With this classic technique, the text would look like:

“The myth of the cave is a metaphor for the human condition. Plato proposes that most people live like prisoners, trapped in their limited beliefs and perceptions, without access to the truth. The search for knowledge and freedom from ignorance requires effort and overcoming internal and external resistance.”

Resume with deep learning 2 (Google):

“Prisoners in the cave since birth, chained and limited to the perception of shadows on the wall (symbolizes ignorance and limitation of sensory knowledge). A freed prisoner explores the cave and the outside world, discovering actual reality (represents the search for knowledge and philosophy). The prisoner returns to the cave to share his knowledge but is ridiculed and misunderstood by the other prisoners (evidence resistance to change and attachment to preconceived beliefs). The myth serves as a metaphor for the human condition, trapped in our limited beliefs and perceptions, with the search for truth being an arduous and often lonely challenge”.

Resume by deep learning 3 (ChatGPT4):

“Plato's myth of the cave describes prisoners chained from birth in a cave, where they see only shadows cast on the wall lit by a fire. When a prisoner is forced out of the cave and discovers the actual reality outside it, he returns to share his knowledge but is ridiculed by the other prisoners who only believe in the shadow reality. This myth metaphorizes the human condition of being trapped in limited perceptions and

established beliefs, highlighting the search for truth as a lonely and often misunderstood challenge.”

IV. POSSIBLE METRICS FOR QUALITY ANALYSIS TEXT SUMMARIZATION

Several potential metrics can be used to evaluate the quality of text summarization [7], particularly in deep learning contexts [8]:

- ROUGE (Recall-Oriented Understudy for Gisting Evaluation): ROUGE measures the overlap between the generated summary and reference summaries regarding n-grams, word sequences, or word pairs. ROUGE-N, ROUGE-L, and ROUGE-S are commonly used to assess recall, precision, and F1-score, making it a standard metric for summarization accuracy.

- BLEU (Bilingual Evaluation Understudy): While initially developed for machine translation, BLEU is sometimes used to evaluate summarization by comparing n-gram overlap between the generated summary and reference text. It focuses on precision, making it useful for measuring how much of the generated summary covers the reference text.

- METEOR (Metric for Evaluation of Translation with Explicit Ordering): METEOR extends BLEU by incorporating stemming and synonym matching, improving its ability to capture semantic equivalence. It helps evaluate the semantic accuracy of summaries.

- Content Overlap: This metric assesses how well the summary captures the essential content of the original text. It can involve manual or automated comparison of critical ideas, entities, or facts mentioned in the original and the summary.

- Compression Ratio: This evaluates the efficiency of summarization by measuring the summary length relative to the original text. A good compression ratio balances brevity with informativeness, ensuring key points are retained without unnecessary verbosity.

- Coherence and Fluency: These subjective metrics assess how logically connected the sentences in the summary are and how natural the language sounds. Human evaluators often rate summaries based on linguistic cohesion and overall readability.

- Semantic Similarity: Using vector-based methods such as cosine similarity, this metric evaluates how semantically similar the generated summary is to the original text, ensuring that the key meaning is preserved even if the wording differs.

- Human Evaluation: Human judgment remains an important metric, particularly for assessing aspects like relevance, coherence, and informativeness that automated metrics may miss. It provides insights into the subjective quality of the summary from an end-user perspective.

Combining all the above metrics can comprehensively evaluate summarization quality, balancing quantitative measures with subjective assessments.

V. CONCLUSION

Text interpretation is subjective, and summarization adds complexity. Studies comparing traditional textual analysis to

machine learning demonstrate potential, especially when integrating advanced visual techniques with conventional methods. According to metrics, and despite challenges like large datasets and model complexity, the future of deep learning-based text summarization is promising, driven by innovations such as Google's Gemini, which shows improvements in accuracy in its free version. However, when comparing paid versions, ChatGPT has demonstrated superior performance. Continued advancements in algorithms will enhance the adaptability and efficiency of these systems, impacting sectors like education, healthcare, finance, and media. Future research should focus on hybrid models that combine the strengths of various techniques and on incorporating real-time processing and deeper contextual understanding. Additionally, developing quantitative metrics will improve the precision, reliability, and reproducibility of evaluations, fostering better comparisons across studies and revealing patterns overlooked by interpretive analysis alone. This approach will contribute to a more comprehensive understanding of the field.

ACKNOWLEDGMENT

The authors would like to thank the Federal University of Technology – CP/PR for the financial support provided for this research.

REFERENCES

- [1] S. Arya, A. Bhaskar and K. Gupta, "Conversational AI: A Treatise About Vying Chatbots," 2024 2nd International Conference on Disruptive Technologies (ICDT), Greater Noida, India, 2024, pp. 929-934, doi: 10.1109/ICDT61202.2024.10489545.
- [2] A. Puri, G. Agrawal, A. Dukare, and M. Jawale, "A Survey and Analysis of Textual Content Based on Exploratory Data Analysis Technique and Opinion Analysis," in 4th International Conference on Computation, Automation and Knowledge Management (ICCAKM), Dubai, United Arab Emirates, 2023, pp. 1-6, doi: 10.1109/ICCAKM58659.2023.10449608.
- [3] A. Rangapur and A. Rangapur, "The Battle of LLMs: A Comparative Study in Conversational QA Tasks," arXiv preprint arXiv:2405.18344, 2024.
- [4] Stanford Encyclopedia of Philosophy, "Plato's Myths and the Mystery Tradition," Stanford Encyclopedia of Philosophy, Available at: <https://plato.stanford.edu/entries/plato-myths/>. Accessed on: May 12, 2024.
- [5] K. D. Garg, V. Khullar, and A. K. Agarwal, "Unsupervised Machine Learning Approach for Extractive Punjabi Text Summarization," 2021 8th International Conference on Signal Processing and Integrated Networks (SPIN), Noida, India, 2021, pp. 750-754.
- [6] S. Ren and Z. Zhang, "Pointer-Generator Abstractive Text Summarization Model with Part of Speech Features," IEEE 10th International Conference on Software Engineering and Service Science (ICSESS), Beijing, China, 2019, pp. 1-4.
- [7] S. R. Rahimi, A. T. Mozhdehi, and M. Abdolahi, "An overview on extractive text summarization," *Proc. IEEE 4th Int. Conf. Knowledge-Based Engineering and Innovation (KBEI)*, IEEE Press, Tehran, Iran, Dec. 2017, pp. 0054-0062, doi:10.1109/KBEI.2017.8324874.
- [8] A. M. Ahmed Zeyad and A. Biradar, "Advancements in the Efficacy of Flan-T5 for Abstractive Text Summarization: A Multi-Dataset Evaluation Using ROUGE and BERTScore," *Proc. 2024 Int. Conf. Advancements in Power, Communication and Intelligent Systems (APCI)*, IEEE Press, Kannur, India, 2024, pp. 1-5, doi:10.1109/APCI61480.2024.10616418.

Attraction-Based Reinforcement Learning: A Real-Time Approach Using Techniques Based on the Animal Behavior

Marcos A. M. Laia, Edimilson B. Santos, Wesley S. Guimarães

Department of Computer Science
Federal University of São João del Rei (UFSJ)
São João del Rei, Minas Gerais - Brazil
e-mail: marcoslaia@ufsj.edu.br,
edimilson.santos@ufsj.edu.br

Márcio Mendonça, Rodrigo H. C. Palácios, Janaína F. S. Gonçalves

PPG of Mechanical Engineering - CP/PG
Technological Federal University of Paraná (UTFPR)
Cornélio Procópio, Paraná - Brazil
e-mail: mendonca@utfpr.edu.br,
rodrigopalacios@utfpr.edu.br, janainaf@utfpr.edu.br

Abstract—Building a virtual world with simulated physical phenomena based on attraction and repulsion rules offers a unique opportunity to move agents according to these rules and plan actions mimicking real-world animal behavior. This paper presents an attraction-based algorithm using the Unscented Kalman Filter (UKF) to learn and predict opponent behavior in real time. The algorithm leverages attraction and repulsion forces to simulate physical interactions, facilitating robust predictions and learning accurately. Agents can optimize their strategies through reinforcement learning by adjusting attraction and repulsion parameters. Our results demonstrate the algorithm's effectiveness in dynamic environments, compared with traditional Q-learning methods, especially in low-frame conditions.

Keywords: *Unscented Kalman Filter; Animal Behavior; Machine Learning; Autonomous Systems.*

I. INTRODUCTION

Creating a virtual environment governed by attraction and repulsion principles enables the simulation of agent movements and decision-making processes that closely resemble real-world animal behavior. Animal movement is a dynamic spatio-temporal process where trajectory data reflect the instantaneous animal position in space and time, and other factors influence movement decisions between these observed positions [1]. This framework allows for accurate simulation of physical interactions and provides a robust foundation for improving predictions and facilitating learning.

By leveraging attraction and repulsion rules in the simulation, we can achieve coherent and realistic modeling of agent behavior. These rules can predict future states of the system by simulating natural interactions and dynamics observed in real-world scenarios. For instance, when an agent (e.g., a robot) moves towards a target (e.g., a ball), the attraction force guides its path, while repulsion forces from obstacles ensure collision avoidance. This combination can be fine-tuned to predict the agent's trajectory accurately. Each agent is assumed to move following a first-order Newtonian law, distinguishing speed and orientation, which

results from the balance of behavioral stimuli defined by direction and weight [2].

These rules can also be utilized for learning purposes. By simulating various scenarios and observing the outcomes, the system can iteratively adjust the parameters governing the attraction and repulsion forces to optimize agents' performance. This approach is particularly effective in reinforcement learning, where agents learn optimal strategies through trial and error within the simulated environment. Interactive multi-agent simulation algorithms compute the trajectories and behaviors of different entities in virtual reality scenarios. However, current methods involve considerable parameter tweaking to generate plausible behaviors [3].

Integrating learning algorithms, such as Artificial Neural Networks (ANN) or Kalman filters (KF), with these physical rules can enhance the system's predictive and adaptive capabilities. For example, a KF can estimate the state of the system (e.g., the positions and velocities of agents) and update predictions based on observed data. This iterative process refines the model's accuracy over time, accommodating both linear and nonlinear dynamics present in the simulation. The KF is commonly applied with ANN for chaotic systems identification [4].

Constructing a virtual world governed by attraction and repulsion rules not only allows for realistic simulation of physical phenomena but also provides a powerful tool for improving predictions and facilitating learning. By continuously refining the parameters and incorporating learning algorithms, the system can evolve to exhibit increasingly sophisticated and accurate behaviors akin to those observed in the natural world [5].

For concrete cases, consider Agre and Chapman's approach, which addresses the complexity, uncertainty, and immediacy of real-world situations. This approach was applied using agents to play Pengo [6]. Agre describes that Pengi, the implementation for playing Pengo, is part of a cognitive architecture theory derived from theories by Drescher, Minsky, and others. Mackworth proposed an architecture aimed at providing solutions for intelligent embedded systems using AI and robotics [7]. Sahota built on

the work of Agre, Chapman, and Mackworth, proposing a simulator with two robots and a ball [8].

Robots can cooperate to perform specific tasks, such as moving an object from one place to another [9]. For more complex tasks, studying the natural behavior of animals performing such tasks can be effective. Anderson and Donath question how a robot can automatically act across various tasks and environments, displaying diverse behaviors [10]. Observing animal behavior in nature for specific tasks without excluding others is beneficial.

In their work, Anderson and Donath identify behaviors of animals with evasion (repulsion) and attraction. Animals may avoid certain locations through repulsion behavior to prevent collisions with moving objects, evade predators, and avoid unsuitable environments. Repulsion can be passive (stopping to avoid collision, like freezing) or active (direct commands to avoid approaching objects). Attraction behavior is fundamental for a robot's movement toward a goal [9].

Using the basic idea of animal instinct, where something that captures attention exerts an attractive force, directs the predator's actions similar to the brain's command to grasp an object on a table. Rejection is also instinctual, where an unwanted object causes repulsion. This behavior can be expressed by attractive and repulsive potential fields [9].

Section II describes the KF and its variations used in this work, while Section III discusses Q-learning and its variation with ANN. Section IV presents the adopted methodology, Section V shows the obtained results, and finally, the conclusions are presented in Section VI.

II. KALMAN FILTER

The Kalman filter is a powerful tool for estimating the state of a dynamic system from noisy measurements. Designed for on-the-fly correction, it obtains precise measurements through sample observations [11]. In its basic form, it is an optimal linear estimator with constraints. The necessary linear functions are:

$$\begin{cases} x_k = Fx_{k-1} + q_k \\ z_k = Hx_k + r_k \end{cases} \quad (1)$$

Here, x_k is the current state, q_k is process noise (zero mean, covariance Q_k), and r_k is observation noise (zero mean, covariance R_k). F and H are transfer matrices. The state and noise distributions are:

$$\begin{cases} x_k \sim N(\bar{x}, P_k) \\ q_k \sim N(0, Q_k) \\ r_k \sim N(0, R_k) \end{cases} \quad (2)$$

The discrete KF involves prediction and update steps:

$$\begin{cases} \hat{x}_k^- = Fx_{k-1} \\ \hat{P}_k^- = F\hat{P}_{k-1}F^T + Q \\ K_k = \frac{\hat{P}_k^- C^T}{C\hat{P}_k^- C^T + R} \\ \hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-) \\ \hat{P}_k = (I - K_k H)\hat{P}_k^- \end{cases} \quad (3)$$

The EKF handles nonlinearities by approximating models with linear functions around the current state [12]:

$$\begin{cases} x_k = f(x_{k-1}) + q_k \\ z_k = h(x_k) + r_k \end{cases} \quad (4)$$

$$F_{k-1} = \frac{\partial f}{\partial x} \quad (5)$$

$$H_k = \frac{\partial h}{\partial x} \quad (6)$$

$$\begin{cases} \hat{x}_k^- = f(x_{k-1}) \\ \hat{P}_k^- = F_{k-1}P_{k-1}F_{k-1}^T + Q \\ K_k = \frac{\hat{P}_k^- H^T}{H\hat{P}_k^- H^T + R} \\ \hat{x}_k = \hat{x}_k^- + K_k[z_k - h(\hat{x}_k^-)] \\ \hat{P}_k = (I - K_k H_k)\hat{P}_k^- \end{cases} \quad (7)$$

The EKF can train neural networks by treating weights as states to be estimated. The nonlinear mapping g is parameterized by the weight vector W :

$$y_k = g(x_k, W) \quad (8)$$

The error is defined by:

$$e_k = d_k - g(x_k, W) \quad (9)$$

The state-space representation is:

$$\begin{cases} W_k = W_{k-1} + v_k \\ y_k = g(x_k, W_k) + e_k \end{cases} \quad (10)$$

The UKF improves estimation for highly nonlinear systems using the unscented transformation [13]. It represents the state distribution with sigma points:

$$\begin{cases} X_i = \bar{x} \\ X_i = \bar{x} + (\sqrt{(L+\lambda)P_x})_i, \text{ para } i = 1, \dots, L \\ X_i = \bar{x} - (\sqrt{(L+\lambda)P_x})_i, \text{ para } i = L+1, \dots, 2L \\ W_o^{(m)} = \frac{\lambda}{L+\lambda} \\ W_o^{(c)} = \frac{\lambda}{L+\lambda} + (1 - \alpha^2 + \beta) \\ W_i^{(m)} = W_i^{(c)} = 1/\{2(L+\lambda)\}, \text{ para } i = 1, \dots, 2L \end{cases} \quad (11)$$

where λ is a scalar parameter defined by the equation:

$$\lambda = \alpha^2(L + \kappa) - L \quad (12)$$

where L is the state vector dimensionality. α determines the spread of sigma points, κ influences the spread, and β incorporates prior distribution knowledge.

KFs are used in navigation, tracking, signal processing, and control systems. They provide robust state estimation and improve accuracy in predictions and measurements. The UKF's ability to handle nonlinearities without Jacobian computations makes it suitable for complex dynamic systems [14].

III. Q-LEARNING FOR REINFORCEMENT LEARNING

Reinforcement Learning (RL) is a computational approach where an agent learns to make decisions by performing actions and receiving feedback from the environment. One of the widely used algorithms in RL is Q-learning, introduced by Watkins and Dayan [15]. Q-learning is a model-free RL algorithm that seeks to learn the value of the optimal policy, which guides the agent's actions to maximize the cumulative reward.

Q-learning is an off-policy RL algorithm that learns the value of an action in a particular state without requiring a model of the environment. The core component of Q-learning is the Q-table, which stores the value (*Q-value*) of each state-action pair. The Q-value update rule is given by:

$$Q(s, a) \leftarrow Q(s, a) + \alpha[r + \gamma \max_{a'} Q(s', a') - Q(s, a)] \quad (13)$$

where $Q(s, a)$ is the Q-value for state s and action a . α is the learning rate ($0 < \alpha \leq 1$). r is the immediate reward received after performing action a in state s . γ is the discount factor ($0 \leq \gamma < 1$). s' is the next state after taking action a . $\gamma \max_{a'} Q(s', a')$ represents the maximum Q-value for the next state s' across all possible actions a' .

To approximate the Q-value function, we use a feedforward neural network. The network takes the current state as input and outputs the Q-values for all possible actions. The network is trained to minimize the Temporal Difference (TD) error:

$$\delta = r + \gamma \max_{a'} Q(s', a') - Q(s, a) \quad (14)$$

The ANN architecture consists of an input layer with a size equal to the state space dimension, one or more hidden layers. An output layer with a number of neurons equal to the number of actions.

IV. METHODOLOGY

A. Integrating the Atari Emulator for Reinforcement Learning Algorithms

The first step was to find an emulator that could provide the necessary structure to test the algorithms. We used the Atari emulator available within the OpenAI framework. OpenAI offers a comprehensive structure for reinforcement learning within the Python environment; however, there is a higher complexity associated with working with matrices and

equations in this environment. To simplify and facilitate the visualization of states and images, we used the Octave environment. To call functions and procedures from Python, we used the pythonic library. The emulator is invoked with a simple call specifying the name of the game, which is an emulated ROM. The game chosen for testing was Boxing, a boxing game between two players with an overhead view where the player must press a button to punch and move to dodge or hit the opponent's nose to score points. The players are confined to the ring area. The first player to score 100 points before the time runs out wins by knockout.

B. Understanding Emulator Inputs and Outputs

The second step was to study what can be sent to the emulator and what can be received from it. The emulator can accept button inputs from the controller to move the player, i.e., up, down, left, right, and an action button, which in this case is to punch the opponent. It accepts binary values such as 1 for pressed and 0 for released. By selecting the button values, they can be sent to the emulator via the `set_button_mask` function. Each iteration of the emulator can be performed using the `step()` function. This function runs one cycle of the game. Since a game runs at 60 frames per second, calling the `step()` function advances the game by one frame.

The emulator provides various outputs, such as memory, which can be utilized using the `get_state()` function. This function returns the current state of the game, allowing it to be saved and later restored using the `set_state(state)` function. In this study, we used the `get_screen()` function to capture the screen at the current game state. This screen capture was used to visualize the current state of the game.

C. Applying Computer Vision Techniques

The third step involved applying computer vision techniques to extract important data for the algorithms. The targets were four objects: the two player sprites and the two score sprites. The player sprites were divided into three points on a two-dimensional Cartesian plane, corresponding to the fighter's head and the two boxing gloves. For the scoring sprites, we used a change detector on the scoreboard to count hits. The hits can be worth one or two points, depending on the proximity of the strike. We used only the hit count as the parameter.

D. Implementation of the Proposed Algorithm

The fourth step was the implementation of the algorithm. For the first algorithm, which we propose in this work, we developed an on-the-fly learning system that observes the moves of the opponent. To achieve this, we mapped the points of the players. Each player has three points, and we created two additional points representing the position of the gloves before throwing a punch. This is a good way to determine if a fighter is approaching the head of the opponent, regardless of whether they have thrown a punch.

