ICN 2024

The Twenty-Third International Conference on Networks

ISBN: 978-1-68558-174-9

May 26 - 30, 2024

Barcelona, Spain

ICN 2024 Editors

Shintaro Mori, Fukuoka University, Japan

Jin-Shyan Lee, National Taipei University of Technology (Taipei Tech.), Taiwan

ICN 2024

Forward

The Twenty-Third International Conference on Networks (ICN 2024), held between May 26-30, 2024 in Barcelona, Spain, continued a series of events organized by and for academic, research and industrial partners.

We solicited both academic, research, and industrial contributions. We welcomed technical papers presenting research and practical results, position papers addressing the pros and cons of specific proposals, such as those being discussed in the standard fora or in industry consortia, survey papers addressing the key problems and solutions on any of the above topics short papers on work in progress, and panel proposals.

The conference had the following tracks:

- Communication
- Networking
- Advances in Software Defined Networking and Network Functions Virtualization
- Next generation networks (NGN) and network management
- Computation and networking
- Topics on Internet Censorship and Surveillance

We take here the opportunity to warmly thank all the members of the ICN 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 ICN 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 ICN 2024 organizing committee for their help in handling the logistics and for their work that made this professional meeting a success.

We hope that ICN 2024 was a successful international forum for the exchange of ideas and results between academia and industry and to promote further progress in the field of networks. We also hope that Barcelona provided a pleasant environment during the conference and everyone saved some time to enjoy the historic charm of the city.

ICN 2024 Chairs

ICN Steering Committee
Pascal Lorenz, University of Haute Alsace, France
Eugen Borcoci, National University of Science and Technology Politehnica Bucharest, Romania
Muath Obaidat, City University of New York, USA
Shintaro Mori, Fukuoka University, Japan

ICN Publicity Chairs
Laura Garcia, Universidad Politécnica de Cartagena, Spain
Sandra Viciano Tudela, Universitat Politecnica de Valencia, Spain

ICN 2024

Committee

ICN Steering Committee

Pascal Lorenz, University of Haute Alsace, France
Eugen Borcoci, National University of Science and Technology Politehnica Bucharest, Romania
Muath Obaidat, City University of New York, USA
Shintaro Mori, Fukuoka University, Japan

ICN 2024 Publicity Chairs

Laura Garcia, Universidad Politécnica de Cartagena, Spain
Sandra Viciano Tudela, Universitat Politecnica de Valencia, Spain

ICN 2024 Technical Program Committee

Luis F. Abanto-Leon, Technische Universität Darmstadt, Germany
Qammer H. Abbasi, University of Glasgow, UK
Khelil Abdelmajid, Landshut University of Applied Sciences, Germany
Alireza Abdollahpouri, University of Kurdistan, Sanandaj, Iran
Abdelmuttlib Ibrahim Abdalla Ahmed, University of Malaya, Malaysia
Ahmedin Mohammed Ahmed, FDRE Ministry of Innovation and Technology (MInT), Ethiopia
Francisco Airton Silva, Federal University of Piauí, Brazil
Sami Marzook Alesawi, King Abdulaziz University | Faculty of Computing and Information Technology at Rabigh, Saudi Arabia
Madyan Alsenwi, Kyung Hee University - Global Campus, South Korea
Reem Alshahrani, Kent State University, USA
Cristian Anghel, Politehnica University of Bucharest, Romania / Pentalog, France
Imran Shafique Ansari, University of Glasgow, Scotland, UK
Andrés Arcia-Moret, Xilinx, Cambridge, UK
Suayb S. Arslan, MEF University, Turkey
Mohammed A. Aseeri, King Abdulaziz City of Science and Technology (KACST), Kingdom of Saudi Arabia
Aishwarya Asesh, Adobe, USA
Michael Atighetchi, BBN Technologies, USA
Jocelyn Aubert, Luxembourg Institute of Science and Technology (LIST), Luxembourg
Marco Aurélio Spohn, Federal University of Fronteira Sul, Brazil
Omran Ayoub, Politecnico di Milano, Italy
Alvaro Barradas, University of Algarve, Portugal
Chemseddine Benkalfate, Quartz Laboratory of ENSEA, France
Luis Bernardo, NOVA University of Lisbon, Portugal
Robert Bestak, Czech Technical University in Prague, Czech Republic
Lucas Bondan, Research and Development Center in Information and Communication Technology (CTIC) of the Brazilian National Research and Educational Network (RNP), Brazil
Eugen Borcoci, National University of Science and Technology Politehnica Bucharest, Romania

Fernando Boronat Seguí, Universitat Politecnica de Valencia-Campus de Gandia, Spain
Radoslav Bortel, Czech Technical University in Prague, Czech Republic
Marilisa Botte, Federico II University of Naples, Italy
Christos Bouras, University of Patras, Greece
An Braeken, Vrije Universiteit Brussel, Belgium
Arslan Brömme, Vattenfall GmbH, Berlin, Germany
Claudia Canali, University of Modena and Reggio Emilia, Italy
Batyr Charyyev, Stevens Institute of Technology, USA
Aizaz Chaudhry, Carleton University, Canada
Hao Che, University of Texas at Arlington, USA
Marc Cheboldaeff, Cap Gemini, Germany
Adil Chekati, University of Tunis El Manar, Tunisia
Bo-Rong Chen, Google, USA
Jundong Chen, Dickinson State University, USA
Sixia Chen, Central Connecticut State University, USA
Yitao Chen, Qualcomm, USA
Yuxuan Chen, Shandong University, China
Andrzej Chydzinski, Silesian University of Technology, Poland
Jorge Crichigno, College of Engineering and Computing | University of South Carolina, USA
Gregorio D'Agostino, ENEA, Italy
Monireh Dabaghchian, George Mason University, USA
Sofiane Dahmane, University of Laghouat, Algeria
Abdulhalim Dandoush, ESME-Sudria engineering school, France
Susumu Date, Cybermedia Center - Osaka University, Japan
Babu R. Dawadi, Tribhuvan University,Nepal
Declan Delaney, University College Dublin, Ireland
Margot Deruyck, Ghent University - IMEC - WAVES, Belgium
Amir Djenna, University of Constantine, Algeria
Pengyuan Du, Facebook Inc., USA
Salahaldeen Duraibi, Jazan University, Saudi Arabia
Zakaria Abou El Houda, University of Montreal, Canada
Basem ElHalawany, Shenzhen University, China / Benha University, Egypt
Gledson Elias, Federal University of Paraíba (UFPB), Brazil
Levent Ertaul, California State University, East Bay, USA
Davide Ferraris, University of Malaga, Spain
Mário Ferreira, University of Aveiro, Portugal
Adriano Fiorese, Santa Catarina State University (UDESC), Brazil
Mathias Fischer, Universität Hamburg, Germany
Edelberto Franco Silva, Universidade Federal de Juiz de Fora, Brazil
Valerio Frascolla, Intel Deutschland GmbH, Neubiberg, Germany
Marco Furini, University of Modena and Reggio Emilia, Italy
Yu Gao, University of St. Thomas, USA
Yun Gao, Nanjing University of Posts and Telecommunications, China
Gourab Ghatak, IIIT-Delhi, India
Saptarshi Ghosh, London South Bank University, UK
Marco Giordani, University of Padova, Italy
Rita Girao-Silva, University of Coimbra & INESC Coimbra, Portugal
Srikrishna Gopu, Meta, USA

Shay Gueron, University of Haifa / Amazon Web Services, Israel
Tina Gui, Anheuser-Busch InBev, Belgium
Tibor Gyires, Illinois State University, USA
Nguyen Tri Hai, Chung-Ang University, Korea
Muhammad Hanif, Hanyang University / Seoul National University of Science and Technology, South Korea
Luoyao Hao, Columbia University, USA
Esteve J. Hassan, Mohawk College of Applied Arts and Technology, Canada
William "Chris" Headley, Virginia Tech National Security Institute | Virginia Polytechnic Institute & State University, USA
Enrique Hernández Orallo, Universidad Politécnica de Valencia, Spain
M. Reza Hoseinyfarahabady, University of Sydney, Australia
Md Shafaeat Hossain, Southern Connecticut State University, USA
Chiu-Han Hsiao, Academia Sinica, Taiwan
Wen-Chen Hu, University of North Dakota, USA
Fatima Hussain, Ryerson University / Royal Bank of Canada, Toronto, Canada
Dragos Ilie, Blekinge Institute of Technology (BTH), Sweden
Gal Itzhak, Technion - Israel Institute of Technology, Israel
Faouzi Jaidi, University of Carthage - Higher School of Communications of Tunis & National School of Engineers of Carthage, Tunisia
Yong Jin, Tokyo Institute of Technology, Japan
Omprakash Kaiwartya, Nottingham Trent University, UK
Faouzi Kamoun, ESPRIT School of Engineering, Tunisia
Kyungtae Kang, Hanyang University, Korea
Binayak Kar, National Taiwan University of Science and Technology, Taiwan
Erdem Karayer, Ege University, Turkey
Kallol Krishna Karmakar, University of Newcastle, Australia
Andrzej Kasprzak, Wrocław University of Science and Technology, Poland
Sokratis K. Katsikas, Norwegian University of Science and Technology, Norway
Hakima Khelifi, Beijing Institute of Technology, China
Pinar Kirci, Istanbul University-Cerrahpasa, Turkey
Sondes Ksibi, University of Carthage | Higher School of Communications of Tunis, Tunisia
Rafael Kunst, University of Vale do Rio dos Sinos (UNISINOS), Brazil
Christo Kurisummoottil-Thomas, Eurecom, France
Mohammed Laroui, Djillali Liabes University, SBA, Algeria & Paris University, France
Vincent Latzko, Technische Universität Dresden, Germany
Riccardo Lazzeretti, Sapienza University of Rome, Italy
Piotr Lechowicz, Wroclaw University of Science and Technology, Poland
Chi-Han Lee, Academia Sinica, Taiwan
Gyu Myoung Lee, Liverpool John Moores University, UK
Jonathan Lejeune, Sorbonne Université | Inria, France
Peilong Li, Elizabethtown College, USA
Kiho Lim, William Paterson University of New Jersey, USA
Lars Lindner, Universidad Autónoma de Baja California, Mexico
Yuchen Liu, North Carolina State University, USA
Jaime Lloret Mauri, Polytechnic University of Valencia, Spain
Pascal Lorenz, University of Haute Alsace, France
Chitradeep Majumdar, University of Liverpool, UK

Zoubir Mammeri, IRIT - Paul Sabatier University, Toulouse, France
D. Manivannan, University of Kentucky, USA
Christopher Mansour, Mercyhurst University, USA
Tagleorge Marques Silveira, Universidade de Aveiro, Portugal
Sreekar Marupaduga, IEEE, USA
Antonio Matencio-Escolar, University of the West of Scotland (UWS), UK
Manuel Mazzara, Innopolis University, Russia
Thijs Metsch, Intel Deutschland GmbH, Germany
Sonia Mettali Gammar, ISAMM ENSI, Tunisia
Umair Mohammad, Florida International University, USA
Ayan Mondal, Univ. Rennes | Inria | CNRS | IRISA, France
Jordi Mongay Batalla, Warsaw University of Technology, Poland
Mario Montagud, University of Valencia & i2CAT Foundation, Spain
Manuela Montangero, Università di Modena e Reggio Emilia, Italy
Marcos Morgenstern, Federal Institute of Education, Science and Technology Farroupilha (IFFar), Rio Grande do Sul, Brazil
Shintaro Mori, Fukuoka University, Japan
Ioannis Moscholios, University of Peloponnese, Greece
Susanna Mosleh, National Institute of Standard and Technology (NIST), USA
Mort Naraghi-Pour, Louisiana State University, USA
Galymzhan Nauryzbayev, Nazarbayev University, Kazakhstan
Hien Quoc Ngo, Queen's University Belfast, UK
Quang Ngoc Nguyen, Waseda University, Tokyo, Japan
Uyen Trang Nguyen, York University, Toronto, Canada
Maciej Nikodem, Wroclaw University of Science and Technology, Poland
Boubakr Nour, Beijing Institute of Technology, China
Muath Obaidat, City University of New York, USA
Olusola Odeyomi, Wichita State University, USA
Lidia Ogiela, AGH University of Science and Technology, Krakow, Poland
Marek R. Ogiela, AGH University of Science and Technology, Krakow, Poland
Urszula Ogiela, AGH University of Science and Technology, Krakow, Poland
Timothy O'Shea, Virginia Tech University & DeepSig Inc., USA
Constantin Paleologu, University Politehnica of Bucharest, Romania
Shashi Raj Pandey, Kyung Hee University - Global Campus, South Korea
Rahul Paropkari, Sprint, USA
Edoardo Persichetti, Florida Atlantic University, USA
Ferdous Pervej, North Carolina State University, Raleigh, USA
Vitaly Petrov, Nokia Bell Labs, Helsinki, Finland
Paulo Pinto, Universidade Nova de Lisboa, Portugal
Agnieszka Piotrowska, Silesian University of Technology, Poland
Ravi Prakash, TU Delft, Netherlands
Cong Pu, Marshall University, USA
Abdellatif Rahmoun, Ecole Supérieure en Informatique, Sid Bel-Abbes, ESI-SBA, Algeria
Shankar Raman, Indian Institute of Technology Madras, India
Kurdman Rasol, Universitat Politècnica de Catalunya (UPC), Spain
Adib Rastegarnia, Purdue University, USA
Claudina Rattaro, Universidad de la República, Montevideo, Uruguay
Danda B. Rawat, Howard University, USA