Using this visualization, we applied a Discrete Kalman Filter (DKF) with three states to estimate the acceleration of the points and thus predict future states, which are the future positions of the players.

$$a^-_k = a_k \quad (14)$$

$$v^-_k = v_k + a_k \quad (15)$$

$$p^-_k = p_k + v^-_k \quad (16)$$

$$\begin{bmatrix} p_k \\ v_k \\ a_k \end{bmatrix} = \begin{bmatrix} 1 & \delta t & 0 \\ 0 & 1 & \delta t \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_{k-1} \\ v_{k-1} \\ a_{k-1} \end{bmatrix} \quad (17)$$

where a is acceleration, v is velocity and p is position.

A second nonlinear filter was used to estimate the intention of the fighters. This algorithm utilizes the UKF to determine the value of the attraction coefficient k (repulsion if the value is negative) between the points of player 1 and player 2.

By observing the five points of each fighter, we obtain k values for each point on each axis. This is because, to land a precise hit, the fighter must align the glove with the opponent's head and then deliver the punch in a straight line. The algorithm was configured to run in real-time. The UKF can map the attraction coefficient function:

$$O'_a = Attraction(O_a, O_b, k) \quad (18)$$

where O_a is the object attracted by object O_b and k is the attraction coefficient. This is a global attraction defined at any point on the field. In this work, the objects are the points of each player. The attraction is process and observation function is given by:

$$\begin{cases} k_k = (k_{k-1}) + q_k \\ p'_1 = Attraction(p_1, p_2, k_k) + r_k \end{cases} \quad (19)$$

where k_k is the actual attraction coefficient and the p'_1 is the future observed state estimated by the DKF. p_1 e p_2 are the states for actual position.

E. Comparative Techniques

To compare the results, we used two other machine learning techniques commonly employed to control players. Typically, these techniques are executed in offline systems, where the player makes a move, and the reward is calculated. The player can avoid low or negative rewards by retracing steps and navigating through the best rewards. In our work, these algorithms had to be adapted for real-time use.

For example, the Q-learning algorithm was modified to remove the exploration rate and was fixed at 6 states, corresponding to one button being pressed at a time or no button pressed, representing a state of no action.

To test Q-learning with artificial neural networks in an on-the-fly system, we used the Extended Kalman Filter (EKF). The EKF was chosen because it requires less memory and processing power compared to the UKF. With many states,

the nonlinear filter becomes heavy and slow due to the computational cost of the sigma points.

V. RESULTS

To generate the results, we used the longest processing time of the algorithms to set a standard for real-time performance. The algorithm with the longest processing time was the EKF with artificial neural networks used to train Q-learning, followed by the UKF, and then Q-learning with tables.

Another factor to note is that the Pythonic library does not have an implementation to transfer a variable or matrix directly from Python to Octave. Transferring matrices proved to be very slow. Transferring an emulator image took about 1.3 seconds, involving a long and complex process. A much faster alternative was to save the image within Python using library calls and then read it in Octave. Using an SSD, the task took about 0.003 seconds.

To simulate a real system, we assumed that out of 60 frames per second (fps) (Figure 1), only 4 frames are received, meaning one frame is received every 15 frames (Figure 2), or 4 fps. For this type of task, not all systems are suitable for tracking, and many details are lost, limiting the results.

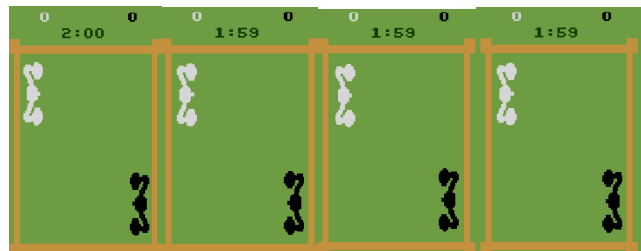


Figure 1. Sprite movement in 60 fps.

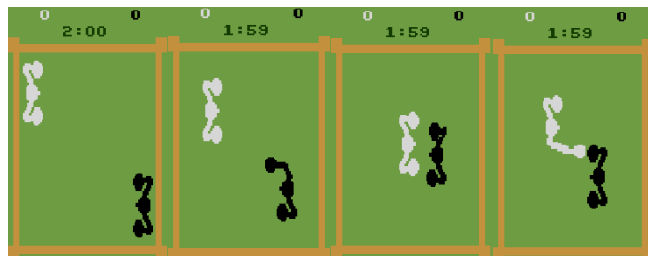


Figure 2. Sprites movement in 4 fps.

For the machine learning approach proposed in this work, we first used the DKF with the variance values $x=[0;0;0]$, $Q=10$, $R=1$ and $P=100$ with Δt set to 4. We then applied the value x twice using $y=F \cdot (F \cdot x)$ to obtain a prediction given the anterior states. This was done with the 5 points of each fighter. The DKF is used to provide a prediction of the movement in order to calculate the degree of attraction of the adversary to a position. Afterward, we applied the UKF with the values $x=[0;0]$, $Q=10$, $R=1$ and $P=100$.

The algorithm for Q-learning was tested in a simulated environment with six predefined states. The Q-table was initialized, and the agent's performance was evaluated over multiple episodes. The learning parameters were set as follows: learning rate $\alpha=0.1$, discount factor $\gamma=0.9$, and exploration rate $\epsilon=0$.

The Q-learning algorithm with EKF and ANN was tested using the following parameters: $Q=0.01$, $R=0.001$, and $P=1$, with pre-defined, randomly initialized weights. The neural network architecture consisted of Input Layer: 12 neurons representing the 3 points of each fighter, Hidden Layer: 13 neurons, Output Layer: 5 neurons for the buttons. The transfer functions used were sigmoid for the hidden layer and linear for the output layer. For the Q-learning prediction values, we used a learning rate α of 0.1 and the following prediction equation:

$$o = p + \alpha(r - p) \quad (20)$$

where o is the observed state for EKF, p is the predicted state (the buttons for controller), r is the reward. The agent successfully learned to navigate the environment and maximize cumulative rewards. The Q-values converged to stable values, indicating the agent had effectively learned the optimal policy.

The reward calculation in the Q-learning algorithm is critical for guiding the agent towards the optimal policy. In this work, the reward function is defined to balance the time spent in the game, the effectiveness of the player's punches, and the actions taken by the player. The reward function is given by:

$$r = -1t_g + hit_1hit_{1acc} - hit_2hit_{acc2} + 5 * B \quad (21)$$

$$hit_{1acc} = hit_{1acc} + 7 \quad (22)$$

$$hit_{2acc} = hit_{2acc} + 3 \quad (23)$$

where t_g is the time counter of the game, hit_1 and hit_2 are the hit detection for player 1 and 2, hit_{1acc} and hit_{2acc} are the accumulative hit (each time, the hit value increase) and B is equal a 1 is for players' movement.

To test the inputs and the capability of the emulator, an agent was created with three actions: Approach, Retreat, and Attack. The player approaches the opponent's head, and when within a certain distance from either the left glove or the right glove, the attack button is pressed. If the opponent's glove is close before the player makes an attack, the player retreats to the upper or lower diagonal, depending on which of the opponent's gloves is closer.

In TABLE I, the results of the interactions are presented. The proposed algorithm can generate both actions and possible new positions for other algorithms. In this work, the actions were combined with Q-learning while the positions were passed to the EKF as states for the input of the Neural Network.

With the programmed player, the game time functions well because decisions are made at 60 frames per second, providing ample time to make decisions. The implementation

was kept simple to test the use of buttons and the integration of the emulator into the Octave environment.

TABLE I. RESULTS OF THE ALGORITHMS IMPLEMENTED

Algorithms	Score	Remaining Time	Total Rewards
Programmed	KO x 74	34 s	45,422
UKF	79 x 78	0 s	27,670
Q-learning	KO x 97	3 s	33,361
EKF	24 x 58	0 s	-18,241
UKF + Q-learning	95 x 96	0 s	30,871
UKF + EKF	13 x 59	0 s	-42,262

By integrating these filtering techniques, we improved the prediction accuracy of the fighters' movements, which subsequently enhanced the performance of our reinforcement learning algorithm.

The implementation with the proposed algorithm of learning from intentions based on the attraction force between various points of the player proved to be a robust technique. This approach allowed the agent to learn intentions and use them to predict the opponent's attacks. Another positive aspect is that the agent did not have a predefined objective; it inferred that its objective was the same as the opponent's.

This capability highlights the algorithm's effectiveness in understanding and adapting to the game's dynamics, enhancing its predictive power and decision-making process in real-time scenarios.

The combination of the DKF and the UKF allowed us to effectively model the dynamic behavior of the fighters in real-time. The DKF provided a robust initial prediction, while the UKF refined these predictions by estimating the nonlinear effects, such as the attraction coefficient k between the fighters' points.

The defined reward function successfully guided the agent towards learning effective strategies in the game. By incorporating both the immediate rewards from actions and the long-term impact of hits landed by the players, the agent learned to balance between offensive and defensive strategies.

The use of EKF provided a robust method for state estimation, enhancing the performance of the Q-learning algorithm. The neural network effectively approximated the Q-values, allowing the agent to make informed decisions in real-time. The EKF algorithm exhibited issues with delayed positions, which theoretically should be better handled. Testing at 60 frames per second (fps) showed better results compared to other Q-learning techniques. However, when reduced to 4 fps, the EKF algorithm demonstrated weaknesses.

This indicates that while the EKF algorithm performs well under high frame rate conditions, its performance degrades with lower frame rates due to the increased latency in position updates. This latency impacts the algorithm's

ability to accurately predict and respond to the dynamic changes in the environment.

The limitation of using Q-learning with 6 states resulted in less natural movement of the player. However, with fewer directions, the algorithm performed better compared to using 9 states (8 directions plus attack). The issue with using Q-learning is that the reward must be known beforehand, and a function must be produced to achieve the optimal value.

By incorporating movement into the reward function, the player is discouraged from standing still and continuously punching. The idea of earning more points through successful attacks encourages the player to be more aggressive and proactive rather than merely avoiding the opponent.

The higher reward for attacking actions motivates the player to engage with the opponent actively. By rewarding movement, the player is incentivized to maneuver strategically rather than remain stationary. The penalty for being hit encourages the player to avoid attacks while planning their own.

VI. CONCLUSIONS

Both implementations have their strengths and limitations. The EKF with neural networks showed potential in understanding and predicting dynamic game scenarios, while Q-learning with fewer states proved to be more efficient in specific conditions. Future work should focus on improving the robustness of these algorithms, particularly in handling lower frame rates and refining the reward structures for better performance.

The reward function was integral in shaping the player's behavior, promoting a balance between offensive and defensive strategies while maintaining an active and engaging playstyle. The careful design of the reward structure ensured that the player optimized both movement and attack to achieve the best results.

This work presents an attraction-based algorithm as a more intuitive solution utilizing the UKF with attraction functions. The initial idea behind this algorithm was to learn the behavior of opponents to predict future positions and actions. However, it proved capable of learning in real-time within a limited sampling state, which would be impractical for humans.

The algorithm's ability to learn and predict in real-time, even with limited sampling data, demonstrates its potential for practical applications in dynamic environments. By modeling attraction forces between key points of the players, the algorithm effectively learns and anticipates opponent behavior, enabling strategic decision-making.

In future work, we plan to extend the application of this algorithm by analyzing video footage, applying the algorithm to analyze videos of players to learn their tactics. This would involve extracting key movement patterns and strategies from recorded gameplay. Testing with different players evaluating the learned strategies against both the implemented player and other players to assess the robustness and adaptability of the algorithm, enhancing the model

improving the attraction model to incorporate more complex behaviors and interactions, potentially including environmental factors and varying opponent skill levels.

The proposed algorithm has potential applications beyond gaming, such as in sports analytics, where understanding and predicting player movements can provide significant strategic advantages. Additionally, it can be applied in robotics for real-time path planning and obstacle avoidance by learning dynamic environments.

REFERENCES

- [1] R. W. Loraamm, "Incorporating Behavior Into Animal Movement Modeling: A Constrained Agent-Based Model For Estimating Visit Probabilities In Space-Time Prisms", *International Journal Of Geographical*, 2019.
- [2] S. Bernardi and M. Scianna, "An agent-based approach for modelling collective dynamics in animal groups distinguishing individual speed and orientation". *Phil. Trans. R. Soc. B 375*: doi.org/10.1098/rstb.2019.0383, 2020.
- [3] J. Ren, W. Xiang, Y. Xiao, R. Yang, D. Manocha and X. Jin, "Heter-Sim: Heterogeneous Multi-Agent Systems Simulation by Interactive Data-Driven Optimization". *IEEE Transactions on Visualization and Computer Graphics*. doi:10.1109/tvcg.2019.2946769, 2019.
- [4] R. J. de Jesus, "Stable Kalman filter and neural network for the chaotic systems identification", *Journal of the Franklin Institute*, doi: 0.1016/j.jfranklin.2017.08.038, 2017.
- [5] Y. Tanaka, "Machine Learning That Reproduces Physical Phenomena from Data", *NTT Technical Review*, vol. 21, no. 10, pp. 15-19, Oct., 2023.
- [6] P. Agre and D. Chapman, "Pengi: An implementation of a theory of activity". In *Proceedings of AAAI-87*, pp. 196-201, 1987.
- [7] A. K. Mackworth, "On seeing robots", in: *Computer Vision: Systems, Theory, and Applications*. Singapore World Scientific, pp. 1-13, 1993.
- [8] M. K. Sahota and A. K. Mackworth, "Can situated robots play soccer?", in *Artificial Intelligence 94*, pp. 249-254. Banff, Alberta, 1994.
- [9] J. T. Weigel, S. Gutmann, M. Dietl, A. Kleiner, and B. Nebel, "CS Freiburg: Coordinating Robots for Successful Soccer Playing". *IEEE Transactions on Robotics and Automation*, vol. 18, no. 5, pp. 685-699, 2002.
- [10] T. L. Anderson and M. Donath, "Animal behavior as a paradigm for developing robot autonomy". *Robotics and Autonomous Systems*, vol. 6(1-2), pp. 145-168. doi:10.1016/s0921-8890(05)80033-8, 1990.
- [11] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35-45, 1960.
- [12] S. J. Julier and J. K. Uhlmann, "A new extension of the Kalman filter to nonlinear systems", *Proceedings of AeroSense: The 11th International Symposium on Aerospace/Defense Sensing, Simulation and Controls*, 1997.
- [13] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation", *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401-422, 2004.
- [14] R. van der Merwe and E. Wan, "Sigma-point Kalman filters for probabilistic inference in dynamic state-space models", *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2001.
- [15] C. J. C. H. Watkins and P. Dayan, Q-learning. *Mach Learn* 8, pp. 279-292. Doi: 10.1007/BF00992698, 1992.

Using Fuzzy Cognitive Maps and Chatbots to Evaluate Student Satisfaction in a University: A Comparison between Strong and Weak AI

Wagner F. Godoy, José A. Fabri, Rodrigo H.C. Palácios

Academic Department of Electrical Engineering
Technological Federal University of Paraná (UTFPR)
Cornélio Procópio, Paraná – Brazil
e-mail: wagner.godoy@utfpr.edu.br, fabri@utfpr.edu.br,
rodrigopalacios@utfpr.edu.br

Márcio Mendonça, Janaína F. S. Gonçalves, Luiz O. M. Moraes

PPG of Mechanical Engineering - CP/PG
Technological Federal University of Paraná (UTFPR)
Cornélio Procópio, Paraná – Brazil
e-mail: mendonca@utfpr.edu.br, janainaf@utfpr.edu.br,
luizotaviomoraes@alunos.utfpr.edu.br

Abstract—The project aims to create a Fuzzy Cognitive Map (FCM) to assess student satisfaction at the Federal Technological University of Paraná, Campus Cornélio Procópio (UTFPR-CP), Brazil. By integrating intelligent techniques, the FCM allows for inference through the causal relationships between variables like teacher training, facilities, housing, and cleanliness. Online surveys were used to quantify these influences and provide an institutional improvement tool. A comparison was made with a Microsoft AI chatbot, which produced similar results, although GPT-4.0 offers more nuanced, human-like understanding and reasoning capabilities. After validating the results, this work ends with a conclusion and addresses future scientific investigations.

Keywords- *Fuzzy Cognitive Maps, Quantitative Analysis, Chatbots, Operational Research.*

I. INTRODUCTION

The text discusses various studies related to education and Fuzzy Cognitive Maps (FCMs). One study at Istanbul Technical University used FCMs to model critical success factors in control engineering, incorporating concepts like student and teacher profiles to assess program success. Another work applied Fuzzy Logic to identify trends in engineering education, focusing on achieving a "Global Dimension" for professional evaluation [9] [23].

One study explored student satisfaction and critical factors for improvement within universities [12]. At the same time, another used the Fuzzy Analytical Hierarchy Process (FAHP) to assess education quality at Nanyang Institute of Technology through four dimensions: government, university, society, and student. The FAHP framework guided quality evaluations by breaking these categories into sub-areas such as funding, curriculum, and student achievements [11][19].

Further research in Central New York examined K-12 school districts' transparency in communicating educational technology information, concluding that reporting needed to be more consistent with best practices in the literature [3]. Knowledge management approaches were also explored to enhance data understanding through cognitive processes like description and interpretation, using formal grammar for linguistic data description [2].

One study mapped educational contributions to human capital in China through Fuzzy Neural Networks (FNN), estimating contributions from various educational levels across different regions. This research emphasized the link between education and technological development [13].

Recent advancements in Artificial Intelligence (AI), such as ChatGPT 4, frequently regarded as a strong AI due to its substantial progress toward Artificial General Intelligence (AGI), more details in the circumscribed section), present new opportunities for educational analysis. Compared to FCMs and Fuzzy Logic-based systems, ChatGPT 4 demonstrates superior capabilities in nuanced comprehension, reasoning, and large-scale data processing, improving identifying patterns, critical success factors, and satisfaction metrics within complex educational systems.

Lastly, sustainability and innovation in Chinese universities were examined, and strategies for integrating sustainability into education were proposed through targeted faculty training, curriculum reform, and practical learning platforms.

The rest of the paper is structured as follows. In Section II, we present the background of fuzzy cognitive maps. Section III presents the development and outcomes of the proposed FCM-based tool. Section IV presents the results and discussions. We conclude in Section V.

II. FUZZY COGNITIVE MAPS: BACKGROUND

Since the pioneering work of Kosko [16], which extended Axelrod's cognitive maps [8] to include Fuzzy Logic, various applications of FCMs have been reported in the literature across different fields of knowledge. Of note are applications in artificial life [14][15], social systems [19], decision-making in the medical field [7], and the works of process control [22], which apply evolutions of FCM in autonomous mobile swarm robotics [4].

The work in [17] focuses on Supply Chain Quality Management (SCQM) modeling, a key research area in both academic and practical contexts. This paper introduces a Multi-Layer Fuzzy Cognitive Map (ML-FCM) model to enhance SCQM by structuring concepts into multi-layered Fuzzy Cognitive Maps (FCMs). Applied to a three-echelon supply chain in the personal care products sector, the model's construction and effectiveness were evaluated using density and strength metrics. The inference process results confirmed

the initial concept selection in each FCM and quantified the impact of each echelon on SCQM performance. Key decision-making variables identified for SCQM improvement included total profit, quality costs, supplier rejections and returns, productivity, and customer rejections and returns.

In [24], a formal adaptation of the original FCM is presented. This new tool is designated as Timed Automata Fuzzy Cognitive Maps (TAFCM), among others. It is observed that these works also present semantic variations of the original proposal, adapting the FCM structure according to the problem being addressed.

III. DEVELOPMENT AND OUTCOMES OF THE PROPOSED FCM-BASED TOOL

Fuzzy Cognitive Maps (FCMs) can be considered a mathematical model of a person or group's "belief structure", allowing the inference or prediction of the consequences this organization of ideas causes in the represented universe. This mathematical model was adapted to include uncertainties through Fuzzy Logic by [14], creating Fuzzy Cognitive Maps. Like the original Cognitive Maps, FCMs are directional graphs (digraphs) where numerical values are variables or Fuzzy sets. The "nodes" of these graphs are linguistic concepts, represented by Fuzzy sets, and each "node" is connected to others through links. Each of these links is associated with a numerical weight, which represents a Fuzzy variable related to the level of causality between the concepts.

$$A_i = \int \left(\sum_{j=1}^n (A_j \times W_{ji}) \right) \tag{1}$$

$$F(x) = \frac{1}{1 + e^{-\lambda x}} \tag{2}$$

To address the problem and obtain quantitative results utilizing expert knowledge, a Fuzzy Cognitive Map (FCM) was developed to assess student satisfaction at the Cornélio Procópio campus of the Federal University of Technology - Paraná. Pursuing continuous quality improvement necessitates using both qualitative and quantitative evaluation tools, with Fuzzy Cognitive Maps representing a suitable approach to this challenge [23].

Fuzzy Cognitive Maps, whether Fuzzy or classical, serve as a modeling tool to represent experts' collective or individual knowledge. It is achieved by assigning values to specific concepts and weights to their causal relationships. The explicit knowledge within these maps is conveyed through the values attached to the concepts and their relationships, while implicit knowledge is embedded in the overall structure of the map. One of the primary difficulties in constructing conceptual maps is converting knowledge into accurate models of human behavior. This challenge is

particularly evident in defining the semantics of the concepts and the precise nature of their interrelationships [20]. Beyond this, careful data acquisition is essential to ensure that raw data is transformed into consistent and meaningful information for the cognitive map.

In general, FCMs integrate aspects of Artificial Neural Networks [5], Fuzzy Logic [6], and Semantic Networks [20], as well as other intelligent computational techniques. Conceptually, an FCM can be understood as a methodology for modeling complex decision-making systems based on causal knowledge derived from the combination of Fuzzy Logic and Artificial Neural Networks [14]. Each concept within an FCM represents a distinct entity, state, variable, or characteristic of the modeled system. This framework provides a structured way to describe and predict the behavior of complex, often poorly understood, systems [21].

A classic example of a Fuzzy Cognitive Map (FCM) is represented by a graph structure with classical weights, as described in [4]. In some cases, however, FCMs may not achieve stability, leading to oscillatory or even chaotic behavior [22]. For systems that behave predictably, it is typically observed that after a finite number of iterations, the values of the concepts converge to either a fixed equilibrium point or a limit cycle (this work). This stabilization process can be seen in the evolution of the FCM depicted in Figure 2, where, after approximately four iterations, the values of the modeled concepts reach a steady state.

In the specific case of control, such as valve actuation used in the work of Mendonça and collaborators [22], the technique of cognitive maps is widely used for structuring complex problems through iterations, as shown in equation (2), until they typically stabilize at a fixed point or a limit cycle [4].

TABLE I. FCM CONCEPTS

- C1: Student Satisfaction Level
- C2: Professor's Performance
- C3: Pedagogic Structure – i.e. Library
- C4: Habitation
- C5: University Restaurant Quality
- C6: Cleaning - Accessibility
- C7: Entertainment and Sports Activities
- C8: Access to the internship

The concepts listed in Table I represent all the variables determining student satisfaction. It is noted that in this study, parking and municipal transportation for faculty/students (with distinct parking facilities) are also considered.

Figure 1 represents a Fuzzy Cognitive Map (FCM), a directed graph where nodes (concepts) are connected by weighted edges, symbolizing the influence between concepts. In this FCM: Nodes (C1 to C8) represent different concepts or variables within the modeled system. For instance, C1 is a central concept connected to several other nodes like C2, C3, C6, and C8, suggesting it plays a significant role in the system's dynamics. Edges between the nodes indicate causal relationships between the concepts. The weights on these edges (represented by numbers along the connecting lines) denote the degree and type of

influence: Positive weights (e.g., 0.80, 0.85) indicate a positive influence of one concept on another. Negative weights (e.g., -0.40, -0.20) suggest a negative influence, where an increase in one concept decreases the other. Directional arrows show the flow of influence from one concept to another. For example, C1 influences C3 with a weight of 0.80, and C2 influences C1 with a weight of 0.46, among others. The color-coded connections likely indicate different types or strengths of relationships between concepts, further detailing the complexity of the system being modeled. Overall, this FCM graph visually represents how each concept is interrelated, with specific concepts like C1 and C2 appearing to hold more central roles based on their numerous connections.

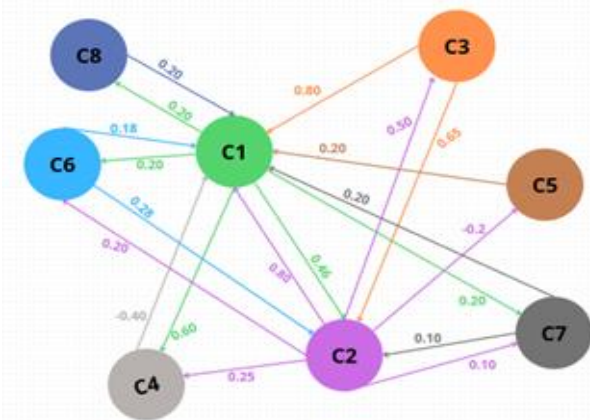


Figure 1. Student Satisfaction Level FCM.

VI. RESULTS AND DISCUSSIONS (FCM, ChatGPT)

A. Fuzzy Cognitive Maps

After modelling the concepts and causal relationships, particularly the numerical values assigned by the students, the model is calculated using Equations 2 and 3. This process allows the system to evolve toward the final values representing the study's objective: a quantitative evaluation of the central theme [10].

Initially, the concept values are set to zero. As shown in Figure 2, the evolution of each modelled concept is represented on a scale from zero to 10, with particular emphasis on the central concept—student satisfaction at UTFPR-CP. The FCM demonstrated stable behaviour, reaching a near-equilibrium point with only minor variations around the result, see Figure 2.

However, specific critical issues emerged during the research and model construction, such as the parking conditions near the university. This challenge is linked to the institution's rapid growth over the past decade and should be addressed by the university administration and government authorities to ensure future improvements [1].

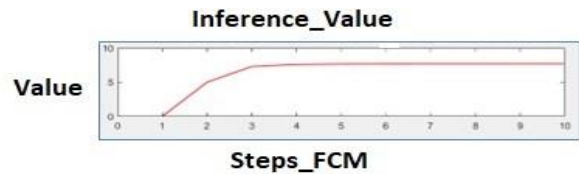


Figure 2. Student Satisfaction Level FCM.

Figure 2 presents the evolution of the Inference Value over successive FCM Steps in a Fuzzy Cognitive Map (FCM) model. The horizontal axis represents the number of Steps FCM taken, while the vertical axis indicates the corresponding Value of the inference at each step.

Key observations from the figure include:

a) *Initial Growth*: At the beginning (Steps 0 to 2), the inference value shows a steep increase, indicating that the system is rapidly adjusting as it begins to process the interactions between the concepts modeled in the FCM.

b) *Stabilization*: After Step 3, the inference value begins to plateau, suggesting that the model is approaching equilibrium. The value stabilizes around 0.8, indicating that the FCM has reached a steady state or a limit cycle where further iterations result in minimal to no significant changes in the concept values.

c) *Minor Variations*: The slight upward adjustment after Step 2 suggests minor refinements in the FCM behavior, but overall, the system stabilizes quickly. This stable convergence implies that the FCM is well-behaved, with the concepts interacting in a way that leads to consistent and predictable outcomes after a few iterations [20]. The final inference value provides a quantitative representation of the central theme being evaluated, likely reflecting a key metric like student satisfaction, as discussed in the context of the study

B. ChatBot 4.0

In 2022, OpenAI introduced a new model of ChatGPT, an NLP system capable of maintaining contextual natural conversations. Since its release, ChatGPT has spurred extensive research into its societal impacts, particularly in education. This article investigates ChatGPT's influence on university education by analyzing over 50 essay papers across various topics, focusing on its use in exams, programming tasks, and term papers. The study assessed misuse levels, including copying ChatGPT's outputs, and examined the effect on student performance, especially given the occasional inaccuracies in its responses. While concerns about academic dishonesty are highlighted, the article also emphasizes the tool's potential benefits, such as aiding learning and improving assignment comprehension. The findings suggest that ChatGPT can assist in generating coherent, informative work but stress the need for new assessment formats that emphasize critical thinking and creativity, areas where AI cannot fully substitute human effort [26].

In short, GPT-4 is a "weak AI" specialized in specific tasks, such as language processing, and it does not possess consciousness, proper understanding, or the ability for independent thought. While it is impressive in language tasks and has shown significant advancements in interaction and content generation, it operates based on patterns and statistics from the texts on which it was trained. It does not have a fundamental understanding of the world or itself.

The GPT-4 version, like other similar AIs, is still not considered a "strong AI" (or AGI - Artificial General Intelligence) [25]. The concept of "strong AI" refers to an artificial intelligence capable of performing any cognitive task a human could, including deep understanding, reasoning, creative adaptation to new problems, and autonomy.

The current version represents a significant advance in artificial intelligence and is closer to the so-called "Strong AI" or Artificial General Intelligence (AGI) than previous versions.

For now, "strong AI" remains within research and theory, and it is uncertain when (or if) this technology will be developed

Question: What is the level of satisfaction of a student with generally good and high-level professors, good university infrastructure in the laboratories, and most students being able to walk to the university? The university and the city offer a reasonable level of entertainment, but the university town is small and there are difficulties in securing internships

The level of satisfaction of a student in such a scenario can be influenced by several factors. Let us break down the specifics to get a clearer picture:

a) Professors and Academics:

High-Level Professors: Generally, students appreciate knowledgeable and experienced professors, which can lead to high satisfaction in terms of the quality of education and academic guidance

Academic Reputation: Good professors can enhance the university's reputation, positively impacting student satisfaction.

b) University Infrastructure:

Good Laboratories: High-quality laboratory facilities are crucial for hands-on learning and research, especially in science and engineering disciplines. This likely increases student satisfaction

Overall Infrastructure: Good infrastructure beyond labs, such as libraries, study spaces, and classrooms, also contributes to a positive learning environment.

c) Proximity to University:

Walking Distance: Most students being able to walk to the university is highly convenient and can significantly improve satisfaction due to the ease of access and saving on transportation costs.

d) Social and Entertainment Aspects:

Reasonable Entertainment Options: While a reasonable level of entertainment is available, the small size of the town might limit diverse social activities, which could slightly

lower satisfaction for students seeking more vibrant social lives.

University Town Size: A smaller town might offer a close-knit community feeling, which some students prefer, while others might find it lacking in variety and excitement.

e) Internship Opportunities:

Difficulty Securing Internships: Challenges in finding internships can be a significant drawback, as internships are essential for gaining practical experience, networking, and improving employability. This factor might lower overall satisfaction

Overall, the student satisfaction level in this scenario would likely be high in terms of academic and infrastructure aspects but might be tempered by the limited entertainment options and difficulties in securing internships. Here's a summarized view:

f) Positive Factors: Highly qualified professors

- Good university infrastructure, particularly well-equipped laboratories

- Proximity to the university for the majority of students

- A reasonable level of entertainment and activities

Negative Factors:

- The university is located in a small town with limited entertainment options

- Difficulties in securing internships, which underscores their importance in the overall student experience

Different weights were assigned to each aspect based on their perceived significance to quantify the impact of these factors. The proposed weighting scheme is as follows:

High-Level Professors: 20%

Good University Infrastructure (Laboratories): 20%

Good University Infrastructure (Overall): 15%

Proximity to University: 15%

Social and Entertainment Aspects (Entertainment): 10%

Social and Entertainment Aspects (Town Size): 10%

Internship Opportunities: 10%

The final assessment is calculated using Equation 3, which integrates the weighted values of each factor into the overall evaluation:

$$\text{Overall} = 1.8 + 1.8 + 1.2 + 1.2 + 0.6 + 0.5 + 0.4 = 7.5 \quad (3)$$

This score provides a comprehensive view of student satisfaction, considering both academic and social factors

g) Interpretation:

An overall satisfaction score of 7.5 out of 10 suggests that students are generally satisfied with their experience at the university. While academic and infrastructural aspects are highly appreciated, the limited entertainment options and challenges in securing internships reduce overall satisfaction.

A similar question applied to the accessible version of ChatGPT 3.5 (free version); the answer was close to 7.4 but with fewer details. That said, the answer from the free version was not included, and its result was to help validate the correlation between FCM and ChatGPT.

IV. CONCLUSION AND FUTURE WORK

The study concluded that student satisfaction at UTFPR-CP reached an overall score of 7.5/10, reflecting strong satisfaction with faculty quality and general infrastructure, while identifying areas for improvement in laboratory and library facilities. Despite recent course expansions and ongoing infrastructure upgrades, gaps remain in resources. The use of Fuzzy Cognitive Maps (FCM) successfully quantified qualitative data, and results were further validated by ChatGPT 3.5 and 4.0, both yielding consistent outcomes. Given the alignment between FCM and AI-based tools, the study is recommended as a continuous assessment tool for guiding university improvements.

Future work should increase sample size, incorporate diverse metrics, and employ longitudinal studies for tracking satisfaction trends, while qualitative methods could provide deeper insights into student experiences.

ACKNOWLEDGMENT

The authors would like to thank the Federal University of Technology – CP/PR for the financial support provided for this research.

REFERENCES

- [1] Y. Gao, J. Yang, and F. Liu, "The Research on the Comprehensive Assessment of Higher Engineering Education Based on Fuzzy Comprehensive Evaluation Method," in *Future Information Technology and Management Engineering, 2009. FITME '09. Second International Conference on*, pp. 124-127, 13-14 Dec. 2009.
- [2] L. Ogiela and M. R. Ogiela, "Knowledge Management Approaches for Data Understanding," in *2016 International Conference on Intelligent Networking and Collaborative Systems (INCoS)*, Ostrava, Czech Republic, 2016, pp. 270-273.
- [3] J. MacLeod and H. H. Yang, "Life-Cycle Efficacy for Educational Technology: Best-Practices for Leading Schools," in *2016 International Conference on Educational Innovation through Technology (EITT)*, Tainan, Taiwan, pp. 139-142, 2016.
- [4] M. Mendonça, H. S. Kondo, L. Botoni de Souza, R. H. C. Palácios, and J. Paulo Lima Silva de Almeida, "Semi-Unknown Environments Exploration Inspired by Swarm Robotics using Fuzzy Cognitive Maps," in *2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, New Orleans, LA, USA, pp. 1-8, 2019.
- [5] S. Haykin, *Neural Networks and Learning Machines*, 3rd ed. Upper Saddle River, NJ: Pearson Education, 2009.
- [6] W. Pedrycz and F. Gomide, *An Introduction to Fuzzy Sets: Analysis and Design*, MIT Press, 1998.
- [7] G. A. Krimpas, N. A. Krimpas, and P. P. Groumpos, "An Overview of Making Decisions on Medical Problems Using Fuzzy Cognitive Maps," in *2023 14th International Conference on Information, Intelligence, Systems & Applications (IISA)*, Volos, Greece, pp. 1-6, 2023.
- [8] R. Axelrod, *Structure of Decision: The Cognitive Maps of Political Elites*. Princeton, NJ: Princeton Univ. Press, 1976.
- [9] L. A. Franco and G. Montibeller, "Facilitated modeling in operational research," *European Journal of Operational Research*, vol. 205, no. 3, pp. 489-500, 2010.
- [10] E. Yesil, C. Ozturk, M. F. Dodurka, and A. Sahin, "Control engineering education critical success factors modeling via Fuzzy Cognitive Maps," in *Information Technology Based Higher Education and Training (ITHET), 2013 International Conference on*, pp. 1-8, 10-12 Oct. 2013.
- [11] Y. Hezhan, "Application of Fuzzy Analytical Hierarchy Process in innovation education quality evaluation of higher education institution," in *Information Management, Innovation Management and Industrial Engineering (ICIII), 2012 International Conference on*, vol. 1, pp. 438-441, Oct. 20-21, 2012.
- [12] S. Yu and K. Zhu, "Fuzzy Neural Network Applications on Estimating the Contribution of Different Education Levels on Human Capital of China," in *Service Systems and Service Management, 2007 International Conference on*, pp. 1-4, June 9-11, 2007.
- [13] P. Xiang, "Research on the Emotion Recognition Based on the Fuzzy Neural Network in the Intelligence Education System," in *Digital Manufacturing and Automation (ICDMA), 2011 Second International Conference on*, pp. 1030-1033, Aug. 5-7, 2011.
- [14] J. A. Dickerson and B. Kosko, "Virtual worlds as Fuzzy cognitive maps," *Presence*, vol. 3, no. 2, pp. 173-189, 1994.
- [15] J. A. Dickerson and B. Kosko, "Virtual worlds as Fuzzy dynamical systems," in *Technology for Multimedia*, B. Sheu, Ed. New York: IEEE Press, 1996.
- [16] B. Kosko, "Fuzzy cognitive maps," *International Journal of Man-Machine Studies*, vol. 24, no. 1, pp. 65-75, 1986.
- [17] J. Cogollo and A. Correa, "Modeling Supply Chain Quality Management using Multi-Layer Fuzzy Cognitive Maps," in *2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, New Orleans, LA, USA, pp. 1-6, 2019.
- [18] J. MacLeod and H. H. Yang, "Life-Cycle Efficacy for Educational Technology: Best-Practices for Leading Schools," in *2016 International Conference on Educational Innovation through Technology (EITT)*, Tainan, Taiwan, pp. 139-142, 2016.
- [19] C. Neocleous, M. Papaioannou, and C. Schizas, "Important issues to be considered in developing fuzzy cognitive maps," in *2011 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2011)*, Taipei, Taiwan, 2011, pp. 662-665.
- [20] E. Papageorgiou, "Learning Algorithms for Fuzzy Cognitive Maps," *IEEE Transactions on Systems and Cybernetics, Part C: Applications and Reviews*, vol. 42, pp. 150-163, 2012.
- [21] E. Papageorgiou, C. Stylios, and P. Groumpos, "Novel for supporting medical decision making of different data types based on Fuzzy Cognitive Map Framework," in *Proceedings of the 29th Annual International Conference of the IEEE EMBS*, Lyon, France, pp. 23-26, Aug. 2007.
- [22] M. Mendonça, B. Angélico, L. V. R. Arruda, and F. Neves, "A dynamic fuzzy cognitive map applied to chemical process supervision," *Engineering Applications of Artificial Intelligence*, vol. 26, pp. 1199-1210, 2013.
- [23] L. B. de Souza, R. V. P. D. Barros, P. P. Soares, J. G. Ferreira, and M. Mendonça, "Fuzzy Cognitive Maps and Weighted Classic Fuzzy Applied on Student Satisfaction Level," in *ICAS 2017*, pp. 60, 2017.
- [24] G. Acampora and V. Loia, "On the Temporal Granularity in Fuzzy Cognitive Maps," *IEEE Transactions on Fuzzy Systems*, vol. 19, no. 6, 2011.
- [25] A. F. Mohammad, B. Clark, R. Agarwal, and S. Summers, "LLM/GPT Generative AI and Artificial General Intelligence (AGI): The Next Frontier," in *2023 Congress in Computer Science, Computer Engineering, & Applied Computing (CSCE)*, Las Vegas, NV, USA, 2023, pp. 413-417.
- [26] A. F. Mohammad, B. Clark, R. Agarwal, and S. Summers, "LLM/GPT Generative AI and Artificial General Intelligence (AGI): The Next Frontier," *2023 Congress in Computer Science, Computer Engineering & Applied Computing (CSCE)*, Las Vegas, NV, USA, 2023, pp. 413-417.