Yenumula B. Reddy, Grambling State University, USA
Ghaya Rekaya, Telecom Paris, France
Eric Renault, IMT-TSP, France
Filip Rezabek, Technical University of Munich, Germany
Farhad Rezazadeh, Centre Tecnològic de Telecomunicacions de Catalunya (CTTC), Spain
Imad Rida, University of Technology of Compiegne, France
Elisa Rojas, University of Alcala, Madrid, Spain
Gerardo Rubino, INRIA, Rennes, France
Rukhsana Ruby, Shenzhen University, China
Marina Ruggieri, University of Roma Tor Vergata, Italy
Abdulhakim Sabur, Arizona State University, USA
Amit Samanta, IIT Kharagpur, India / Max Planck Institute for Software Systems, Germany
Rodrigo Sanches Miani, Universidade Federal de Uberlândia, Brazil
Masahiro Sasabe, Graduate School of Science and Technology - Nara Institute of Science and Technology, Japan
Qi Shi, Liverpool John Moores University, UK
Yuankun Shi, Intel, China
Megumi Shibuya, The University of Electro-Communications, Japan
Kostas Stamos, University of Patras, Greece
Cristian Lucian Stanciu, University Politehnica of Bucharest, Romania
Prasad Talasila, Aarhus University, Denmark
Sudeep Tanwar, Institute of Technology | Nirma University, Ahmedabad, India
Giorgio Terracina, Università della Calabria, Italy
Vasileios Theodorou, Intracom Telecom, Greece
Eirini Eleni Tsiropoulou, University of New Mexico, USA
Abu Barkat Ullah, University of Canberra, Australia
Dalton C. G. Valadares, IFPE, Brazil
Rob van der Mei, Centre for Mathematics and Computer Science (CWI), Amsterdam, Netherlands
Costas Vassilakis, University of the Peloponnese, Greece
Quoc-Tuan Vien, Middlesex University, UK
César Viho, IRISA - ISTIC/Université Rennes 1, France
Calin Vladeanu, University Politehnica of Bucharest, Romania
Dmitriy Volkov, eQualit.ie, Canada
Xianzhi Wang, University of Technology Sydney, Australia
Bernd E. Wolfinger, University of Hamburg, Germany
Longfei Wu, Fayetteville State University, USA
Hong Yang, Nokia Bell Labs, Murray Hill, USA
Daqing Yun, Harrisburg University, USA
Habib Zaidi, Geneva University Hospital | Geneva University, Department of Radiology & Medical Informatics, Switzerland
Mariusz Żal, Poznan University of Technology, Poland
Pavol Zavarsky, Framatome, Canada
Sherali Zeadally, University of Kentucky, USA
Tengchan Zeng, Virginia Tech, Blacksburg, USA
Shengzhi Zhang, Boston University | MET College, USA
Shuai Zhang, Aalborg University, Denmark
Zhenghao Zhang, Florida State University, USA
Zhu Zhengyu, Zheng Zhou University, China

Taieb Znati, University of Pittsburgh, USA
Doukha Zouina, University of Science and Technology Houari-Boumediene (USTHB), Algeria

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 or 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 is 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

mmWave UAV-assisted Information-Centric Wireless Sensor Network for Disaster-Resilient Smart Cities: Preliminary Evaluation and Demonstration 1
Shintaro Mori

Automating SDN-ACLs with User Groups and Authentication Events 5
Florian Griesser, Atsushi Shinoda, Hirokazu Hasegawa, and Hajime Shimada

Application of a Deep Reinforcement Learning Algorithm to Virtual Machine Migration Control in Multi-Stage Information Processing Systems 13
Yukinobu Fukushima, Yuki Koujitani, Kazutoshi Nakane, Yuta Tarutani, Celimuge Wu, Yusheng Ji, Tokumi Yokohira, and Tutomu Murase

Measurement of the Per-flow Burst Ratio Parameter in IP Networks 19
Dominik Samociuk

Localization of Mobile Devices in Future Wireless Networks 26
Caleb Ludinga Lodi and Ronald Beaubrun

Applications of Computer Vision to Posture Corrections and Eye Disease Prevention 36
Fu-Yu Chen and Jin-Shyan Lee

mmWave UAV-assisted Information-Centric Wireless Sensor Network for Disaster-Resilient Smart Cities: Preliminary Evaluation and Demonstration

Shintaro Mori
Department of Electronics Engineering and Computer Science
Fukuoka University
8-19-1 Nanakuma, Jonan-ku, Fukuoka 814-0180, Japan
E-mail: smori@fukuoka-u.ac.jp

***Abstract*—This paper presents an information-centric wireless sensor network-based ecosystem for smart-city applications. The proposed scheme aims for an integrated non-terrestrial wireless network using unmanned aerial vehicles with higher frequency bands for future broadband wireless communication in disaster-resilient smart cities. To demonstrate the feasibility of the proposed scheme, we performed a preliminary evaluation in terms of network performance, including throughput and jitter in the application and TCP layers. In addition, as one of the scenarios to be applied for disaster-information sharing systems, we demonstrated a video-streaming test through an on-site experiment, which will explore a new wireless networking technology in promising mmWave bands.**

Keywords—millimeter-wave; unmanned aerial vehicle; information-centric wireless sensor networking

I. INTRODUCTION

Emerging technologies, such as the Internet of Things (IoT), metaverse, and artificial intelligence, enable crowd sensing in central city areas. Thanks to the massive amount of valuable information they provide, problems related to urbanization, social needs, and governmental structures can be mitigated. Smart cities are a new paradigm that can lead to the provision of smart services centered around healthcare, transportation, energy, and natural disasters. This makes cities greener, safer, and friendlier for residents [1]. To take advantage of one of the essential elemental technologies underpinning the social infrastructure, we propose Wireless Sensor Networking (WSN) technology for disaster-resistant smart cities. Such technologies usually provide daily services, but in our approach, disaster-related information is shared using the same system when a disaster occurs. This approach brings two advantages: economic efficiency (i.e., we can eliminate the necessity of an exclusive disaster-communication and networking system) and availability improvement (i.e., the system can be available in emergencies because it is already in place for performing daily operations). At the same time, we need to make sure the proposed scheme can provide a high data rate and low latency with stable connectivity to establish a new sustainable smart-city ecosystem.