Leveraging Large Language Models for Enhanced Personalised User Experience in Smart Homes

Jordan Rey-Jouanchicot^{*†‡}, André Bottaro^{*}, Eric Campo[‡], Jean-Léon Bouraoui^{*},
Nadine Vigouroux[†], Frédéric Vella[†]

^{*}Orange Innovation, Blagnac, 31700, France

[†]IRIT, University of Toulouse, CNRS, Toulouse INP, UT3, UT2J, Toulouse, France

[‡]LAAS-CNRS, University of Toulouse, CNRS, UT2J, Toulouse, France

e-mail: { jordan.reyjouanchicot@orange.com | andre.bottaro@orange.com | eric.campo@laas.fr |
| jeanleon.bouraoui@orange.com | nadine.vigouroux@irit.fr | frederic.vella@irit.fr }

Abstract—Smart home automation systems aim to improve the comfort and convenience of users in their living environment. However, adapting automation to user needs remains a challenge. Indeed, many systems rely on hand-crafted routines. This paper presents an original smart home architecture leveraging Large Language Models (LLMs) and user preferences to push the boundaries of personalisation and intuitiveness in the home environment. This article explores a human-centred approach that uses the general knowledge provided by LLMs to learn and facilitate interactions with the environment. The advantages of the proposed model are demonstrated on a set of scenarios, as well as a comparative analysis with various LLM implementations. Some metrics are assessed to determine the system’s ability to maintain comfort, safety, and user preferences. The paper details the approach to real-world implementation and evaluation. The proposed approach shows up to 52.3% increase in average grade, and with an average processing time reduced by 35.6% on Starling 7B Alpha LLM. In addition, performance is 26.4% better than the results of the larger models without preferences, with almost 20 times faster processing time.

Keywords—Artificial Intelligence; Decision-making; Adaptivity; Smart Home Automation System; Modelisation

I. INTRODUCTION

Networks of devices are deployed to assist human beings in their daily activities, using notions of context and knowledge to decide on the best actions to take [1][2]. Indeed, many houses are covered by wireless and wired networks and equipped with electronic devices allowing the occupants to control their environment for comfort, entertainment, security, energy management, and elderly care.

However, Smart Home Automation Systems are still missing the aim of autonomously taking the best action in every situation. Aligning automation routines to meet every need for every home configuration, every set of devices, available or not, functional or not, remains a challenge. In fact, most systems today are configured for simple routines that occur frequently. For instance, setting the home for wake-up or departure time by playing music, ringing a bell, acting on lights, shutters, heating, ventilation, and air conditioning. More precise needs in less frequent situations are not covered. Furthermore, if some devices are missing from a routine, no decision is made to use alternative devices. While artificial intelligence can play a role in learning about situations and

the associated actions taken by users [3], it still requires time to understand users’ habits, and is never able to cover the wide range of situations encountered at home.

One of the main technical challenges of ubiquitous computing is the ability to set up a system knowing the wide variety of users’ potential needs, and how it can adapt to the wide range of functions of existing devices, with devices and locations whose configuration can be very different. As their execution environment is particularly dynamic, applications need to be aware of their context and act appropriately.

Large language models [4] have the potential to give this general knowledge at once to applications, such as smart home. In essence, these models have acquired a vast amount of knowledge. They are trained on a diverse range of textual sources, enabling them to cover a variety of topics, facts, and concepts, here the expected actions of home devices to meet user needs in a wide range of situations.

Retrieval-Augmented Generation (RAG) [5] is a technique that improves the accuracy of LLMs (Large Language Models), by giving them access to more targeted and precise information. This is achieved by using a retrieval component that searches a specific database and introduces it into the model, along with users’ query. To do this, an embedding model is needed to represent the sentences in vector form, so that we can find the results closest to the query.

This paper proposes a software architecture integrating the general LLM knowledge available today into a smart home automation system. The LLM is placed at the centre of the home’s decision-making system and participates in the reaction to every event to deduce the next best actions. This paper investigates the inclusion of preferences with a LLM for smart home automation, The contribution also includes a user-centred representation of smart home states and actions in natural language. Finally, experiments are carried out using several LLMs with different prompting styles for decision-making in the smart home.

Section II presents related work in the literature. The proposed software architecture is detailed in Section III, in particular the integration of LLMs for decision-making combined with preferences. Afterwards, Section IV describes a benchmark dedicated to LLMs and prompting styles with a

set of home scenarios. The results are described in Section V and discussed in Section VI. Finally, Section VII is devoted to conclusion and future work.

II. RELATED WORK

This topic is recent and few papers have been published on LLMs for smart homes. The following papers are selected for their approaches using the knowledge of user preferences on decision-making in smart home automation systems. This section cannot be exhaustive in this larger domain.

Oliveira et al. [6] propose a multi-agent environment with a Belief-Desire-Intention cognitive model, to support adaptivity and preferences transparently in a smart home environment. Any possible interaction could be modeled in such a way as to allow alternative proposals, but this requires considerable work in semantic representation to be fully complete.

Another paper takes advantage of general knowledge information to improve system adaptivity to preferences, [7] proposes an approach using three Knowledge-Based systems, one with general knowledge, one with skill knowledge, and one with contextual information, such as device location, and generating rule models for a middleware platform based on all this information. This approach requires really strict semantic modelling to handle most scenarios, and cannot take advantage of some information that is implicitly given in context.

Shuvo [8] proposes an Actor-to-Critic (A2C)-based algorithm adapted to decision-making in smart homes for energy consumption. In this work, at each step, an A2C algorithm is applied to each device to select the best action, using as inputs the activity and price of electricity at that time. A key element is that the set of actions for each appliance depends on the category associated with the appliance, adding initial knowledge to the model to ensure action based on the importance of the appliance. This system requires learning, and any change in device availability leading to different optimal decisions will require many steps before adapting. In addition, the larger the context, the more difficult it will be to converge for each scenario to learn the best-suited device state. Zhang et al. [9] also propose a system for energy consumption in smart homes, it combines a set of control rules and reinforcement learning to reduce the adaptation time.

Peng et al. [10] describe an approach for decision-making in home automation using deep reinforcement learning. It showed the ability to learn when to turn on a light but with a really limited context supported, so the application is limited to learning when to turn on a light with schedules of 15 minutes, which is quite a large period for this application.

The first approach to smart home automation with LLM [11][12] uses a JSON (JavaScript Object Notation) [13] data representation from a smart home middleware platform and experiments with an LLM to select an action based on a user request. This approach is a first step towards the use of the general knowledge provided by these models. However, it does not support user preferences, and the idea is to select actions based on an initial user request, rather than proactively. These works have proposed ways of managing decision-making, but

are limited by the contextual data supported. Works using symbolic AI methods can show great adaptability, but at the cost of extensive semantic modelling and with requirements to make them adaptable to future changes in preferences. A new method is proposed to support contextual data while adding preferences.

III. PROPOSED ARCHITECTURE

This study proposes a new architecture for decision-making in smart home automation systems. The system uses Large Language Models and proposes methods to add user preferences, in order to select an action according to context and users. It aims to be a proactive system. At every event occurring in the home, the system proposes actions on devices to align the change home state with user needs and preferences. The system supports different types of data thanks to LLMs ability to process data while generating a textual representation of the home based on device states and the action list. This reason leads directly to the use of RAG or directly injecting the knowledge into the prompts for retrieving updated preferences at every execution time of the AI, instead of fine-tuning.

A home simulator is implemented, it takes information on the configurations of sensors and then generates a textual representation, it is also used to generate the list of actions from its data.

This section details the main components designed for the proposed system architecture. Figure 1 shows this architecture, including the simulator.

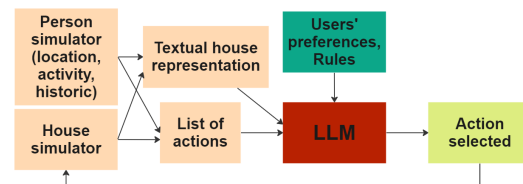


Figure 1. The implemented architecture.

The system generates a user-centred text description of the home and a list of actions, with control over connected devices. It filters the number of relevant actions that can be taken in every situation. For example, it limits some actions that may be prohibited to guarantee user safety. This list of actions is used by the model to select the optimal action.

Concerning contextual data representation editing:

- User positions are listed with their current activity and the history of previous activities.
- The history of previous actions performed in the house is supported.
- All rooms, sensors, and actuators are presented using their names, which the user optionally gives.
- Some sensors and actuators give more global data and control, e.g., temperature sensors, humidity sensors, air-conditioning, etc.

The implementation supports some device categories with a dedicated natural representation, to generate more natural


```

Require: userid, devices_list
1: for all device_name, device_kind, device_location, device_state in devices_list do
2:   if device_kind = "actuator" then
3:     if user_location[userid] = device_location or device_state = 1 then
4:       devices.append(device_name is device_state)
5:       action_vector.append(1)
6:     end if
7:   end if
8: end for
9: devices.append("Interact with user")
10: action_vector.append(2)
11: devices.append("No action required")
12: action_vector.append(0)
13: return (action_vector, devices)

```

Figure 2. Dynamic Devices: Action builder Algorithm.

sentences adapted to some types of data: lights, CO2 sensors, smart curtains, etc. It could also support any additional data sensors with a generic representation template, using the device name in the smart home environment and data status. Using meaningful naming added by the users helps the system to understand the usage of the device.

The action proposal Algorithm 2 considerably reduces the set of possible actions. It assumes that only devices that are in the room or global devices can be switched on, but that all switched-on devices can be switched off. As far as the list of possible actions is concerned, the idea is to filter the actions supported according to some conditions and device types. This approach is made possible by LLM's native support for a change in the output action space, without requiring training.

Any type of device can be easily supported: it simply has to be added to the representation and the LLM will ingest the data thanks to its internal knowledge, learned during model pre-training. This knowledge allows the LLM to get a natural human description of the home including biased contextual data unlike many conventional home automation systems. The latter do not take advantages of information, such as the names of lights or rooms. This data is transmitted to the LLM using one of the prompting styles, prompting being the way to call the LLM with contextual arguments. The different styles of prompts will be described in the following section.

A common aim for all the prompting styles is to take advantage of the knowledge of the overall world provided by this Large Language model, to handle changes or even new types of sensors added to the representation.

The user preferences and rules block represents a database containing information about the system's basic rules, generalities about human preferences and specific user preferences.

A benchmark is established with predefined context-aware scenarios for evaluation purposes.

Regarding LLMs, recognized top-performing models (relative to their parameter count at the time of evaluation) from reputable research labs were selected. Selected models were run using a local inference engine backend.

IV. EXPERIMENTATION

Different objectives are defined for the experimentations.

- Evaluate the improvements provided by adding user preferences, with various techniques of doing so.
- Evaluate the improvements of natural language representation of a smart home automation state over a JSON representation, as LLMs are trained on natural language corpus. Even if they contain other kinds of data like code.

In alignment with the objectives of experimentations, different metrics are used for these evaluations:

- Grades: The grade evaluates the response of a model combined to a prompting style when running a scenario. Grade values are defined in Table I.
- Processing time: Total runtime, including the construction of the context data and action list representation, the inferences with the prompting styles, the use of RAG if the prompting style uses it, and the processing of the formatted LLM response.

Eleven evaluation scenarios are defined as starting points, with predefined actions graded 0-2 based on user satisfaction. Table I presents the scenarios by name and associated reward values. A category is associated with each scenario, the goal being to group the scenario with the name of the main evaluated ability. The database of preferences and rules is defined in

TABLE I. SCENARIO RESPONSES WITH GRADES, ASSOCIATED ANSWERS, AND EVALUATION CATEGORIES.

Scenario Name	Grade	Associated Answer	Category
Out of bed at night	2	Turn on auxiliary light or main light with reduced luminosity level	Safety
	1	Turn on main light	
	0	Everything else	
Watching TV: late evening	2	Turn on auxiliary light or main light with reduced luminosity level	Comfort
	1	Turn on main light, open curtains, discuss	
	0	Everything else	
Out from bed issue with CO2	2	Inform user of risk	Safety
	1	Do an action and inform the user of risk	
	0	Everything else	
Going back to bed at night	2	Turn on auxiliary light or main light with reduced luminosity level	Safety
	1	Turn on main light	
	0	Everything else	
Evening sleeping: TV ON	2	Turn off TV	Preference
	1	Turn off anything on	
	0	Everything else	
At dinner watching TV	2	Turn on auxiliary light or main light with reduced luminosity level, open curtains	Preference
	1	Turn off the main light, do nothing	
	0	Everything else	
User out: TV is on	2	Turn off TV, turn off HVAC	Comfort
	1	Turn off all lights	
	0	Everything else	
Too low temperature	2	Turn on HVAC	Preference
	1	Open Curtains	
	0	Everything else	
Low luminosity day	2	Open curtains	Preference
	1	Turn on any light in the room	
	0	Everything else	
Failed curtains	2	Turn on any light of the room	Comfort
	1	Open curtains	
	0	Everything else	
Forgot to turn off lights	2	Turn off any lights, or HVAC	Preference
	1		
	0	Everything else	

a single file for all scenarios. These data are naturally written sentences. At the end of each one, an information level is

recorded: Rules, Preferences, Generality. The idea is to transmit to the LLM the importance of each data through keywords. Generality is considered the least important, Preferences the second most important, and Rules the most important. The database includes some preferences, generality and some rules. It is designed to handle some scenarios, help in some others but does not provide a solution for all scenarios. The data are fed into a vector database so that RAG may be used instead of prompts to convey them to the LLM. Four different prompting styles are compared on all the scenarios:

- direct: A system prompt and a prompt to request direct answers in the specified format.
- directPref: A system prompt with the preferences, rules and generality from the database and a prompt to request answers in the specified format.
- OpenQuestion: a two-steps chain. First a system prompt and a prompt to request "a list of 3 main problems". For each of the 3 problems, the LLM is invited to use RAG to get the 3 closest preferences. Then a prompt to request answers in the specified format.
- ThreeQuestion: a three-steps chain. First a system prompt and a prompt to request "a list of 3 main problems". For each of the 3 problems the LLM is invited to use RAG to get the 3 closest preferences. Then a prompt to request answers in the specified format (twice). Finally a prompt to select the best answer in the specified format.

A common point between all the prompting styles is the action expected in the output: as mentioned above, most supported devices, such as lights or HVAC systems, are considered as switches in the action list; therefore, in the output of all prompting styles in addition to a "reasoning" and an "action" key, three optional keys are available: temperature, luminosity and explanation. It enables the model to respectively modify the temperature of an HVAC system when executing a related action, modify the luminosity by dimming a light or give an explanation to transmit a sentence to the user.

Two ways of representing the state of the house data are implemented, both using the same input data:

- JSON: A JSON representation
- Textual: A fully natural textual representation

The implementation of the system is evaluated using various open-sources LLMs, including:

- Starling Alpha 7B [14]- 8bpw (bits per weight)
- Qwen 1.5 14B [15] - 5bpw
- Qwen 1.5 72B [15] - 3.5bpw

The three models are selected for their performance and for covering the three main open-source model sizes. They are used to evaluate the impact of proposed prompting methods and data representation.

Qwen 1.5 72B, with around 72 billion parameters, is currently one of the best models available open-source. Starling 7B Alpha, with around 7 billion parameters, is an excellent smaller model. It is based on Mistral 7B, an efficient model for its size on various benchmarks. Qwen 1.5 14B model, a smaller version of Qwen 1.5 72B, is selected to add an

intermediary model.

All these models are used with versions that are quantized [16], a technique used to reduce inference time and memory footprint, the quantization chosen for each model is given in bits per weight (bpw). With their quantization, they require around 8GB, 12GB, and 44GB of memory respectively.

Every model is evaluated on local instances, served locally with an engine-based API backend, using TabbyAPI [17], based on ExLLamaV2 multiple GPUs (graphics processing units) and without automatic splitting, using a workstation equipped with a Ryzen 9 7950x, 96GB of DDR5 memory running at 5600mhz and 2 Nvidia RTX 4090, each with 24 GB dedicated memory, running Ubuntu 23.10.

Experiments are carried out beforehand on various uncontrolled scenarios to define LLM parameters. With the sole aim of reducing non-determinism from one cycle to the next, the final parameters modified from the default engine parameters are as follows:

- max_tokens, maximum number of tokens in output: 300
- min_p, minimum percentage value that a token must reach to be considered (Value is scaled based maximum token probability): 0.05
- temperature, parameter that regulates the randomness: 0.2

The RAG is implemented using Langchain [18] with an inference engine from HuggingFace [19], to locally execute an embedding model: BAAI/bge-large-en-v1.5 [20], and an Elasticsearch [21], local instance is used as vector database, both instances running on main processor and associated memory.

To evaluate system performance, each scenario is executed 10 times with each prompting style, and a grade is given to each response. Results then count the total number of points for each prompting category in general, and also for each defined metric associated with each question. The system's complete processing time is also measured, in order to estimate the average latency of the different prompting styles.

The theoretical random action grade for each scenario is calculated as a baseline using the following (1).

$$\text{grade}_s = \frac{\text{num_actions_graded_1}_s + 2 \times \text{num_actions_graded_2}_s}{\text{total_number_of_actions}_s} \quad (1)$$

With num_actions_graded_X the number of actions associated with grade X. Figure 3 shows an example of a scenario data representation using proposed natural language textual representation, it shows an example of the generated contextual representation transmitted to the LLM on scenario 1.

As previously mentioned, several data types have been included in the representations: numerical values, boolean values, and character strings. Using an LLM to process context allows the system to support all these data types.

Regarding the action list building algorithm, in one of the scenarios for example, it reduces the set of actions from more than 18 to 6 actions, which represents a 3-fold reduced set in this case. This reduced set takes into account the fact that lighting appliances are basic switches and cannot be set with

Current State of the House:
 User 1 is in the Livingroom. User is watching TV.
 Previously: User was watching TV

Livingroom: Curtains are Closed.
 Lights: main, floor lamp are respectively Off, Off.
 There is a TV in the room and its state is on.
 CO2 level in room is 513ppm.

.....
 House was cleaned today, expected cleaning one time a week.
 Centralized HVAC system is on with objective to 20°C.
 Entrance smart Door is locked.
 Time: 10:21 PM
 Global house temperature is 20°C, outside temperature is 5°C.

Figure 3. Extract of natural data representation on the scenario "Out of bed at night".

different luminosity levels. No further studies assessed the improvement provided by the action filtering for the LLMs.

This limited set of actions makes it interesting to use a model randomly selecting one action from the set as a baseline. With a reduced action space, a random answer may be good.

V. RESULTS

This section will present the results of our simulations, focusing on various aspects. Initially, we will explore the impact of data representation, followed by an analysis of the results concerning user preferences.

A. Data representation

The first analysis of the results focuses on the advantage of using a natural representation versus a JSON representation. Table II shows an average difference between results using

TABLE II. COMPARING THE TWO CONTEXTUAL REPRESENTATIONS: AVERAGE GRADE OF MODELS PER EXECUTION PER SCENARIO.

Model	JSON	Natural
Qwen 1.5 72B	1.25	1.18
Qwen 1.5 14B	0.63	0.99
Starling Alpha 7B	1.03	1.18

JSON representation, or the most natural representation. It shows that the larger models are almost stable independently of the contextual representations, and on average even slightly better using JSON representation, by around 5.9%. However, on smaller models, the natural language representation greatly improves performances, with a 14.6% increase in average performance using Starling 7B, and a 57.1% increase using Qwen 1.5 14B. In terms of processing time, results are quite similar on average for models with both representations. On average, across all models, it leads to an increase in performance of 21.9%, despite the results on the larger Qwen model, making the natural representation more efficient.

B. User preferences

Figure 4 depicts the average grades obtained by the 3 chosen models on their responses to scenarios with the 4 different prompting styles and the 2 distinct representation types.

It first shows that Qwen 72B model is much more stable than the other two and that the prompting style does not have

as much impact on response quality. Regarding the two smaller models, larger inconsistency in results can be noted as varying with the chosen prompting styles, particularly with Qwen 14B model. The results of LLMs are compared with a baseline corresponding to the random choice of an action. The results remain on average behind any model without preferences (37.1% below the worst result with the "direct" prompting style). However, thanks to the algorithm reducing the set of actions and the fact that multiple responses are acceptable on each scenario, the random baseline obtains grades that are sometimes better than LLMs.