Millimeter-Wave (mmWave) communications have been recognized as a revolutionary new research domain in future mobile networking technologies, which can support a wider bandwidth compared to current mainstream spectrums, such as ultra-high frequency and microwave bands. Due to the vast spectrum bandwidth, mmWave communication enables a multi-gigabit data transfer [2], and the spectrum is globally assigned (for example, in 28, 38, and 60 GHz in the cellular network utilized in the 3GPP-FR2 [3]). Therefore, mmWave communication is positioned at the forefront of the global frontier and is an essential element in any discussion on next-generation wireless communications. As a global standardized system, one of the most important technology is a Wireless Local Area Network (WLAN), which helps smartphone users connect to the local network. In contrast to other mmWave communication systems, such as local 5G or private 5G, the IEEE 802.11 family has the advantage of widespread user terminals, which yields economic benefits in common device usage in the phase of smart-city deployment. IEEE 802.11 ay is the latest version of mmWave communications and operates under the point-to-point and point-to-multi-point topologies in indoor and outdoor environments on the unlicensed 60-GHz bands. IEEE 802.11 ay has been specified to improve the legacy IEEE 802.11 ad while guaranteeing backward compatibility for legacy users.

IEEE 802.11 ay supports mesh networks, which can provide a cost-efficient broadband wireless solution to replace fiber optical networks in city areas. The IEEE 802.11 ay-compliant mesh network can be deployed using a combination of Distribution Nodes (DNs) and Client Nodes (CNs), i.e., multiple DNs are linked to each other to form a backhaul mesh network, and end-users can access the network via CNs. The mesh network is structured and works based on multi-hop communication and dynamic controls, such as finding the most efficient path for information en route, i.e., if one DN goes down, another can immediately take over its role, thereby improving the network's availability. As such, the mesh network has suitable features for a network that supports disaster-resistant smart cities. As a commercial product, Meta (Facebook) offer Terragraph (TG) as an IEEE 802.11 ay-compliant mesh network [4]. TG aims to provide an alternative low-cost solution for operators to provide a similar cellular network or regional internet service on the unlicensed 60-GHz band.

Natural disasters (earthquakes, typhoons, hurricanes, floods, and other geologic processes) can potentially cut or destroy the existing territorial wireless network infrastructure in a disaster area. In this situation, it is a serious challenge to provide quick and temporary alternative wireless connectivity, but one solution is to use Unmanned Aerial Vehicles (UAVs),

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

Copyright (c) IARIA, 2024. ISBN: 978-1-68558-174-9

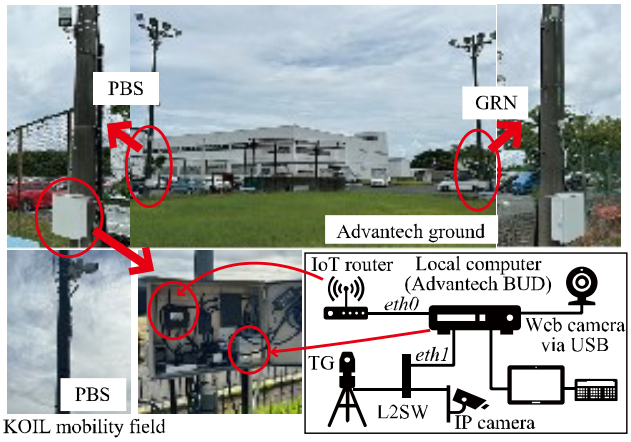

Figure 1. Overview of the test field developed in [12][13][14].

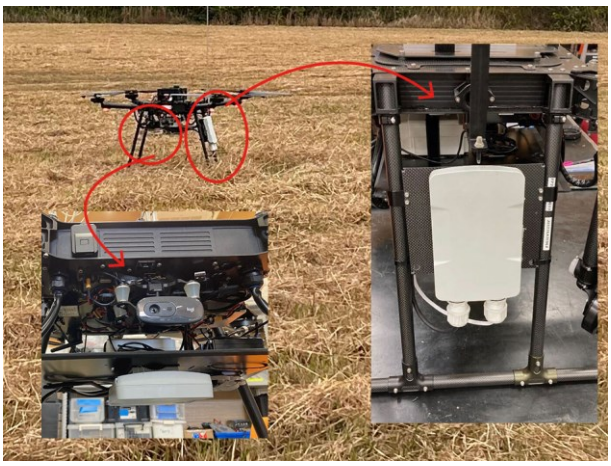

Figure 2. Overview of developed ARN device.

which improve the coverage by acting as aerial base stations [5][6]. Recall for a moment the characteristics of mmWave communications: the radio propagates straightforwardly, and therefore is significantly attenuated by penetration, atmosphere (oxygen), heavy rain, and moisture-containing material. Fortunately, UAVs can establish more reliable Line-of-Sight (LoS) links for ground end-users, which leads to a better communication channel. Therefore, UAV-assisted mmWave wireless networks can provide a broadband wireless network offering wide-area coverage even if a disaster strikes.

As we look at the network layer, the protocol suite must be designed based on an autonomous and decentralized network architecture. Information-Centric Networking (ICN) is a remarkable candidate for a future network architecture that shifts from host locations to content data [7]. This is beneficial in data-intensive applications optimized for content retrieval in an autonomous-decentralized ad-hoc network environment. ICN names the data instead of the address, i.e., the end-users can discover and obtain the data via names, and this naming-based retrieval achieves a location-free structure. Another vital feature of ICN is in-network caching, i.e., the data are copied and stored in the cache memories on the network nodes, which can be helpful for further data retrieval. In ICN systems, the data are handled separately by individual content units, i.e., the data can be self-certified and encrypted by their producer, which contributes to security improvement. Applying ICN to WSN, which yields Information-Centric Wireless Sensor Network (ICWSN) [8], positively affects network performance, such as boosting data delivery and improving data fetching delay. Therefore, ICWSN has the potential to solve the challenges arising from the case where most WSN devices are resource-constrained with radio frequency, processing resource, energy, and memory limitations. To provide information related to disasters through ad-hoc wireless networks, the data abstraction resulting from ICN design contributes to easy data spreading.

For mmWave UAV communications, Sanchez et al. [9] formulated a stochastic channel model for mmWave UAV communications under hovering conditions. Gapeyenko et al. [10] investigated the use of aerial relay nodes for dynamic routing to mitigate the effect of obstacles on the radio links. Masaoka et al. [11] investigated mmWave UAV-assisted communications for remotely operating and flying unmanned devices regardless of ground conditions, achieving high-speed data transmission. The use of mmWave bands is growing, and this study is positioned as a prior effort to them.

Consequently, this paper investigates an ecosystem to support application services for disaster-resilient smart cities. The proposed scheme is constructed based on the mmWave UAV-assisted ICWSN architecture. As part of our ongoing work [12][13][14], we have been developing a test field for ICWSN in the mmWave band. In this work, we describe the developed test field and aerial node using an industrial UAV. As a contribution of this paper, to illustrate the effectiveness of the proposed scheme, we evaluate network performance and the feasibility of the scheme. The demonstration includes a real-time video streaming application on the mmWave UAV-assisted ICWSN system as a scenario that shares a disaster-area information.

The remainder of this paper is organized as follows. Section II of this paper gives a brief overview of the development of the proposed ICWSN test field. Section III describes the proposed scheme. Section IV presents the evaluation results and discussion. Section V discusses related work. We conclude in Section VI with a brief summary and mention of future work.

II. DEVELOPMENT OF ICWSN TEST FIELD

We have been developing the testbed device and test field to evaluate a mmWave ICWSN framework [12][13][14], as shown in Figure 1. The test fields were constructed at the KOIL mobility field (Kashiwa, Chiba) and the baseball field in Advantech Japan (Nogata, Fukuoka). In this paper, we focus more on the baseball field because this is where we conducted the experiment. The framework is composed of a group of Sensor Nodes (SNs), Relay Nodes (RNs) (classified into Ground RN (GRN) and Aerial RN (ARN)), and a Private (self-operated) Base Station (PBS). We implemented the PBS and RN devices, and the control computer used an industrial

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

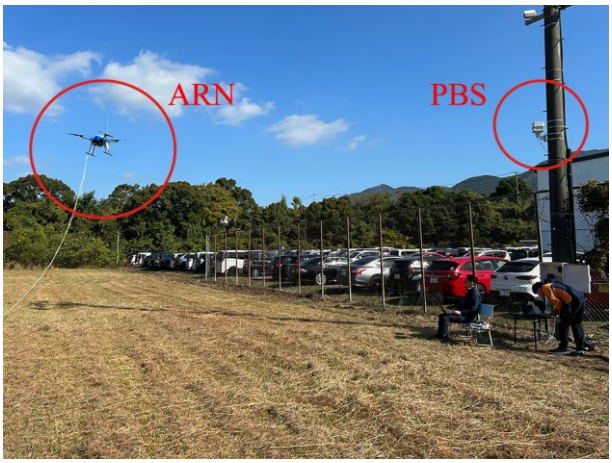

Figure 3. Field view of experimental site

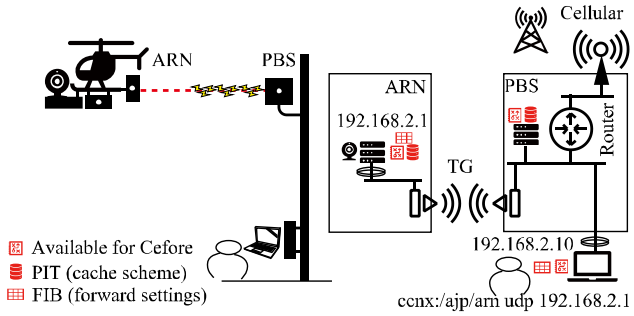

Figure 4. Network model of experimental site.

Advantech BUD (two-core 1.8 GHz Intel Atom CPU, 4 GB RAM, Ubuntu 20.04 OS, and extended processing and communications modules) device for reliability and robustness.

For the mmWave TG communication system, we used the BeMap MLTG-DN for PBS and MLTG-CN [15] for RN and SN. Note that MLTG-DNs can transmit up to distances of about 300 m, and all routes must be LoS with no foliage, walls, or other obstacles in the link. Their maximum transmission power, i.e., Effective Isotropic Radiated Power (EIRP), is 45 dBm, and the antenna consists of a phased array with 64 elements. In the beamforming method, the steering angle is [–45°, 45°] in the azimuth plane and [–25°, 25°] in the elevation plane, and it selects an index of direction among predefined beams. The MLTG-DN and MLTG-CN support the SC-PHY mode with the adaptive rate control of 1–12 in IEEE 802.11 ad/ay and have four channels, consisting of 58.32, 60.48, 62.64, and 64.80 GHz (central) frequency bands with 2.16-GHz bandwidth.

III. Development of ARN Device

The ARN device consists of a control computer, camera, and MLTG-CN mounted on the UAV, as shown in Figure 2. The control computer used the BUD device, the same as the previously mentioned testbed device. The camera and MLTG-CN were connected to the computer via the Universal Serial Bus (USB) and Ethernet (wired LAN) cables, respectively. As shown in Figure 3, due to Japan's Radio Act and Civil

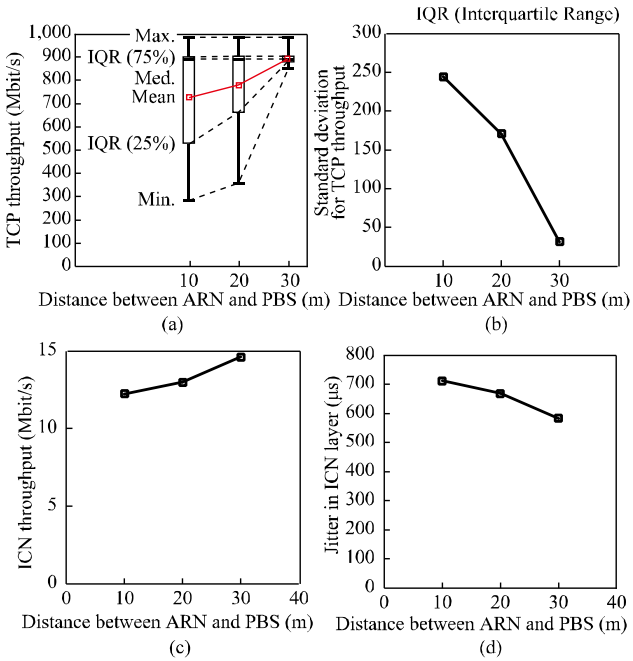

Figure 5. Experimental results of network performance: (a) TCP throughput. (b) Standard deviation for TCP throughput. (c) ICN throughput. (d) Jitter versus distance between nodes.

Aeronautics Act regulations, the UAV flew with a captive flight (not free). Note that the mooring rope is not only used to anchor the UAV to the ground but is also bundled with a LAN cable in parallel for power supply to the MLTG-CN via PoE. Figure 4 shows the network model of the experiment. As an end-user terminal, a PC (two-core 1.3 GHz Intel Core i5U CPU, 8 GB RAM, and Ubuntu 20.04 OS) was directly connected to the PBS, and the static IP addresses were assigned for ARN and the PC. The ICN platform used Cefore [16], which is a ccnx-compliant protocol stack. Note that we only install Cefore in the control computer of ARN and PC; the data can be exchanged via the "cefnetd" and "csmgrd" daemon processes from the application program. Namely, the system can perform based on the ccnx-based procedure, including naming, caching, and data management.