The sequence of multiple questions requires the model to be consistent and to respect the expected instruction format. Furthermore, given that only prompt engineering is used to ensure the format, some prompting styles with multiple questions may lead to invalid responses, forcing the model to take a default action in the proposed setting. This default action is set to do nothing and to inform the user that it has failed to act. This reduces the performance of some models with some prompting styles. In Figure 4, failure ratio measures the ratio of invalid responses. Qwen 14B results are especially below this baseline because they failed to answer in a large number of scenarios

On average, "directPref" prompt models achieved a gain of 11.3% over basic prompt with natural representation, and even 20.0% with JSON representation.

The best results are achieved with JSON representation for "OpenQuestion" and "ThreeQuestion", leading to a 28.6% improvement over the "direct" prompting style.

With the natural representation, Qwen 14B model gives almost stable results with all prompting styles except "OpenQuestion". The main difference with "OpenQuestion" seems to be linked to a high failure rate, as shown in Figure 4. With JSON representation, failures are so high with all advanced prompting styles that the best results are obtained with the most basic prompt. In this case, "direct" prompting style is 30.7% more efficient than the "directPref" prompting style.

On Starling 7B, with both representations, the best results are obtained with "directPref", on the Starling representation, its results are 52.3% better than the direct representation and respectively 8.1% and 6.3% better than "OpenQuestion" and "ThreeQuestion" prompting styles.

In addition, the average processing time is reduced by 35.6% on Starling 7B Alpha using the "directPref" prompting style instead of "direct".

With JSON representation, the averaged grade results follow the same trend, but with lower overall values.

Table 4 shows that RAG-based prompting models can lead to better results than a single prompt containing all the data, as it is visible Qwen 72B. However, to date, LLM inference is still slow, and the use of complex prompts leads to a loss of accuracy in the test scenarios with smaller models.

Adding inference time to the balance with Table III highlights Starling 7B Alpha results with "directPref". It outperforms almost all others except Qwen 72B with JSON

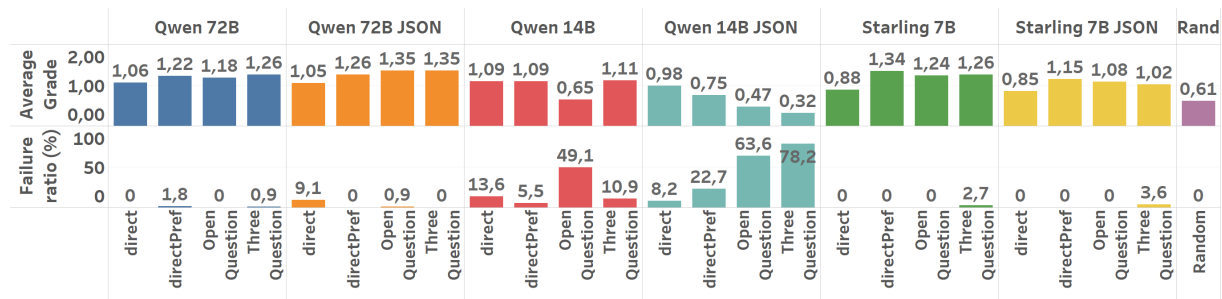


Figure 4. Average grades by model, data representation, and prompting styles.

TABLE III. COMPARISON OF AVERAGE GRADES FOR EACH PROMPTING STYLE AND MODELS WITH INFERENCE TIME.

Model	Prompting Style	Average grade	Proces. time (s)
Qwen 72	direct	1.06	9.04
	directPref	1.22	7.01
	OpenQuestion	1.18	24.43
	ThreeQuestion	1.26	42.13
Qwen 72 JSON	direct	1.05	9.02
	directPref	1.26	6.71
	OpenQuestion	1.35	25.20
	ThreeQuestion	1.35	41.47
Qwen 14	direct	1.09	1.22
	directPref	1.09	1.16
	OpenQuestion	0.65	8.83
	ThreeQuestion	1.11	15.54
Qwen 14 JSON	direct	0.98	1.28
	directPref	0.75	1.21
	OpenQuestion	0.47	9.39
	ThreeQuestion	0.32	16.62
Starling	direct	0.88	0.73
	directPref	1.34	0.47
	OpenQuestion	1.24	3.98
	ThreeQuestion	1.26	6.23
Starling JSON	direct	0.85	0.48
	directPref	1.15	0.48
	OpenQuestion	1.08	3.77
	ThreeQuestion	1.02	6.20

representation and prompt chaining, but is 53.6 times faster, with a lower grade of just 0.7%.

With Qwen 72B, the use of more complex models (OpenQuestion, ThreeQuestion) can lead to better results, as seen previously in particular with JSON representation, but this comes with a trade-off: inference time. Qwen 72B model is already much slower due to its number of parameters, and due to the number of operations required to produce a single token. For instance, an inference with "directPref" takes 6.86 seconds on average (JSON and textual representation), whereas using "OpenQuestion" (which is around 60% faster than "ThreeQuestion") is 3.6 times slower.

In the current state of this type of hardware, it is impossible to consider them as a viable alternative for managing a smart home automation system, with such reaction times.

VI. DISCUSSION

This section mainly discusses the results of the LLMs and prompting techniques chosen in order to make choices for real experiments. The advantages of this study are highlighted as well as the new challenges that are raised.

As seen previously, using larger LLM, such as Qwen 72B, allows greater stability in preference-free scenarios. Indeed, Qwen 72 is 19.9% better than Starling 7B in this case. However there are drawbacks, the first being the inference time as mentioned, and the second being the hardware infrastructure required. The quantized version of Qwen 1.5 72B requires 44GB of memory compared with around 8GB for Starling 7B Alpha with lower quantization. Compared with Qwen 72B model using JSON representation, Starling 7B Alpha with "directPref" takes advantage of natural representation and achieves almost similar performances with much more complex prompting techniques. This makes the approach of using Starling 7B Alpha with this prompting style a good choice for future work. It gives similar performances concerning grades, and has a relatively low inference time (Average: 0.47s). In addition, compared to Qwen 1.5 72B with natural representation and no preferences, Starling 7B Alpha's performance is 26.4% better using "directPref", with a processing time almost 20 times faster.

The results show the advantage of adhering to preferences. Drawbacks appear, however, with the additional average computation time for "OpenQuestion" and "ThreeQuestion" prompting styles, which use RAG. Using RAG brings no advantage in most cases. This is certainly due to the relatively small database. If the system required a larger database of preferences and rules, the results might have been different as it would not have been possible to give them directly through "directPref" prompting style. Based on current results, the best choice for a use case with a larger database would remain Starling 7B Alpha, with "OpenQuestion" prompting style and natural representation, as it provides results that are aligned with users preferences, with only 8.1% less average grade than Qwen 1.5 72B, while keeping an acceptable average inference time (3.98 seconds vs 25.20 seconds).

The use of smaller models poses the challenge of enforcing the output format. New methods have recently been proposed, such as Outlines [22], to strengthen the output grammar. They should be tested soon in order to analyze the cost in inference time and the improvement of the results of this paper.

Another approach is fine-tuning. Applying it to user preferences is impractical as they are constantly changing. Applying it to smart home domain would improve the analysis of contextual data while allowing the system to ensure the output

format.

Although the proposed system is not comparable to other works due to its completely new approach, it has the advantages of its nature, as well as drawbacks. Firstly, it is not deterministic. Each configuration has been run 10 times, but some results are different each time, which means that the results of the system cannot be certified and a safety layer must be developed to secure some actions. Secondly, LLM queries are slow and require significant computing and memory resources: adaptability without training comes at a cost.

VII. CONCLUSION AND FUTURE WORK

This paper presents a new architecture for a smart home automation system, using LLMs with user preferences to enhance personalised user experiences. This approach leverages the general knowledge provided by LLMs and combines it with naturally written rules and preferences to make contextually relevant decisions in line with user preferences. This architecture is proactive, able to adapt to any change in the environment thanks to the robustness provided by LLMs.

The experimental results demonstrate the potential of this architecture to improve alignment with user preferences compared with an implementation without user preferences, showing up to 52.3% performance increase, with an average processing time reduced by 35.6% on Starling 7B Alpha LLM using "directPref" prompt style compared to "direct". Additionally, the performance is 26.4% better than that of the largest models evaluated without preferences, while also achieving a processing time that is nearly 20 times faster.

The study showed that, particularly with small models, using a natural representation instead of a JSON one leads to an increase in performance, with an average 21.9% increase.

Although the system shows promising results on a set of defined scenarios, it also presents challenges due to stochastic behaviour and a slower inference time compared to traditional machine learning methods. These drawbacks are offset by the system's ability to adapt dynamically - without retraining - to changes in preferences, appliances, and home configuration.

Future work will focus on implementing the system in a real-world smart home middleware system, such as OpenHAB to evaluate its performance with real users.

REFERENCES

- [1] M. Weiser, "The computer for the 21st century," *SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 3, no. 3, pp. 3–11, Jul. 1999, ISSN: 1559-1662. DOI: 10.1145/329124.329126.
- [2] A. K. Dey, "Understanding and Using Context," *Personal and Ubiquitous Computing*, vol. 5, no. 1, pp. 4–7, Feb. 2001, ISSN: 1617-4909. DOI: 10.1007/s007790170019.
- [3] P. Rashidi and D. Cook, "Keeping the Resident in the Loop: Adapting the Smart Home to the User," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 39, no. 5, pp. 949–959, Sep. 2009, ISSN: 1083-4427, 1558-2426. DOI: 10.1109/TSMCA.2009.2025137.
- [4] A. Radford *et al.*, *Language models are unsupervised multitask learners*, 2019.
- [5] P. Lewis *et al.*, "Retrieval-augmented generation for knowledge-intensive nlp tasks," in *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, Eds., vol. 33, Curran Associates, Inc., 2020, pp. 9459–9474.
- [6] P. F. Oliveira, P. Novais, and P. Matos, "Adaptive environment system to manage comfort preferences and conflicts," in *Optimization, Learning Algorithms and Applications*, Cham: Springer International Publishing, 2022, pp. 685–700, ISBN: 978-3-031-23236-7.
- [7] A. R. Ruiz *et al.*, "Leveraging commonsense reasoning towards a smarter Smart Home," en, *Procedia Computer Science*, vol. 192, pp. 666–675, 2021, ISSN: 18770509. DOI: 10.1016/j.procs.2021.08.069.
- [8] S. S. Shuvo and Y. Yilmaz, "Home energy recommendation system (hers): A deep reinforcement learning method based on residents' feedback and activity," *IEEE Transactions on Smart Grid*, vol. 13, no. 4, pp. 2812–2821, 2022. DOI: 10.1109/TSG.2022.3158814.
- [9] Z. Zhang, Y. Su, M. Tan, and R. Cao, "Fusing domain knowledge and reinforcement learning for home integrated demand response online optimization," *Engineering Applications of Artificial Intelligence*, vol. 121, p. 105995, 2023, ISSN: 0952-1976. DOI: <https://doi.org/10.1016/j.engappai.2023.105995>.
- [10] Z. Peng, X. Li, and F. Yan, "An adaptive deep learning model for smart home autonomous system," in *2020 International Conference on Intelligent Transportation, Big Data and Smart City (ICITBS)*, Los Alamitos, CA, USA: IEEE Computer Society, Jan. 2020, pp. 707–710. DOI: 10.1109/ICITBS49701.2020.00156.
- [11] E. King, H. Yu, S. Lee, and C. Julien, "'Get ready for a party': Exploring smarter smart spaces with help from large language models," *Unpublished*, 2023. DOI: 10.48550/ARXIV.2303.14143.
- [12] E. King, H. Yu, S. Lee, and C. Julien, "Sasha: Creative goal-oriented reasoning in smart homes with large language models," *Unpublished*, 2023. DOI: 10.48550/ARXIV.2305.09802.
- [13] F. Pezoa, J. L. Reutter, F. Suarez, M. Ugarte, and D. Vrgoč, "Foundations of json schema," in *Proceedings of the 25th International Conference on World Wide Web*, International World Wide Web Conferences Steering Committee, Association for Computing Machinery, 2016, pp. 263–273.
- [14] B. Zhu, E. Frick, T. Wu, H. Zhu, and J. Jiao, *Starling-7b: Improving llm helpfulness and harmlessness with rlaiif*, Nov. 2023.
- [15] J. Bai *et al.*, "Qwen technical report," *Unpublished*, 2023.
- [16] Turboderp, *Exllamav2*, <https://github.com/turboderp/exllamav2>, 2023.
- [17] TheRoyallab, *Tabbyapi*, <https://github.com/theroyallab/tabbyAPI>, 2023.
- [18] O. Topsakal and T. C. Akinci, "Creating large language model applications utilizing langchain: A primer on developing llm apps fast," in *International Conference on Applied Engineering and Natural Sciences*, vol. 1, All Sciences Academy, 2023, pp. 1050–1056.
- [19] HuggingFace, *Text embeddings inference*, <https://github.com/huggingface/text-embeddings-inference>, 2023.
- [20] S. Xiao, Z. Liu, P. Zhang, and N. Muennighoff, *C-pack: Packaged resources to advance general chinese embedding*, 2023. eprint: 2309.07597 (cs.CL).
- [21] Elastic, *Elasticsearch*, <https://github.com/elastic/elasticsearch>, 2015.
- [22] B. T. Willard and R. Louf, "Efficient guided generation for llms," *arXiv preprint arXiv:2307.09702*, 2023.

One Cash

A New Generation of E-wallet for Mobile Money

Macaire Ngomo

Department of Engineering, Research, Innovation
CM IT Conseil
Romilly sur Seine, France
e-mail: macaire.ngomo@gmail.com

Clavy Viannet Mbou Mouyé

Engineering Department
Imagine Technology
Brazzaville, Congo
e-mail: clavymbou12@gmail.com

Abstract— Launched in Congo in 2012 by mobile phone operators, Mobile Money is a digital financial service that offers the possibility of opening an account for free and carrying out financial transactions easily and flexibly. Subscribers can then carry out several types of operations: pay bills, transfer money, receive money, deposit money, etc. In Congo, in addition to the fact that financial transactions are still very expensive, the solutions offered by operators currently only allow transactions to be carried out between subscribers of the same mobile operator. In addition to this lack of a gateway between operators, the old Unstructured Supplementary Service Data (USSD) or Subscriber Identity Module (SIM) hierarchical menu interfaces used do not facilitate or encourage the use of the solutions offered. A single transaction can require up to seven or eight screens and the entry of two to three sequences of numbers. In this article, we propose an innovative solution, called "One Cash", which provides a concrete and practical solution to these major concerns, by integrating the Application Programming Interfaces (APIs) of mobile and/or financial operators and by exploiting the richness and flexibility of smartphone interfaces to simplify use. In particular, the content is presented in a richer and clearer way than a simple text; the number of manipulation steps has been reduced with the proposed solution, thus reducing the cognitive load of the user. We highlight the interest of a diversified offer of m-payment services including a simplified solution offered by the operator alone, and more elaborate services, in particular by associating technological and financial partners such as banks. The proposed architecture is easily adaptable to respond to future extensions and improvements of the system.

Keywords—Ubiquitous computing support; e-payment; m-payment; mobile money; mobile telephony; User Interface/eXperience User; usage; useful; transactions; financial inclusion; Unstructured Supplementary Service Data (USSD); lower-cost and user-friendly solutions.

I. INTRODUCTION

Mobile banking [4]-[7][10][12][15][20][26] is a technological ecosystem that allows, from a mobile terminal [21], to access financial services: mobile payment, money transfer, mobile banking and mobile finance (microcredit, microinsurance or savings). Transactions can be done via the Internet, directly between two phones or between a phone and any other type of terminal [5]. The impact of mobile banking is totally different between developing and

developed countries. In developing countries, while the low penetration rate of banks and the number of mobile phones sold has enabled mobile phone providers to understand the marketing and financial interest of this means of communication, in developed countries the impact remains minimal since mobile banking is considered an additive channel, not dissociable from a bank account. Offering a much narrower range of services than those available in developing countries and incorporating few or no payment solutions, these services remain incomplete and do not allow unbanked populations to access a first level of banking inclusion [23]. It is, therefore, difficult to propose mobile banking as a solution to unbanking in developed countries. It is also important to note that access to banks in developed countries is facilitated by a high penetration rate and a good network of branches. With 8% of the population unbanked in Organisation for Economic Co-operation and Development (OECD) countries [7], banks do not see the financial interest in developing platforms and managing these customers for obvious profitability reasons.

The African continent has thus become one of the main areas for the deployment of mobile money and e-payment. Under-banking partly explains the success of mobile payment services in Congo and in developing countries.

A. *Local context of the banking sector, issues of bancarisation in Congo and e-payment system in Congo-Brazzaville and in Africa*

(a) *The banking sector and the challenges of banking in Congo*

After more than three decades of crisis, the banking landscape in the Republic of Congo has been transformed from a fragile context marked by repeated crises and bankruptcies to a reassuring paradigm that is solid, integrated and increasingly systemic. With about ten active credit institutions (excluding microfinance institutions and development banks), we can distinguish: pan-African groups (UBA Bank, EcoBank, BGFI Bank), Moroccan groups (LCB Bank, CDCO Bank, BCI Bank), the French bank "Société Générale Congo", and finally Congolese banks, notably BSCA Bank, Banque Postale, BCH Bank and BESCO Bank. The situation here is dominated by foreign banks, which control almost 90% of the market. The Republic of Congo is a Central African country with a population of 5 million (5.518 million (2020)), of which, despite this strong presence of banking players, barely 13% is banked. This percentage

remains low. This means that 87% of this population is excluded from traditional banks, according to a paper in one of the largest Congolese newspapers, “Brazzaville Dispatches” in French “Les Dépêches de Brazzaville” (2016), on the situation of Congolese banks. When commercial banks are inaccessible to part of the population, they are not considered efficient institutions [1][2][8][11].

(b) Weaknesses in electronic communication infrastructure

In addition, Congo has major weaknesses in terms of electronic communication infrastructure. Optic fiber Internet coverage only serves three urban areas (Brazzaville, Pointe-Noire and Oyo). This is despite Congo's connection to the West Africa Cable System (WACS) submarine fibre-optic cable since 2012 and cross-border interconnections with Gabon and the Democratic Republic of the Congo (DRC).

(c) The combination of banking and technological shortcomings

The combination of banking and technological shortcomings is the breeding ground and the driving force behind mobile money in this country. The Congolese mobile phone market is made up of two mobile network operators, Mobile Telephone Networks (MTN) and Airtel [3]. The operator Azur, although still licensed, is no longer operational. Launched in 2012 by the two mobile telephone operators Mobile Telephone Networks (MTN) and Airtel, mobile money, commonly known as “MoMo”, is a digital banking financial service offered by these operators through mobile phones. Congolese consumers have quickly and massively adopted this new technology based on mobile telephony, which combines the power of information technology [31] and microfinance practices. Faced with the problems of Internet coverage throughout the country, on the one hand, and the difficult conditions and administrative red tape inherent in opening an account in traditional banks in Congo, on the other, mobile telephony has revolutionised the financial ecosystem and economic habits thanks to mobile money. Mobile money allows the holder to transfer money and/or credit from their user account to another account, whether or not they are a subscriber, using mobile phone technology. For some time now, this service has also been used to pay certain bills.

(d) Local electronic payment offers

Airtel Payments Bank is a leading Indian payment bank headquartered in New Delhi. The company is a subsidiary of Bharti Airtel. In Congo, Airtel Money is a service of Airtel that allows money transactions to be made using a telephone. In May 2022, Mobile Telephone Networks (MTN) was pleased to inform all its stakeholders that the subsidiary Mobile Money Congo (MMC) had been created. As an electronic payment company in Central Africa, Mobile Money Congo (MMC) is now authorised to provide payment services. The approval marks a new milestone in the development of the Mobile Money service in Africa and establishes this service as a major driver in the acceleration of financial inclusion [8][13][17][25] for more than a decade.