IV. Experimental Results

Let d denote the distance between ARN and PBS. To establish the link between two, UAVs hovered at the location where d = 10 m and at the height of 5 m, the same as that of PBS, and the antenna surfaces between SN and ARN and between ARN and PBS facing each other. Under this condition, the TG link can be reconstructed, including the beamforming direction, by restarting MLTG-CN (on the ARN). Note that this procedure can be accomplished by restarting MLTG-DN (on the PBS), but it takes more time to reconfigure than when using MLTG-CN. Figure 5 shows the results of network performance. TCP throughput was measured every 1 s for 30 s using iPerf3. ICN throughput was calculated based on the time intervals when the data provider commits static data using the "cefputfile" command from the control application, and then the receiver retrieves the data

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

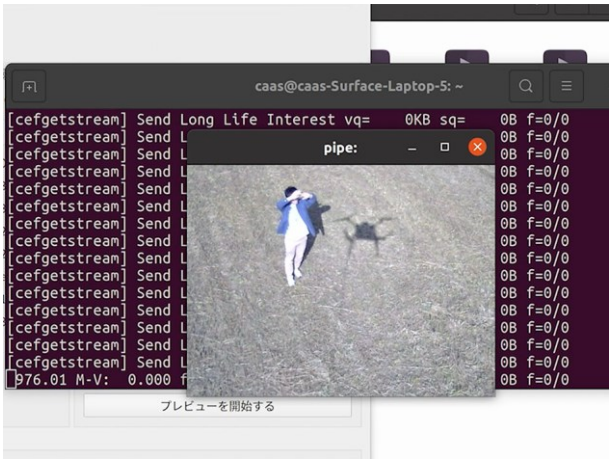

Figure 6. Demonstration of video streaming application.

using the "cefgetfile" command. The ICN throughput was the mean value of the three measures for three different file fetches.

As shown in Figure 5(a), the average TCP throughput is 891 Mbit/s (in median value) and 735, 787, and 899 Mbit/s (in mean value) in the cases where d = 10, 20, and 30 m, respectively. Note that, in the physical layer, the TG can support the data transfer rate up to 1,925–4,620 Mbit/s; nevertheless, MLTG only supports Gigabit Ethernet (GbE); so this wired interface causes a bottleneck. Figure 5(b) shows the variance of TCP throughput. The standard deviation decreases when d increases because the UAV moves vertically and horizontally (including roll and pitch), even if it is stably hovering in a fixed position. This movement affects the mmWave feature (i.e., straight radio propagation and directional beamforming), which can be relatively small for far distances of d. As shown in Figures 5(c) and (d), the average ICN throughputs are 12.2, 13.0, and 14.6 Mbit/s, and the average jitters are 712, 669, and 583 μs for d = 10, 20, and 30 m, respectively. These results have the same characteristics as that of TCP evaluations in Figures 5(a) and (b). The ICN throughput is much lower than that of TCP because the latency causing mmWave propagations will affect the ICN layer, and Cefore cannot optimally work, which is for wired LANs.

To demonstrate the provision of information on the disaster-stricken area, the ARN performed live video broadcasting from the sky to the PC (connected to ground PBS). On the basis of the literature [17], the ARN provided the streaming video from the camera mounted on the UAV using the "cefputstream" command, and the PC received it using the command of "cefgetstream" command. Figure 6 shows a screenshot of the PC during the demonstration, where it is clear that the streamed video can be received, although the static photo cannot represent its motion.

V. CONCLUSION

In this paper, we evaluated the network performance and demonstrated a high-capacity application of the video-streaming application. We can obtain the fundamental performance and show the scheme's feasibility. In future work, we plan to deploy the proposed eco-system in an actual city and we should construct stable mmWave-band networks.

ACKNOWLEDGMENT

This work was partly supported by NICT Japan, Grant No. JPJ012368C05601. We are grateful to Dr. Kenji Kanai for helpful discussions, and to Advantech Japan, BeMap, Haft, Panasonic, and TEAD for their help with the experiments.

REFERENCES

[1] P. Mishra and G. Singh, "6G-IoT framework for sustainable smart city: Vision and challenges," *IEEE Consumer Elect. Mag.*, pp. 1–8, Aug. 2023.

[2] K. Aldubaikhy, W. Wu, N. Zhang, N. Cheng, and X. Shen, "mmWave IEEE 802.11ay for 5G fixed wireless access," *IEEE Wireless Commun.*, vol. 27, no. 2, pp. 88–95, Apr. 2020.

[3] M. Cudak, A. Ghosh, A. Ghosh, and J. Andrews, "Integrated access and backhaul: A key enabler for 5G millimeter-wave deployments," *IEEE Commun. Mag.*, vol. 59, no. 4, pp. 88–94, Apr. 2021.

[4] A. Nordrum, "Facebook pushes networking tech: The company's Terragraph technology will soon be available in commercial gear," *IEEE Spectrum*, vol. 56, no. 4, pp. 8–9, Apr. 2019.

[5] Q. T. Do, D. S. Lakew, A. T. Tran, D. T. Hua, and S. Cho, "A review on recent approaches in mmWave UAV-aided communication networks and open issues," *Proc. ICOIN 2023*, Bangkok, Thailand, Feb. 2023, pp. 728–731, doi: 10.1109/ICOIN56518.2023.10049043.

[6] M. T. Dabiri, M. Hasna, N. Zorba, T. Khattab, and K. A. Qaraqe, "Enabling long mmWave aerial backhaul links via fixed-wing UAVs: Performance and design," *IEEE Trans. Commun.*, vol. 71, no. 10, pp. 6146–6161, Oct. 2023.

[7] B. Ahlgren, C. Dannewitz, C. Imbrenda, D. Kutscher, and B. Ohlman, "A survey of information-centric networking," *IEEE Commun. Mag.*, vol. 50, no. 7, pp. 26–36, July 2012.

[8] L. C. M. Hurali and A. P. Patil, "Application areas of information-centric networking: State-of-the-art and challenges," *IEEE Access*, vol. 10, pp. 122431–122446, Nov. 2022.

[9] S. G. Sanchez, S. Mohanti, D. Jaisinghani, and K. R. Chowdhury, "Millimeter-wave base stations in the sky: An experimental study of UAV-to-ground communications," *IEEE Trans. Mobile Comp.*, vol. 21, no. 2, pp. 644–662, Feb. 2022.

[10] M. Gapeyenko, et al., "Flexible and reliable UAV-assisted backhaul operation in 5G mmWave cellular networks," *IEEE J. Sel. Areas in Commun.*, vol. 36, no. 11, pp. 2486–2496, Nov. 2018.

[11] R. Masaoka and G. K. Tran, "Construction and demonstration of access link for millimeter wave UAV base station network," *Proc. ICUFN 2023*, Paris, France, Jul., 2023, pp. 163–167, doi: 10.1109/ICUFN57995.2023.10200815.

[12] S. Mori, "Information-centric wireless sensor networks for smart-city-as-a service: Concept proposal, testbed development, and fundamental evaluation," *Proc. IEEE CCNC 2023*, Las Vegas, NV, USA, Jan. 2023, pp. 945–946, doi: 10.1109/CCNC51644.2023.10060577.