B. Technological context: the strategic advantage and power of smartphone interfaces

In this context, the problem of improving the services offered to the target populations, accessibility, ease of use and interoperability of systems can, therefore, has been raised [22]- [24]. Little known and abandoned in Europe in favour of mobile web applications since the arrival of the smartphone, Unstructured Supplementary Service Data (USSD) technology continues to be massively used throughout the world and particularly in Africa. Where there is no Internet network and where certain services are difficult to access, USSD opens up the possibility of offering mass digital services. The USSD is a source of innovation. Beyond the basic functions of this protocol, entrepreneurs have quickly understood the development potential of this technology and have used it to respond to numerous concrete uses (USSD and agriculture: informing farmers, use of USSD in health, USSD and education: revising lessons, consulting results, USSD and insurance, USSD and transport, etc.) [18][19]. Its simplicity is one of its advantages. In practice, the user sends a code that connects him in real time to a telephone operator to exchange data. Unlike a Short Message Service (SMS), there is no storage in USSD, the information is only available during the opening of the session and then disappears. USSD is, therefore, a real tool for democratising digital uses. However, it is a technology that remains relatively basic. Firstly, this is the case because each session has a cost (network fees) and a limited opening time (time session). As there is no data backup, the user has to start from scratch. This is costly, but also relatively limited in terms of customer experience: it is impossible to perform long tasks. The main competitor is, therefore, the rise of the mobile Internet. However, while it is clear that its use will become increasingly limited in the future, USSD technology undoubtedly has many years ahead of it. Although Internet penetration is growing rapidly across the African continent, it is still low compared to the rest of the world. This is an obstacle to the development of more advanced technology for a significant proportion of the population. Mobile broadband is still very expensive in some African countries. The power of smartphone technologies and interfaces [31] can contribute significantly to the improvement of services. It is true that smartphones are not yet accessible to many people. But the number of people with smartphones has increased significantly. This increase concerns all strata of the population. African countries in particular are rapidly urbanising. As in Asia, more financial transactions are done via a smartphone than through traditional banking channels. This may include many low-income people who would otherwise not have access to formal financial services. However, interface design could have an even greater impact. Today, the old hierarchical USSD or SIM toolkit menu interfaces alienate many users. A simple cash transaction can require seven to eight menu screens and the entry of two-to-three-digit sequences. Errors are too easy. It is too painful to use often or for everyday purposes. Many stay away completely. Smartphone interfaces not only have the potential to simplify basic

transactions, but they can also solve many other obstacles. The richer graphical interfaces of smartphones can present services more clearly than simple text. Information is presented in an understandable and intuitive way. Ease of use can persuade users to make transactions more frequently.

Unfortunately, early uses of smartphones did not exploit the full capabilities of smartphones. New applications often mimic the hierarchical menus of the USSD, rely too heavily on text or simply try to replicate a laptop interface on a much smaller smartphone screen. To take full advantage of the transformation potential, it is important to invest in User Interface design and User eXperience (UX). Interfaces are a strategic asset. A report on #DesignInTech [16] highlights the growing importance of design in technology. In recent years, experts have been collaborating to collectively help develop a consensus set of principles based on experience, including [9]. This is seen as an emerging area of research that can harness the power of smartphones to better serve users. The User Interface (UI), therefore, plays a key role in optimising accessibility and improving the user experience.

C. The goal of the study

The goal of this research is to propose some advances in the form of new architectures and new solutions. Specifically, we propose a mobile and web-based platform and client application that provides solutions to the problems described, including ease of use and access for all segments of the population. It is a technological solution designed in a particular context of use to improve financial and commercial services or uses in Congo.

II. ARCHITECTURE AND FUNCTIONALITIES

A. Architecture

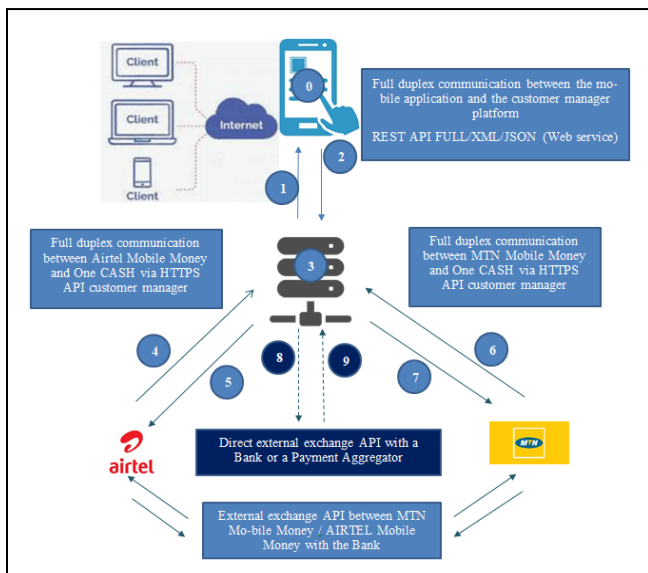


Figure 1. One Cash general architecture and functioning.

One Cash is a payment aggregator facilitating mobile and online payments, an e-payment solution via mobile or web to boost e-commerce and sales. An Application Programming Interface (API) allows the One Cash solution to be integrated into other applications, in particular e-commerce and matchmaking solutions. It includes a web-based management platform in the BackOffice and a mobile application and web interface for managing and viewing transactions in the FrontOffice. It is designed to allow easy payments on mobile and online via e-commerce platforms.

The functioning of the application (Figure 1) is described below.

(a) **The 0, 1, 2, 3 scheme:** For data consultation and information update type transactions, this scheme allows balance consultation, account linking, etc.

(b) **The 0, 1, 2, 3, 4, 5 scheme:** This scheme defines a transaction of type Money Deposit (CASH IN) or Money Withdrawal (CASH OUT) between One Cash and Airtel Money:

o CASH IN case: When a One Cash customer wants to debit his Airtel Money account to credit his One Cash account in order to make an online payment.

o CASH OUT case: When a One Cash customer wants to debit his One Cash account to credit an Airtel Money account.

(c) **The 0, 1, 2, 3, 6, 7 scheme:** This scheme defines a CASH IN or CASH OUT type transaction between One Cash and MTN Mobile Money:

o CASH IN case: When a One Cash customer wants to debit his MTN Mobile Money account to credit his One Cash account in order to make an online payment.

o CASH OUT case: When a One Cash customer wants to debit their One Cash account to credit an MTN Mobile Money account.

Transactions between MTN Mobile Money or Airtel Money and banking systems are handled by external APIs and are, therefore, not handled by One Cash.

(d) **The 0, 1, 2, 3, 8, 9 scheme:** This scheme defines a CASH IN or CASH OUT transaction between One Cash and the Bank. The solution consists of three main modules:

- Client module: One Cash Client (Mobile application: connected or disconnected mode - limited functionality). The mobile application will allow users to make payments from their mobile or via a web interface.
- Server module: One Cash Web-based administration platform (One Cash Manager).
- Dashboard & Customer Manager: The dashboard will be used to manage customer transaction accounts for One Cash.
- Server API: One Cash API interface.

The One Cash API makes it possible to integrate the One Cash solution with the rest of the internal applications of the cluster network (management platform for a central purchasing and e-commerce unit, contact platform for the payment of services) or with other external e-commerce applications through a subscription agreement (Figure 2).

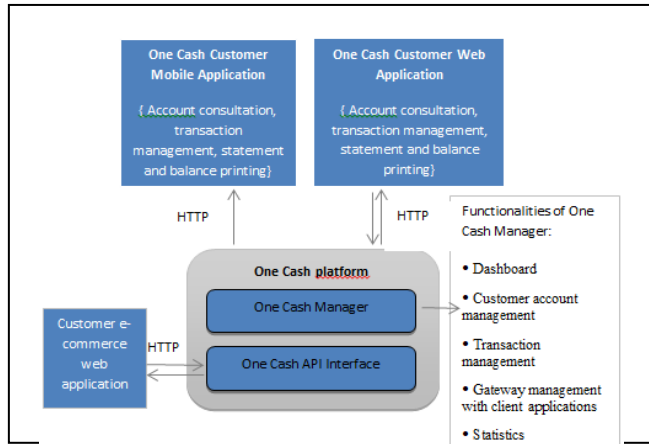


Figure 2. One Cash general architecture.

B. Services offered (functionalities)

1) One Cash Customer

Regardless of the customer application used, Mobile or Web, the One Cash Customer Mobile / One Cash Customer Web application offers the following services:

- Creation of ONE CASH accounts
- Airtel Money to One Cash Transactions (AM2OC) & One Cash to Airtel Money Transactions (OC2AM),
- Mobile Money to One Cash (MM2OC) & One Cash to Mobile Money (OC2MM) transactions,
- Pay online including via QR Code,
- View transaction history,
- Transfer money between two One Cash accounts.

2) The Dashboard & The Customer Manager

- Account management,
- Transaction management,
- Transaction monitoring: Dashboard, Consultation of transaction history (all types of transactions).

In the rest of this paper, we present the interface of the mobile application. The web client application offers the same functionality as the mobile client application. For security reasons, the administration module will only be briefly described in this paper.

III. SOME USE CASES OF ONE CASH INTERFACE

A. One Cash Customer Mobile interface

One Cash Customer Mobile is an Android mobile application that can be downloaded from the Google Play Store or App Store. Once you have downloaded the

application, simply follow the installation instructions which only take two to three minutes. At the end of the installation, the application icon will appear on the smartphone screen, as shown in Figure 3 below:

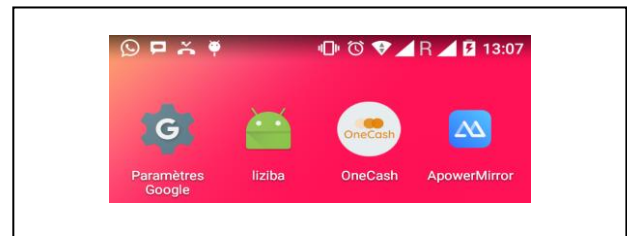


Figure 3. One Cash launch icon after installing the app.

Launching the application is simple using the One Cash icon above. The application then starts up. Once connected, the application displays the user dashboard, as shown in Figure 4.

The dashboard presents:

- The name of the user,
- The available balance,
- Transaction history.

For a debit, the colour of the amount debited and thus transferred to another account is red. For a credit, the colour of the amount received is green. To obtain the details of a transaction, simply click on the corresponding line.

The screen on the right shows the amount of the transaction, the date of the transaction, the type of transaction (sending or receiving), the cost of the transaction if applicable, the sender and the receiver, each identified by their name and One Cash ID.

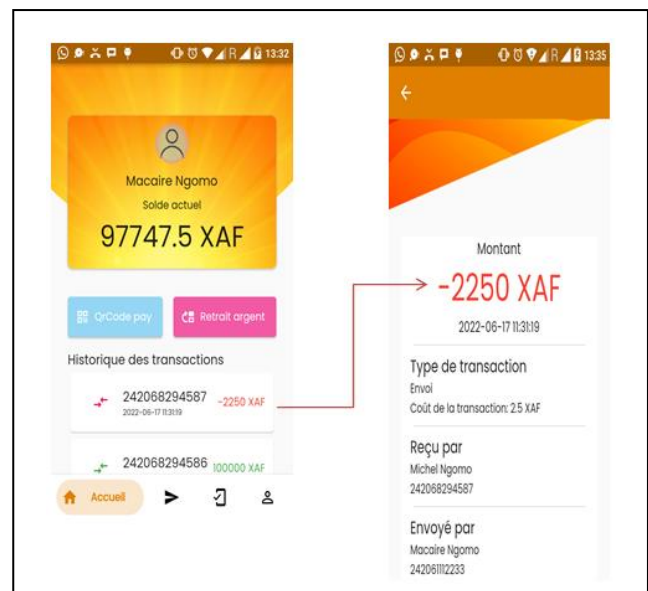


Figure 4. One Cash interface for customer (XAF = CFA franc).

1) Funding / Reloading a One Cash account

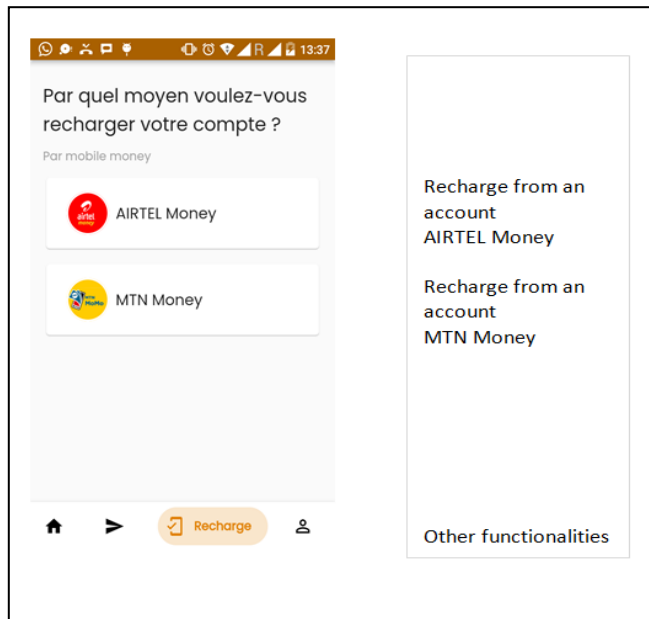


Figure 5. One Cash account funding interface.

One Cash allows the user to fund his One Cash account by debiting his MTN Mobile Money or Airtel Money account. The One Cash account funding is done by the user from his MTN Money or Airtel Money account (Figure 5). To do so, the user simply selects one of the Mobile Money accounts and enters the amount to be transferred to his One Cash account.

The Mobile Money account must have an amount greater than or equal to the amount to be transferred. The user has the same level of security as his Mobile Money account. For security reasons, the user must confirm the transaction by entering the PIN code of his Mobile Money account.

2) Withdrawal or transfer of money to an MTN MONEY or Airtel MONEY account

To withdraw money, the user must transfer from their One Cash account to a Mobile Money account and then make a withdrawal from a kiosk. To do this, the user uses the withdrawal button available on a home screen (Figure 6).

By selecting this operation, they enter the amount to be transferred. They have the same level of security as their Mobile Money account.

For security reasons, they must confirm the transaction by entering the PIN code of their Mobile Money account.

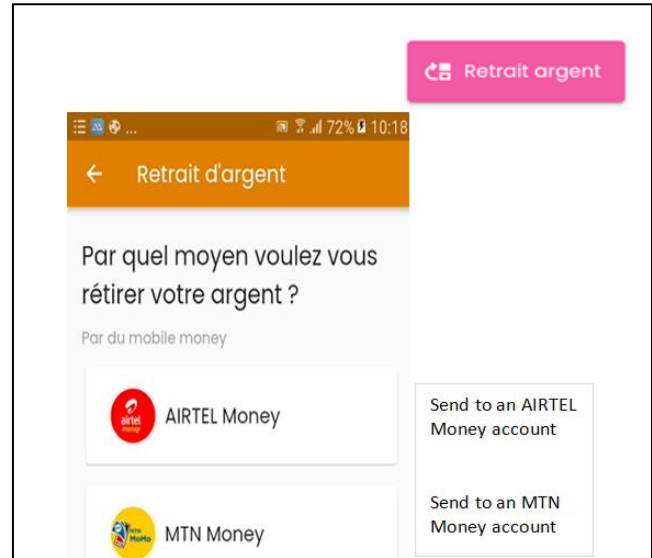


Figure 6. One Cash Transfer interface to an MTN MONEY or Airtel MONEY account.

3) Send or transfer money to another One Cash account

To transfer money from a One Cash account to another One Cash third party account, simply use the transfer button on the home screen. The user can then enter the phone number associated with a One Cash account and the amount to be transferred to the account.

4) Transfer of money by QR Code (QR CODE PAY)

Sending money to a third-party One Cash account can be done by QR Code. To do this, the user uses the QR Code Pay function in the application (Figure 7).

After scanning the QR Code, the phone number of the beneficiary is read by the sending phone. The user then enters the amount to be transferred and confirms the entry. For security reasons, he confirms the transaction by entering his One Cash secret code. If the operation is successful, the application displays the transfer confirmation message.



Figure 7. One Cash QR Code.

B. One Cash Manager interface

1) Dashboard & One Cash Manager

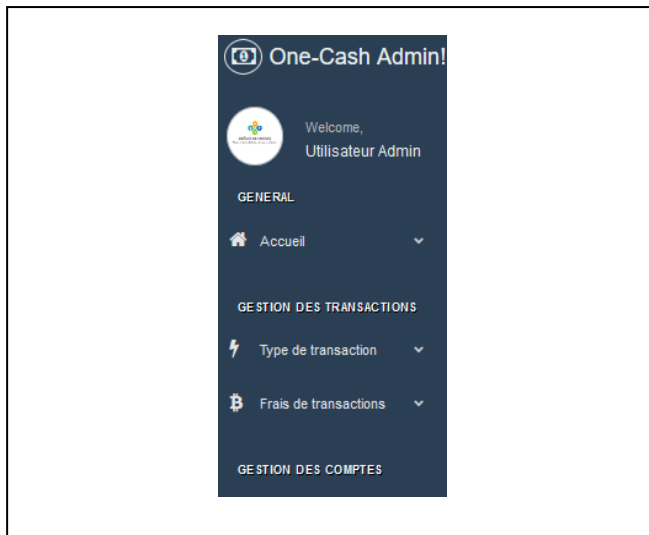


Figure 8. Diagram of master and slave.

As already mentioned, for security and confidentiality reasons, the One Cash Manager administration module will be presented only partially and briefly in this paper. One Cash Manager is a web-based platform designed according to architecture models that meet the requirements of robustness, scalability, interoperability, availability, security, reliability and resilience. It is based on a high-level technological foundation.

One Cash Manager offers a set of services for data administration, application management, transaction management, user account management, transaction cost management, platform configuration, etc. (Figure 8).

This administration module is only accessible to the solution managers.

C. Examples of concrete experimentation

In order to test the effectiveness of the solution, the first concrete experimental use of the One Cash was made within the framework of a program to strengthen the commercial and entrepreneurial capacities of Congolese companies in 2002. This operation was organized to allow nearly 4,500 members of the community of Congolese agribusiness clusters to carry out financial transactions between themselves, without additional costs other than Internet connection fees, only for operations carried out within the network. The second concrete experimental use is within a community of students for the payment of small student life services. Like mobile money or noCash platforms, One Cash is being tested for example in a university cafe-restaurant to allow the payment of restaurant bills or to give tips from a phone in a few clicks without any cash, or to pay for

document reproduction services. Since the regulations that can govern this type of system are not yet ready in Congo, the uses of the One Cash solution remain at the experimental level with voluntary communities of farmers or students who carry out transactions between themselves or with small service providers, by first making a deposit with a trusted third party. These experiments allow us to test different aspects (ergonomics, robustness, scalability, interoperability, availability, security, reliability, resilience, etc.).

IV. CONCLUSION AND FUTURE WORK

There is no doubt that the unbanked market is a market with great potential. Mobile banking can only have a certain future, especially in developing countries. It seems to be the solution to disseminate financial services and bank a new population [17]. Numerous studies [4]-[6][14][17] have shown the strong potential of the mobile phone to provide access to financial services to the unbanked in developing countries while lowering transaction costs. As a high potential market [32], mobile banking seems to have a very favourable future in developing countries. It is, therefore, important to pursue the development of services aimed at mass access and satisfaction of target populations, in line with technological evolution [16][23][24][27]-[31]. This study is in line with the work on the use of ICTs in Africa and in particular in Congo. Its goal is to propose some advances in the form of new architectures and new solutions. Specifically, we propose a platform and a client application in mobile and web versions that exploits the strategic advantages and power of smartphone interfaces to provide solutions to the problems described above, in particular ease of use, transactions that are not limited to a single mobile telephone network, and access to all segments of the population. The number of manipulation steps has been reduced with the proposed solution, thus reducing the user's cognitive load. It is a technological solution designed in a particular context of use to improve financial and commercial services or uses in Congo.

While in this first version of the solution, funding and withdrawal operations can only be carried out between a One Cash account and a Mobile Money account, we plan to extend it by integrating banking APIs and thus allowing direct exchanges between a One Cash account and a bank account. The advantages are significant both for users and for banks, which have been excluded from the mobile banking market since its launch in Congo in 2012 and are only beginning to access it via agreements with the two mobile phone operators. Access to this service for banks will offer great flexibility. The user will be able to carry out transactions remotely, between his One Cash Account and his bank account. In Congo-Brazzaville, some financial institutions are beginning to pool services with mobile phone operators. Unfortunately, transactions are still only one-way, from the operator to the bank. It is not yet possible to fund a mobile money account from a bank account. It is important to develop not only this interoperability between operators, financial institutions and payment aggregators, but also for these institutions to facilitate financial operations thanks to the potential of the new services. So, there is an opportunity

to be seized now. There is a real need for companies to integrate mobile money APIs and financial institutions into their applications. This trend towards digitisation of transactions through mobile money will accelerate the adoption of mobile money APIs and initiate a change in the digitisation strategy. Until now, providers' applications did not communicate with the outside world. With this interconnection of applications, new partnerships and new services will emerge. Ultimately, the end customer's experience will be greatly improved. They will benefit from a simpler, less costly service that offers new possibilities. Moreover, in this way, new uses will be established and will become the norm in terms of customer service. Mobile money will thus contribute to the acceleration of digitalisation by establishing the use of mobile money APIs as the norm. Then companies will really start to integrate these APIs into their services. Companies will then gradually evolve towards a company that assembles services in the same chain to build new services. This allows for greater agility, lower costs, and building the best possible services for the end customer. This will of course lead to new points of tension in the customer experience with new uses. New players will then emerge to solve these problems, or existing players will evolve their services to do so. In both cases, the end-customer experience will be continuously improved. The proposed system architecture is easily adaptable to meet future extensions and improvements.

REFERENCES

- [1] F. Allen, E. Carletti, R. Cull, J. Qian, L. Senbet, and P. Valenzuela, "Improving Access to Banking: Evidence from Kenya", Forthcoming, *Review of Finance*, August 28, 2020. Available at SSRN: <https://ssrn.com/abstract=3688708> (consulted on 2024-05-08).
- [2] F. Allen, E. Carletti, R. Cull, J. Qian, L. Senbet, and P. Valenzuela, "Improving Access to Banking: Evidence from Kenya", *Review of Finance*, vol. 25, issue 2, pp. 403–447, <https://doi.org/10.1093/rof/rf>, March 2021.
- [3] ARPCE, "Le marché de la téléphonie mobile au Congo", <https://www.arpce.cg/article/marche-telephonie-mobile>. (consulted on 2023-08-15).
- [4] D. Assadi, and A. Cudi. "Le Mobile Banking est-il efficace pour offrir des services financiers aux non-bancarisés? Une étude exploratoire", IN : ITEM2, *Microfinance : de la charité aux affaires*, Dijon, France, 2010.
- [5] D. Assadi and A. Cudi, "Le potentiel d'inclusion financière du Mobile Banking. Une étude exploratoire", *Management & Avenir*, vol. 6(46), pp. 227-243, 2011, DOI: 10.3917/mav.046.0227. URL : <https://www.cairn.info/revue-management-et-avenir-2011-6-page-227.htm>. (consulted on 2024-03-15).
- [6] S.-J. Barnes B. Corbitt, "Mobile banking : concept and potential", *International Journal of Mobile Communications*, vol. 1, no. 3, pp. 273-288, 2003.
- [7] C. Beshouri and J. Gravråk, "Capturing the promise of mobile banking in emerging markets", *McKinsey Quarterly*, 2010, https://www.mckinseyquarterly.com/Capturing_the_promise_of_mobile_banking_in_emerging_markets_2539 (consulted on 2024-05-15).
- [8] L. Chaix and D. Torre, "Le double rôle du paiement mobile dans les pays en développement", *Revue Économique*, Vol. 66(4), pp. 709-733, July 2015. DOI:10.3917/reco.664.0703 (consulted on 2023-11-25).
- [9] G. Chen, "The Power of Smartphone Interfaces for Mobile Money", October 2016, <https://www.cgap.org/blog/power-smartphone-interfaces-mobile-money> (consulted on 2023-12-20).
- [10] S. De Sousa, "The role of payment systems in reaching the unbanked", *Journal of Payments Strategy and Systems*, vol. 4, no. 2, pp. 148-155, 2010.
- [11] P. Dupas, A. Keats, S. Green, and J. Robinson, "Supply and Demand Challenges in Banking the Rural Poor, Evidence from Kenya", NBER, Working Paper, 17851, 2012.
- [12] P. Fiorina, "Les services mobiles dans la relation client", In: Investance Institute, *A new customer relationship Management : between loyalty and profitability*, Paris, France, Mars 2009.
- [13] P. Kengue Mayamou and M. Bidan, "Chapitre 27. Mobile Money (MoMo) ou l'histoire d'un portefeuille électronique devenu star au Congo-Brazzaville et ailleurs", Dans : Soufyane Frimousse éd., *Africa Positive Impact. Agir pour un meilleur impact sociétal*. Caen, EMS Editions, Académie des Sciences de Management de Paris, pp. 287-293., 2020, DOI : 10.3917/ems.frimo.2020.01.0287. URL : <https://www.cairn.info/--9782376873976-page-287.htm>. (consulted on 2024-04-10).
- [14] M. Koriko, "Determinants of Ict Adoption", In *The Financial Sector In Sub-saharan Africa*, SSRG International Journal of Economics and Management Studies, vol. 7, no. 9, pp. 48-55, Sep. 2020. Crossref, <https://doi.org/10.14445/23939125/IJEMS-V7I9P106>.
- [15] S. Loaba, "The impact of mobile banking services on saving behavior in West Africa", *Global Finance Journal* 53(4):100620, February 2021, DOI:10.1016/j.gfj.2021.100620.
- [16] J. Maeda, "Design in Tech Report 2016", At SXSW : <https://designintech.report/2016/03/13/design-in-tech-report-2016/>. (consulted on 2024-05-10).
- [17] P. T. Massanga and T. Miere, "Des opérateurs de téléphonie mobile aux opérateurs financiers : usage du mobile money au Congo-Brazzaville", *Revue française des sciences de l'information et de la communication*, Sep. 2020. URL: <http://journals.openedition.org/rfsic/9767> ; DOI : <https://doi.org/10.4000/rfsic.9767>. (consulted on 2024-05-15).
- [18] MONISNAP, "L'USSD, une technologie méconnue qui fait des merveilles en Afrique", October 2019. Available at: <https://www.monisnap.com/fr/blog/article/ussd-en-afrique/> (consulted on 2023-12-25).
- [19] Myriadgroup, "L'USSD - des services mobiles qui renforcent l'efficacité des fermiers et transforment l'agriculture sur le continent", Available at: <https://connect.myriadgroup.com/lussd-et-agriculture/> (consulted on 2023-12-25).
- [20] K. Moran, "Pilot Survey Highlights", *FinScope Mobile Banking*, South Africa, 2006.
- [21] A. Ndiaye and W. Adoum, "Mobile-banking, un potentiel d'inclusion financière au Sénégal : le rôle de la proximité géographique des points de service de transfert d'argent sur l'adoption", *African Economic Conference 2020*, <https://aec.afdb.org/en/papers/mobile-banking-un-potentiel-dinclusion-financiere-au-senegal-le-role-de-la-proximite-geographique-des-points-de-services-de>

- transfert-dargent-sur-ladoption-413 (consulted on 2024-04-15).
- [22] M. Ngomo and H. Abdulrab, "Application Service Provider System: Using Web Services to Provide Interoperability between Learning Management Systems", International Conference WTAS 2006 (Web Technologies, Applications, and Services), July 17-19, 2006, Calgary, Alberta, Canada, Editor(s): J.T. Yao; pp. 119-125, 2006.
 - [23] M. Ngomo and H. Abdulrab, "Application service provider system: the new way to provide interoperability between learning management systems", International Conference Applied Computing 2007, IADISI2007, Salamanca, Spain, pp. 18-20, February 2007.
 - [24] M. Ngomo, H. Abdulrab, and L. Oubahssi, "Application Service Provider System : a new concept to provide interoperability between learning management systems", Proceedings of E-Learn 2005 World Conference (World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education), Vancouver, (Canada); Research/Technical Showcase, pp. 2763-2769, 2005.
 - [25] G.-G. Panigyrakis, P.-K. Theodoridis, and C.-A. Veloutsou, "All customers are not treated equally : Financial exclusion in isolated Greek islands", Journal of Financial Services Marketing, vol. 7, pp. 54–66, 2002.
 - [26] D. Porteous, "The enabling environment for mobile banking in Africa", DFID, Etats-Unis, DFID, 2006.
 - [27] E. Scornavacca and S.-J. Barnes, "M-banking services in Japan: a strategic perspective", International Journal of Mobile Communications, vol. 2, no. 1, pp. 51-66, 2004.
 - [28] K. Sodokin, M. Couchoro, and M. Koriko, "Information and communication technologies, artificial intelligence, and the fight against money laundering in Africa", Strat Change, May 2021.
 - [29] M. Suoranta, M. Mattila, and J. Munnukka, "Technology-based services: a study on the drivers and inhibitors of mobile banking", International Management and Decision Making, vol. 6, no. 1, pp. 33-46, 2005.
 - [30] G. Varshney and P. Madan, "A study of functionality Dilemma and Barriers to Optimal usage of M-Commerce", Curie Journal, vol. 3, no. 1, pp. 60-73, 2010.
 - [31] T.-T. Wei, G. Marthandan, A.-Y.-L. Chong, K.-B. Ooi, and S. Arumugam, "What drives Malaysian m-commerce adoption ? An empirical analysis", Industrial Management & Data Systems, vol. 109, no. 3, pp. 370-388, 2009.
 - [32] C. Weitenberner, M. Chapman, G. Miranda, S. Tobar, E. Wagoner, D. Akasaki, and K. Seal, "United States and South East Asian mobile markets : a comparative analysis of infrastructure and cultural differences", International Journal of Mobile Marketing, vol. 1, no. 1, pp. 73-88, 2006.

Real-Time Emotion Assessment System in Smart Classrooms Using Wearable Bracelets

Edgar Batista , Laia Cot, Valeria Pérez, Antoni Martínez-Ballesté 

Smart Technologies Research Group

Department of Computer Engineering and Mathematics

Universitat Rovira i Virgili

Tarragona, Spain

e-mail: {edgar.batista, antoni.martinez}@urv.cat

{laia.cot, valeria.perez}@estudiants.urv.cat

Abstract—Smart classrooms are the next frontier in education to accelerate and improve both teaching and learning processes. These environments integrate sensors, ubiquitous computing systems, Artificial Intelligence (AI) techniques, and high-speed data networks to create more cognitive, effective, interactive, and adaptive learning settings. Among other applications, students' emotions are key to maintain a positive learning atmosphere. Unlike traditional video analysis approaches, emotions can be detected through physiological data, such as electrodermal activity, heart rate, and skin temperature, which can be collected using non-invasive and user-friendly wearable devices. To this end, this paper presents a real-time emotion assessment system intended for both students and teachers within smart classrooms. A prototype has been developed using EmotiBit bracelets and a Raspberry Pi. Initial testing in laboratory settings has shown the system could smoothly run in off-the-shelf technology. However, the deployment of this system in real-world classrooms reveals several challenges that must be addressed, namely data volume, battery duration, data security concerns, and the lack of training datasets for the AI model.

Keywords—smart classroom; emotions; bracelets; wearables; EmotiBit; artificial intelligence; education.

I. INTRODUCTION

The adoption of the Information and Communication Technologies (ICT) paves the way for the creation of smart environments capable of collecting and analysing data from users and their context, facilitating the adjustment of processes in real-time and providing stakeholders with valuable insights. In the field of education, this technology can lead to significant advancements, as the concept of smart classrooms. Such environments prioritise students-centred learning, enhance teaching effectiveness, and relieve teachers from routine tasks, thereby reducing their stress and preventing burnout. To this end, smart classrooms leverage technology to enhance the quality of teaching and learning processes by collecting and processing information from students, teachers, and the environment, thus enabling adequate decision-making [1]. Smart classrooms can be implemented using several distributed, interconnected systems that autonomously gather data from embedded sensors, which are seamlessly integrated within the physical learning environment [2][3]. Among such systems, the *Classroom Agent* is responsible for collecting and analysing information from the individuals and their context using Artificial Intelligence (AI)

while providing stakeholders with visualisation and interaction tools.

Among the numerous opportunities within smart classrooms, understanding students' emotions can foster a positive emotional climate in the classroom that results in improved academic performance [4]. Affective states, encompassing a broad range of feelings, moods, and emotions, can be detected and monitored by the smart classroom to track the overall mood of students during specific activities or analyse which activities perform better at different times of the day. Various theoretical models have been proposed to define and categorise emotions. The most popular models are the Ekman's model of the six basic emotions, which identifies universal emotions, such as joy, sadness, fear, surprise, disgust, and anger [5], and the Russell's circumplex model, which organises emotions in a two-dimensional space based on valence (pleasant-unpleasant) and arousal (low-high) [6].

Most proposals for detecting emotions, moods, and feelings in smart classrooms rely on video analysis, conducted either in real-time or post-lecture. For instance, the studies in [7][8][9] proposed real-time emotion recognition systems, where the latter assigns numerical scores to students based on their concentration levels. However, video-based solutions might be limited by lighting conditions, occlusions, or even individual differences in expressing emotions. Moreover, implementing these solutions in real classrooms would require robust privacy-preserving methods to safeguard the privacy and confidentiality of underage students in accordance to ethical and legal standards. These constraints could be relaxed by detecting emotions using alternative data sources, such as wearables combined with air quality sensors and acoustic sensors. Specifically, recent technological advancements in physiological sensing have enabled novel methods for detecting and monitoring emotions through wearable technology. For instance, sensor-equipped wearables capable of measuring Electrodermal Activity (EDA), Heart Rate (HR), Heart Rate Variability (HRV), skin temperature (SKT), and electroencephalography (EEG) can help track students' emotional states in real-time [10][11][12]. With usability in mind, bracelets are generally accepted by users for detecting biological markers. However, the majority operate within a closed model paradigm, such as the research-oriented Empatica, Shimmer, and BIOPAC bracelets, requiring

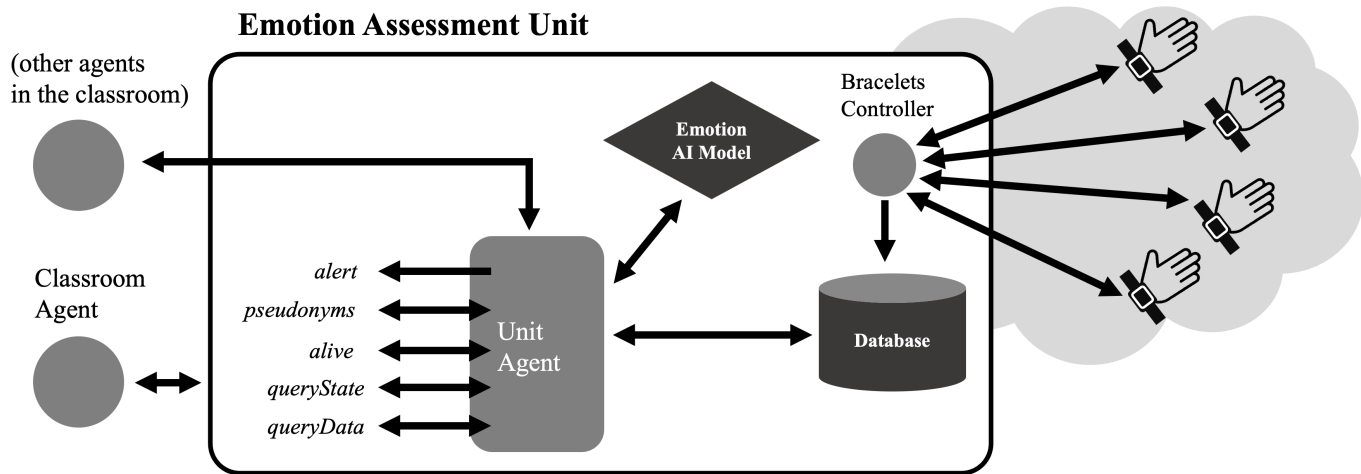


Figure 1. High-level diagram of the proposed Emotion Assessment Unit.

proprietary applications to collect and visualise the data. In contrast, open-source devices, such as EmotiBit [13][14], facilitate the wireless, real-time transmission of scientifically validated data from a variety of integrated sensors.

In this paper, we present a first approach to a real-time emotion assessment system for both students and teachers within smart classrooms. Our proposal involves real-time data collection gathered from EmotiBit devices and employs an AI model that might be used for emotion assessment. Additionally, this system is equipped with capabilities to interact with other distributed systems within the smart classroom ecosystem. The remainder of the paper is organised as follows: Section II describes the architecture of our proposal, Section III elaborates on the implementation and testing of the system, Section IV discusses the results and, finally, Section V concludes the article and provides further research lines.

II. OUR PROPOSAL

This section describes the design of the Emotion Assessment Unit, aiming to identify, track, and respond to the emotional states of the classroom population. The unit collects data from bracelets, which are then processed by an AI model. Moreover, the unit responds to request from the Classroom Agent, which can ultimately inform teachers in a timely manner. The architecture of our proposal is illustrated in Figure 1.

A. Components

Our system comprises the following components:

- **Bracelets Controller:** Responsible for interacting with all the bracelet devices within the smart classroom through a specific communication protocol.
- **Database:** Responsible for storing the data collected from the bracelet devices.
- **Emotion AI model:** This subsystem assesses the emotional states of the classroom population based on specific data, including physiological parameters, such as EDA, HR, and SKT, but also contextual data, such as ambient

noise and air temperature, obtained from other agents in the smart classroom. In this first approach, this component only addresses the variability of physiological parameters that indicate emotional changes.

- **Unit Agent:** Core component of the unit consisting of a software agent that interacts with the other components in the unit and with other agents in the smart classroom. It takes decisions about the emotional states of students and teachers, and responds to request from the Classroom Agent.

For the sake of brevity, we only focus on the Classroom Agent, which is considered the central component of the smart classroom. Other agents, responsible for monitoring other aspects of the smart classroom, such as the air quality or the noise, are also integral to this ecosystem but are not discussed.

B. Functionalities

Functionalities are categorised into three groups: (i) bracelet operational functionalities, (ii) Classroom Agent coordination functionalities, and (iii) management functionalities.

The first group, managed by the Bracelets Controller, ensures the proper management of wearables in the smart classroom:

- **Bracelets discovery:** Detection of all bracelet devices in the smart classroom network, *i.e.*, identifying each device's IP/MAC address.
- **Bracelet assignment:** Temporal assignment of each bracelet device to its wearer.
- **Start/Stop monitoring:** After assignment, the Bracelets Controller starts collecting data from the devices and stores them in the database until the functionality is deactivated.

The second group includes actions performed by the Unit Agent upon requests from the Classroom Agent:

- **alive:** Verifies whether the Emotion Assessment Unit is up and running.
- **pseudonyms(*pseudonyms*):** To protect individuals' privacy, the system uses pseudonyms instead of real identi-

fiers. This function updates the pseudonym list whenever the Classroom Agent generates a new set.

- **queryState(pseudonyms):** As the Classroom Agent may request the emotional states of students and teachers specified in the *pseudonyms* list, the Unit Agent retrieves the latest data from the database, sends them to the Emotion AI model, and returns the resulting emotional state assessments.
- **queryData(tsInitial, tsFinal):** As the Classroom Agent may retrieve data for further analysis, the Unit Agent queries the database to retrieve information stored between the specified timestamps.

The third group comprises proactive functionalities conducted by the Unit Agent to manage the system effectively:

- **Emotion update:** Periodical assessment of the emotional states of the smart classroom population using the AI model with the latest data stored in the database. Upon detecting significant changes (*e.g.*, increased number of stressed students), the Classroom Agent is alerted.
- **Battery alert:** Regular checking of the battery levels of the bracelet devices, alerting the Classroom Agent of low battery levels.
- **Database reduction:** To maintain optimal database performance, a maximum number of records in the database is limited. If it is exceeded, the oldest records are deleted. Note that the queryData functionality can be used if the Classroom Agent intends to store raw data values.