[13] S. Mori, "Test-field development for ICWSNs and preliminary evaluation for mmWave-band wireless communications," *Proc IEEE CCNC 2024*, Las Vegas, NV, USA, Jan. 2024, pp. 1–2, doi: 10.1109/CCNC51664.2024.10454799.

[14] S. Mori, "Energy-efficient cooperative caching scheme for green ICWSN: Preliminary analysis and testbed development," *Proc. ACM MobiCom WS NET4us 2023*, Madrid, Spain, Oct. 2023, pp. 207–212, doi: 10.1145/3615991.3616406.

[15] BeMap, https://www.bemap.co.jp/ (retrieved: Apr. 2024).

[16] Cefore, https://cefore.net/ (retrieved: Apr. 2024).

[17] K. Matsuzono and H. Asaeda, "NMRTS: Content name-based mobile real-time streaming," *IEEE Commun. Mag.*, vol. 54, no. 8, pp. 92–98, Aug. 2016.

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

Automating SDN-ACLs with User Groups and Authentication Events

Florian Grießer*, Atsushi Shinoda†, Hirokazu Hasegawa‡, Hajime Shimada†
* *School of Computation, Information and Technology, Technical University Munich, Munich, Germany*
†*Graduate School of Informatics, Nagoya University, Nagoya, Japan*
‡*Center for Strategic Cyber Resilience R&D, National Institute of Informatics, Tokyo, Japan*
§*Information Technology Center, Nagoya University, Nagoya, Japan*
florian.griesser@tum.de, shinoda@net.itc.nagoya-u.ac.jp, hasegawa@nii.ac.jp, shimada@itc.nagoya-u.ac.jp

***Abstract*—Due to emerging cybersecurity threats, traditional networks struggle to adapt to new challenges because of their static nature and need for manual adjustments. In contrast, the inherent flexibility and rapid adaptability of Software-defined Networks (SDN) present an opportunity to overcome these limitations. Leveraging this potential, we propose a novel approach for automatically generating Access Control Lists (ACLs) within SDN environments. The system centralizes Access Control to the User Database and automatically generates derived rules, thus reducing administrators' manual work. By implementing Port Access Control, we can ensure that only authentic clients can access network resources. As a second feature, the system can change ACLs to block traffic or forward traffic to an Intrusion Detection System (IDS) for deeper inspection in case of suspicious activity like failed login attempts. To demonstrate the effectiveness, we evaluated the system in two use cases, initial client connection and dynamic adaption to authentication events, to test and compare the implementation to other systems. The evaluation proved that we can reduce manual processes and enhance the security of a network by dynamically generating ACLs to isolate clients.**

***Index Terms*—Software-defined Networking; Authentication; Access Control Lists**

I. Introduction

Digital transformation has exponentially increased the complexity of network architectures, presenting unprecedented challenges in maintaining robust security frameworks. In this ever-evolving digital landscape, cybersecurity threats have become more sophisticated, leveraging the linkage of modern infrastructures to exploit vulnerabilities at an alarming rate. Traditional network security mechanisms, which mainly rely on static configurations and manual oversight, are increasingly proving inadequate against this backdrop of dynamic and evolving threats [1]. The inherent limitations of these conventional approaches, characterized by their inflexibility and slow response times, underline the urgent need for more adaptable, responsive security measures.

Software-Defined Networking is a revolutionary paradigm that promises to redefine network management and security [2]. At its core, SDN separates the network's control logic from the underlying hardware, facilitating a centralized and programmable framework that transcends traditional hardware limitations [3]. This separation enhances network flexibility and management and introduces dynamism and adaptability, which were unachievable with conventional network architectures until now. According to a report by Global Market Insights, the SDN Market, valued at USD 28.2 billion in 2023, is expected to experience significant growth, with a projected expansion rate exceeding 17% annually from 2024 to 2032 [4]. Through SDN's capabilities, networks gain the flexibility to adapt swiftly to evolving security demands. This flexibility enables the immediate implementation of tailored security measures and configurations to effectively counter new threats, as illustrated in the study by Ali et al. [5].

Furthermore, our contribution is complemented by the work of Yakasai et al. in FlowIdentity, which advances virtualized network access control within SDN through a role-based firewall [6]. We also draw upon the architectural insights of Casado et al. in Ethane, demonstrating the power of centralized policy enforcement [7], and the innovative approach of Mattos et al. in AuthFlow, focusing on authentication and access control mechanisms in SDN environments [8].

Furthermore, this approach was refined by incorporating a sophisticated analysis of authentication logs, drawing upon the work of Xing et al. in SnortFlow, which explores an OpenFlow-based intrusion prevention system in cloud environments [9], and the study by Le et al. on a flexible network-based intrusion detection and prevention system on Software-Defined Networks [10].

The paper progresses from reviewing related SDN security work in Section II to foundational concepts in Section III. Section IV describes our system for automating ACLs, while Section V covers its implementation. Section VI evaluates the system's performance against existing methods, and Section VII encapsulates concluding thoughts and future directions, rounding off our discussion.

II. Related Work

Network security and access control advancements are crucial in the evolving landscape of SDN. The following studies demonstrate that emerging technologies and frameworks are pivotal in addressing these challenges.

A. Intrusion Detection with Authentication Events

In the study by Chu et al. [11], "ALERT-ID," an intrusion detection system for large-scale network infrastructures, is presented. The system distinguishes between normal operations and potential security threats through real-time analysis of authentication, authorization, and accounting (AAA) system logs. It employs behavioral models built on historical access patterns and user profiles, efficiently identifying potential intrusions and misuse. Notably, ALERT-ID balances the need

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

for thorough security monitoring with a manageable false alarm rate, demonstrating the importance of dynamic security measures in complex network environments.

B. Dynamic Access Control in SDN

Transitioning to dynamic access control, the work by Nayak et al. introduces "Resonance: Dynamic Access Control for Enterprise Networks" [12]. Resonance implements dynamic security policies with a registration phase, complemented by real-time monitoring and inference mechanisms specified by administrator rules.

Further extending the concept of network security, the study by Martins et al. [13] introduces an innovative access control architecture for SDN, leveraging the ITU X.812 standard. This framework incorporates Role-Based Access Control (RBAC) with traffic prioritization rules, significantly advancing towards more granular and role-specific access control mechanisms in network environments. A notable feature of this architecture is its reliance on predefined rules, which are carefully established and mapped to user roles by administrators. This approach requires administrators to proactively define comprehensive security and access parameters, ensuring a tailored access control environment but also necessitating substantial initial setup and configuration efforts.

These studies highlight a significant progression in network security approaches, evolving from intrusion detection systems to dynamic and sophisticated role-based access control models. Our forthcoming work aims to advance the field by introducing a new system that significantly enhances these foundational methodologies and requires less administrative work.