C. Emotion AI Model

In this first approach, our model processes an array of data representing a succession of physiological values for a specific parameter within a t -seconds timeframe (where t is a parameter). The model analyses these values to identify distinct patterns: *peaks* (characterised by a sudden increase followed by a decrease within a short period), *valleys* (a sudden decrease followed by an increase within a short period), *increases* (a steady rise over time), *decreases* (a steady decline over time), and *constant behaviour* (a stable state with no significant variations). Consequently, the model returns a 5-item array, each of them representing the significance of these patterns occurring in the input array.

III. DEVELOPMENT

This section elaborates on the development of a prototype of the proposed system to evaluate its feasibility in real-world classrooms.

A. Sensors

In our implementation, EmotiBit bracelets were employed for data collection (see Figure 2). These devices provide real-time physiological data, including EDA, HR, HRV, SKT, and oxygen saturation, among others, which have been scientifically validated [15]. Each EmotiBit is equipped with a battery and wireless connectivity capabilities. To manage the collected data effectively, a dedicated tool called Oscilloscope is provided to identify all EmotiBits within a network and select a specific

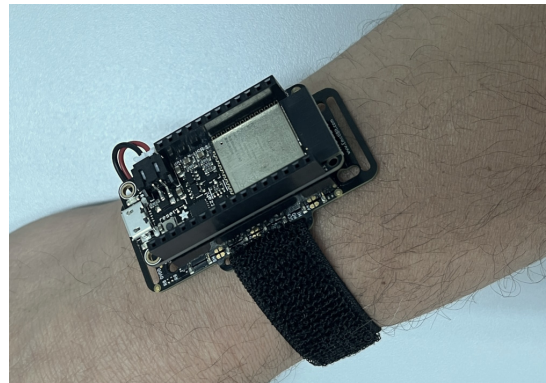


Figure 2. An EmotiBit sensor placed as a wearable bracelet.

device for data capture. Once selected, the Oscilloscope begins capturing the transmitted data in real-time and displays them graphically [16]. Unfortunately, since the Oscilloscope can only capture data from a single device at a time, and our goal is to process data from multiple devices, we opted to develop a software that emulates the Oscilloscope functionality, *i.e.*, the Bracelets Controller. This required analysing the open-source documentation and using Wireshark to understand the communication protocol between the EmotiBits and the Oscilloscope. Specifically, the EmotiBit ecosystem operates on three different network channels through Transmission Control Protocol (TCP) and User Datagram Protocol (UDP):

- **Advertising channel (via UDP):** This channel is used for device discovery through the so-called *Discovery Protocol*. In a nutshell, the Bracelets Controller broadcasts a ‘Hello EmotiBit’ message across the network and awaits ‘Hello Host’ responses from the devices. These responses contain a unique EmotiBit identifier for each device.
- **Data channel (via UDP):** This channel facilitates the transfer of data from the EmotiBit to the Bracelets Controller.
- **Control channel (via TCP):** This channel manages various operational aspects of the EmotiBit, such as initiating and terminating data exchange. To initiate data collection, the Bracelets Controller sends an ‘EmotiBit Connect’ message to each of the detected devices to specify the ports used for data collection.

The data messages sent by the EmotiBit contain the timestamp, the packet number, and the number of data points encapsulated in the payload. Indeed, each packet can contain up to sixteen different variables. Data acquisition occurs at different times depending on the variable: for instance, the sampling frequency is 15 Hz for EDA and 7.5 Hz for SKT [15]. Finally, the packets also contain the version and a reliability tag reserved for future updates.

B. Emotion Assessment Unit

All components of the Emotion Assessment Unit, namely the Bracelets Controller, the Unit Agent, the database, and the Emotion AI model, run on a Raspberry Pi 4 Model B



Figure 3. Prototype implementation of the solution with the Raspberry Pi, two EmotiBit bracelets, an access point, and the front-end web application.

board. The EmotiBits and the Emotion Assessment Unit are connected to the same network using a wireless LAN (WLAN) enabled by an access point. Hence, the Raspberry Pi board and each EmotiBit obtain their IP addresses via classical Dynamic Host Configuration Protocol (DHCP). Once connected, the Bracelets Controller can then initiate the Discovery Protocol to the broadcast address of the WLAN and awaits responses from the devices to launch the complete system. A picture of the prototype implementation is shown in Figure 3.

The Unit Agent, the Emotion AI Model, and the Bracelets Controller are implemented in Python. The latter uses lightweight process to handle unicast connections with each EmotiBit. Data are stored in an SQLite database to maintain a lightweight structure. All bracelet operational functionalities, explained in Section II-B, are managed using a local web front-end implemented in Flask, a Python web application framework. It is worth noting that, in a real smart classroom scenario, these functionalities would be managed through a front-end agent of the smart classroom. Similarly, the Classroom Agent's requests are handled via web services implemented with Flask. For instance, the list of pseudonyms is received through the *pseudonyms* function, and the association between EmotiBit identifiers and pseudonyms is stored in the database to ensure traceability.

Regarding data collection, the Bracelets Controller stores only the values for EDA, HR, and SKT in the database. Each value is stored as a new record in the database. Table I provides an example of data received in two consecutive packets: the first packet contained several data points, from which the Bracelets Controller retained EDA and HR, resulting in the insertion of two new records into the database, and the second packet contained a single data point on SKT, leading to the insertion of one new record.

C. Testing

Overall, the emotion assessment system works smoothly based on tests conducted in the laboratory with four EmotiBits simultaneously. Data are transmitted at the specific frequencies programmed in the sensors' firmware. The Bracelets Controller filters out values not related to emotions, and relevant records

are inserted into the database. Moreover, when the Classroom Agent requests data, the web service processes these requests and retrieves the required information from the database without any performance issues. This agent also receives responses from the Emotion AI model with no noticeable delay for queries performed at a rate of one per minute.

Regarding database size, we observed that collecting all 16 values transmitted by a single EmotiBit per hour results in approximately 25 MB of data. If only EDA, HR, and SKT are stored, the storage requirement decreases significantly to approximately 4 MB per hour. Considering the application of EmotiBits in a typical primary school in Catalonia, where the average student-to-teacher ratio is 20 students per class [17] and each student wears an EmotiBit device for 5 hours per day, the storage requirement for the entire class is around 2 GB per week. Given the vast amount of data, the database reduction functionality described in Section II-B is crucial for maintaining the efficiency and performance of the database. However, having real-time processing in mind, this operation should be requested by the Unit Agent when no data are being collected from the devices, for instance, post-lecture.

Regarding battery life, the EmotiBit device offers two operational modes that significantly impact its duration: (i) a normal mode, where the device operates at full capacity, using all its sensors and transmitting data wirelessly in real-time, and (ii) a low-power mode with no transmission but storage on an SD card. Hence, only the normal battery fits our requirements. Battery performance measurements were repeated three times to ensure reliability and consistency. In normal mode, the device's battery lasts approximately 3.25 hours, which can be impractical for a typical school day. In low-power mode, the battery life extends up to 9 hours.

IV. DISCUSSION

The implementation and testing of the proposed system works adequately in laboratory settings from a technological perspective. However, to fully deploy this solution in a real-world classroom, several aspects should be considered.

A. Computational Capabilities

The current prototype considers four EmotiBit devices, but real-world classroom deployments would necessitate a greater number. This demands managing and processing more data, so requiring increased computational capabilities. Currently, the system operates on a Raspberry Pi, which despite being a cost-effective computer, is constrained by limited resources. To address the demand for increased computational power, the Raspberry Pi could be seamlessly replaced by an Intel NUC,

TABLE I. SAMPLE OF DATA STORED IN THE DATABASE.

Timestamp	EmotiBit_Identifier	Data_Type	Value
2024-06-23 18:15:09	MD-V5-000014	EDA	0.030191
2024-06-23 18:15:09	MD-V5-000014	HR	72
2024-06-23 18:15:10	MD-V5-000014	SKT	36.232

a more powerful computer with a small form factor that could be integrated into a smart classroom.

B. Data Volume

The amount of data stored can grow rapidly unless adopting adequate implementation and maintenance processes. Two optimisations could be applied to reduce the data volume. First, it may not be necessary to insert a new record into the SQLite database for every data packet received by the Bracelets Controller. Instead, a more efficient approach could involve inserting new data entries only after a defined time interval has elapsed since the last insertion. Notwithstanding, the implementation of this enhancement is not trivial, as it must go in hand with the data requirements of the Emotion AI model. And second, data storage could be further optimised by adopting binary format storage. For instance, values like HR could be stored in one byte each, rather than allocating one byte per digit.

C. Battery Duration

Battery life is a significant concern when deploying our proposal in real-world classrooms. To address this issue, given the open-source nature of the EmotiBit platform, a viable strategy involves re-configuring the device's firmware to lower the sampling rate or deactivate sensors that are not essential to our objectives. This would effectively reduce the frequency of data transmission packets, thereby extending battery longevity.

D. Data Security

When analysing the network traffic between the EmotiBits and the Oscilloscope, it became evident that the payload of the packets is transmitted without any encryption. Hence, anyone (legitimate or intruder) within the WLAN could potentially access and analyse the transmitted data. More critically, device authentication is not considered. Hence, a malicious intruder could impersonate a fake EmotiBit simply by responding to the 'Hello EmotiBit' message during the Discovery Protocol, and generating data packets with fake values. Overall, there is significant room for improvement in this area to make the system robust against unauthorised access and potential data tampering [18].

E. Emotion AI Model

Assessing emotions and stress involves interpreting complex physiological and psychological signals. Besides the use of non-invasive devices like bracelets, electroencephalography and electrocardiograms can provide valuable data but they require invasive equipment. Moreover, audio analysis can contribute to emotion detection but it might require using individual microphones. In both cases, specialised equipment that is not practical for classroom environments is needed [19]. Although facial recognition could enhance the accuracy of emotion detection, several privacy and ethical concerns arise.

In our first approach to emotion assessment, we have addressed the detection of changes in physiological signals considering the most significant findings from the literature:

- An increase in EDA is related to positive valence and higher arousal and, hence, could be used as a preliminary indicator that students are receptive to learning [11].
- An increase in HR is related to momentary stress, which is an interesting state to monitor in both students and teachers [20][21].
- A decrease in SKT is related to higher arousal. Warmer SKT are hence associated to calm states [22][23].

In our prototype implementation, we have chosen to develop an AI model based on simple heuristics and algorithms. However, since the assessment of emotions is a complex issue, a more sophisticated AI model based on Machine Learning (ML) should be considered in the future. In addition to considering merely physiological parameters, our AI model is expected to also take contextual information into account (such as temperature, humidity, air quality, and noise level) and, to a certain extent, obtain complementary information from video sources.

Notwithstanding, in order to train the ML model, we need a dataset that relates our target predictions (*e.g.*, stress, valence, arousal) to a combination of values of EDA, SKT, air quality, noise level, time of the day, and type of learning activity performed, among others. We searched for databases on emotions that could be used to train a ML model. Despite the availability of databases [24]–[30], they include information from adult subjects (collected in controlled environments not related to learning) and, hence, their suitability to our purposes should be further studied.

V. CONCLUSION AND FUTURE WORK

Smart classrooms promote positive learning environments to enhance educational experiences. In this context, students' moods and emotions play a key role to achieve improved academic performance and information retention. In this paper, we have proposed an emotion assessment system able to collect and analyse emotions in real-time across the classroom population. The system builds upon wearable technology in the form of bracelets for physiological data collection, as well as AI models for emotion assessment. Our system, which can be seamlessly integrated into smart classrooms, works autonomously with minimal user intervention. The feasibility of our system has been successfully validated in laboratory settings, and its deployment in real-world classrooms will be processed once the identified challenges are addressed.

Future research in this area will concentrate on assessing our proposal in real-world settings using, at least, 20 EmotiBit devices to evaluate the Raspberry Pi's performance. Moreover, EmotiBit devices should be re-programmed to (i) incorporate a data security layer ensuring data confidentiality, integrity, and authentication, and (ii) optimise its firmware to increase battery life. Last but not least, the Emotion AI model should be further refined with comprehensive training datasets. Ideally, this model should be trained using data from different sources, namely physiological data (obtained from the bracelets), video footage (obtained from AI-equipped cameras) and contextual data (air quality, noise level...).

ACKNOWLEDGMENTS

This research has been supported by Ministerio de Ciencia, Innovación y Universidades, Gobierno de España (Agencia Estatal de Investigación, Fondo Europeo de Desarrollo Regional -FEDER-, European Union) under the research grant PID2021-127409OB-C33 CONDOR. Also, this research has been partially founded by AGAUR research group 2021SGR-00111: “ASCLEPIUS: Smart Technology for Smart Healthcare”.

REFERENCES

- [1] O. Unciti and R. Palau, “Teacher decision making tool: Development of a prototype to facilitate teacher decision making in the classroom”, *Journal of Technology and Science Education*, vol. 13, no. 3, pp. 740–760, 2023.
- [2] A. Martínez-Ballesté *et al.*, “A Proposal for the Smart Classroom Infrastructure using IoT and Artificial Intelligence”, in *48th IEEE Int. Conf. on Computers, Software, and Applications (COMPSAC)*, in press, 2024, pp. 1–6.
- [3] J. Machin, E. Batista, and A. Solanas, “Multi-agent Systems for Pervasive Electronics: A Case Study in School Classrooms”, in *Applications in Electronics Pervading Industry, Environment and Society (ApplePies)*, F. Bellotti *et al.*, Eds., ser. LNEE, Springer, vol. 1110, Cham, 2023, pp. 474–479.
- [4] J. Tan, J. Mao, Y. Jiang, and M. Gao, “The Influence of Academic Emotions on Learning Effects: A Systematic Review”, *Int. Journal of Environmental Research and Public Health*, vol. 18, no. 18, p. 9678, 2021.
- [5] P. Ekman, “Are There Basic Emotions?”, *Psychological Review*, vol. 99, no. 3, pp. 550–553, 1992.
- [6] J. A. Russell, “A Circumplex Model of Affect”, *Journal of Personality and Social Psychology*, vol. 39, no. 6, pp. 1161–1178, 1980.
- [7] E. Di Lascio, S. Gashi, and S. Santini, “Unobtrusive Assessment of Students’ Emotional Engagement during Lectures Using Electrodermal Activity Sensors”, *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)*, vol. 2, no. 3, pp. 1–21, 2018.
- [8] L. Li, C. P. Chen, L. Wang, K. Liang, and W. Bao, “Exploring Artificial Intelligence in Smart Education: Real-Time Classroom Behavior Analysis with Embedded Devices”, *Sustainability*, vol. 15, no. 10, p. 7940, 2023.
- [9] M. M. A. Parambil, L. Ali, F. Alnajjar, and M. Gochoo, “Smart Classroom: A Deep Learning Approach towards Attention Assessment through Class Behavior Detection”, in *Advances in Science and Engineering Technology Int. Conf. (ASET)*, Dubai, United Arab Emirates, 2022, pp. 1–6.
- [10] C. Llurba, G. Fretes, and R. Palau, “Pilot study of real-time Emotional Recognition technology for Secondary school students”, *Interaction Design and Architecture(s)*, vol. 52, pp. 61–80, 2022.
- [11] A. Horvers, N. Tombeng, T. Bosse, A. W. Lazonder, and I. Molenaar, “Detecting Emotions through Electrodermal Activity in Learning Contexts: A Systematic Review”, *Sensors*, vol. 21, no. 23, p. 7869, 2021.
- [12] W. Meng, Y. Cai, L. T. Yang, and W.-Y. Chiu, “Hybrid Emotion-aware Monitoring System based on Brainwaves for Internet of Medical Things”, *IEEE Internet of Things Journal*, vol. 8, no. 21, pp. 16014–16022, 2021.
- [13] S. M. Montgomery, N. Nair, P. Chen, and S. Dikker, “Introducing EmotiBit, an open-source multi-modal sensor for measuring research-grade physiological signals”, *Science Talks*, vol. 6, p. 100 181, 2023.
- [14] EmotiBit, *EmotiBit: Wearable biometric sensing for any project*, <https://www.emotibit.com>, [retrieved: August, 2024].
- [15] S. M. Montgomery, N. Nair, P. Chen, and S. Dikker, “Validating EmotiBit, an open-source multi-modal sensor for capturing research-grade physiological signals from anywhere on the body”, *Measurement: Sensors*, vol. 32, p. 101 075, 2024.
- [16] EmotiBit, *EmotiBit Documentation*, https://github.com/EmotiBit/EmotiBit_Docs, [retrieved: August, 2024].
- [17] Generalitat de Catalunya, *Education Ratios per Teaching Staff*, <https://educacio.gencat.cat/ca/departament/estadistiques/indicadors/sistema-educatiu/escolaritzacio/ratios/index.html>, in Catalan. [retrieved: August, 2024].
- [18] E. Batista, M. A. Moncusi, P. López-Aguilar, A. Martínez-Ballesté, and A. Solanas, “Sensors for Context-Aware Smart Healthcare: A Security Perspective”, *Sensors*, vol. 21, no. 20, p. 6886, 2021.
- [19] S. S. Tippannavar *et al.*, “Advances and Challenges in Human Emotion Recognition Systems: A Comprehensive Review”, *Journal of Trends in Computer Science and Smart Technology*, vol. 5, no. 4, pp. 367–387, 2023.
- [20] W. L. Sergio, V. Ströele, M. Dantas, R. Braga, and D. D. Macedo, “Enhancing well-being in modern education: A comprehensive eHealth proposal for managing stress and anxiety based on machine learning”, *Internet of Things*, vol. 25, p. 101 055, 2024.
- [21] N. Michels *et al.*, “Children’s heart rate variability as stress indicator: Association with reported stress and cortisol”, *Biological Psychology*, vol. 94, no. 2, pp. 433–440, 2013.
- [22] M. W. Park, C. J. Kim, M. Hwang, and E. C. Lee, “Individual Emotion Classification between Happiness and Sadness by Analyzing Photoplethysmography and Skin Temperature”, in *4th World Congress on Software Engineering*, IEEE, 2013, pp. 190–194.
- [23] R. Nakanishi and K. Imai-Matsumura, “Facial skin temperature decreases in infants with joyful expression”, *Infant Behavior and Development*, vol. 31, no. 1, pp. 137–144, 2008.
- [24] P. Schmidt, A. Reiss, R. Dürichen, C. Marberger, and K. Van Laerhoven, “Introducing WESAD, a Multimodal Dataset for Wearable Stress and Affect Detection”, in *Proc. 20th ACM Int. Conf. on Multimodal Interaction (ICMI)*, Boulder, USA, 2018, pp. 400–408.
- [25] S. Koldijk, M. Sappelli, S. Verberne, M. A. Neerincx, and W. Kraaij, “The SWELL Knowledge Work Dataset for Stress and User Modeling Research”, in *Proc. 16th Int. Conf. on Multimodal Interaction (ICMI)*, Istanbul, Turkey, 2014, pp. 291–298.
- [26] V. Markova, T. Ganchev, and K. Kalinkov, “CLAS: A Database for Cognitive Load, Affect and Stress Recognition”, in *Proc. Int. Conf. on Biomedical Innovations and Applications (BIA)*, Varna, Bulgaria, 2019, pp. 1–4.
- [27] K. Sharma, C. Castellini, E. L. van den Broek, A. Albu-Schäffer, and F. Schwenker, “A dataset of continuous affect annotations and physiological signals for emotion analysis”, *Scientific Data*, vol. 6, no. 1, p. 196, 2019.
- [28] S. Koelstra *et al.*, “DEAP: A Database for Emotion Analysis Using Physiological Signals”, *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 18–31, 2012.
- [29] J. A. Miranda-Correa, M. K. Abadi, N. Sebe, and I. Patras, “AMIGOS: A Dataset for Affect, Personality and Mood Research on Individuals and Groups”, *IEEE Transactions on Affective Computing*, vol. 12, no. 2, pp. 479–493, 2021.
- [30] M. Soleymani, J. Lichtenauer, T. Pun, and M. Pantic, “A Multimodal Database for Affect Recognition and Implicit Tagging”, *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 42–55, 2012.