III. Preliminary Concepts

This section introduces foundational concepts relevant to network management and security, including OpenFlow switches in SDN, 802.1X Port-Based Network Access Control, access management with Active Directory and LDAP Authentication, and the Extensible Authentication Protocol over LAN (EAPOL).

A. The Role of OpenFlow Switches in SDN

SDN represents a paradigm shift in network management, with OpenFlow switches being a cornerstone of this architecture. These programmable switches enable dynamic network control, efficient routing, and robust access control mechanisms [14].

Despite their advantages, OpenFlow switches introduce security challenges, such as controller security and flow table vulnerabilities [15].

B. Strengthening Network Defenses with 802.1X Port-Based Network Access Control

IEEE 802.1X Port-Based Network Access Control significantly strengthens network security by implementing stringent access control measures. It restricts network entry to verified devices and users, ensuring high network integrity and protection [16].

The interaction follows a precise sequence: initiation, identity presentation, and authentication verification, utilizing Extensible Authentication Protocol (EAP) over LAN [17]. This structured process confirms the identity of the devices and users and mitigates potential replay and impersonation threats.

In summary, 802.1X Port-Based Network Access Control is a foundational network security pillar, effectively managing access through rigorous authentication and encryption practices [17].

C. Access Management with Active Directory and LDAP Authentication

Active Directory (AD) [18] and Lightweight Directory Access Protocol (LDAP) [19] Authentication play vital roles in access management within network environments. AD streamlines user group and role management in Windows networks, while LDAP provides a unified authentication framework across multiple services.

A centralized user database with Access Rights and user Groups and Roles is employed in a typical company network, often based on Active Directory. This centralized management offers scalability, ease of administration, and seamless integration, making it indispensable for effective access control and user management.

IV. Proposed System

This section presents a comprehensive overview of our proposed system, designed to significantly enhance network security and efficiency through advanced ACL management and authentication mechanisms.

A. Previous works Problem and Approach

Building on our established framework for automating ACL generation through statistical analysis of communication patterns, this work seeks to leverage further and enhance the existing infrastructure. Our initial efforts in [20] laid the foundational groundwork for this approach, which saw significant development and refinement in subsequent studies, such as [21]. A primary challenge identified in our exploration was addressing the need for authentication proof for IP addresses.

We have refined our approach to take advantage of a typical user database, like Active Directory [18], a standard part of a company network. This centralized database offers a significant advantage because user and group management and their corresponding resource access permissions are already handled there. We aim to leverage this existing infrastructure to streamline the process and eliminate redundant tasks for administrators.

B. Architecture of the proposed system

In the proposed system, we enhance network security through Port Access Control, which limits network port access exclusively to authenticated users. This approach is grounded in a security model where each port is individually secured and requires authentication before granting access. As a result, each user undergoes an authentication process, enhancing

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

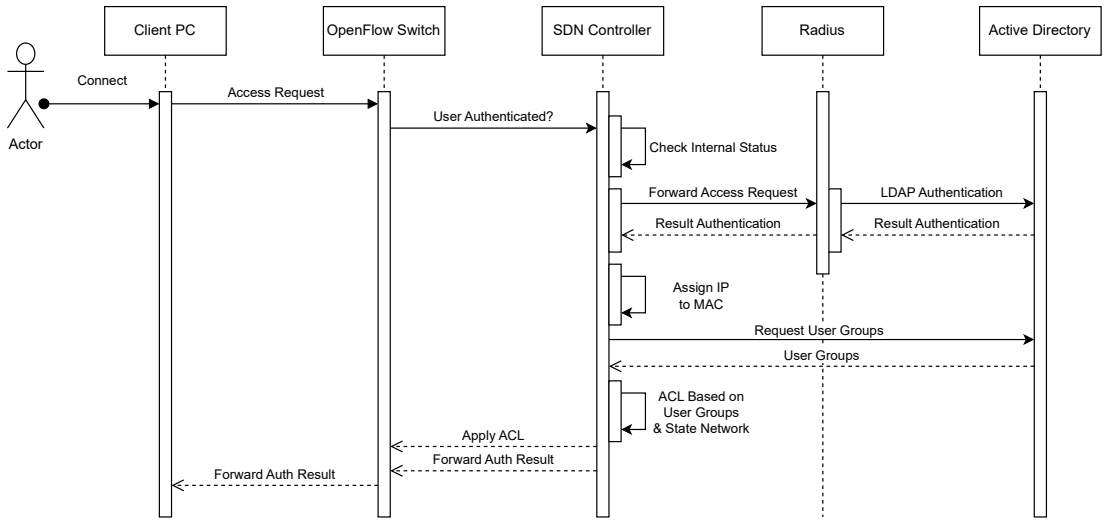

Figure 1. Dynamic ACL Adaption based on Authentication Events

the network's overall security posture. The process for user connection is designed with precision to ensure a secure and efficient authentication mechanism and can be found in Figure 1.

Initially, the system is configured only to allow EAPOL messages, which are then directed to an authenticator component. This step ensures that there is no communication before the client authenticates.

The SDN Controller checks its internal state for preconfigured users based on MAC and IP addresses. This information is crucial for comparing against new data received during authentication. Authentication messages for EAPOL are forwarded to a Radius server, which validates the credentials against a common user database, typically an active directory.

Upon successful authentication, the system assigns an IP address to the specific MAC address by inserting a record in a DHCP server. This procedure ensures that the assigned IP address corresponds to a specific MAC address and is associated with a specific port. The system then generates User-Specific Access Control Lists tied to a particular port by requesting user groups for the specific username from the Active Directory via LDAP. The fundamental idea is that users in the same group as a specific server should also have access to that server. For instance, if the user 'Ben' is a member of the 'Mail' group, to which our Mailserver also belongs, we understand that we need to create ACLs that permit this specific traffic. Consequently, we establish a whitelist to allow this connection and block all other traffic.

The system can identify users labeled as servers, which differ from standard clients, through a unique identifier group assigned explicitly to servers. The same concept would also be possible with a specific role depending on the existing usage in a company network, but in our case, we focused on groups.

Thus, any user belonging to this shared group is required to be able to establish a connection with the designated server. A port scan is conducted for servers to identify open ports and protocols. This information is linked to the user group of the server. For clients, the system constructs ACLs based on user groups and existing database information about servers, ensuring only communication between the user's MAC & IP address and the server's IP address and port. Since, for the system, only the port is necessary, it is also possible to apply this to multiple SDN switches since the ACL is only bound to a port on a switch.

Administrators can create templates for specific scenarios, such as restricting SSH access to only administrators. This template can be created by adding a template for port 22 associated with, for example, the user group "Administrator". These templates are scanned before actual ACL generation, with a higher priority than user roles and ports discovered during the generation process. Additionally, templates are essential for managing Internet traffic, with administrators defining traffic routing rules that cannot be determined using available information.

Since the update of active users and the checking for new ports on the servers occur periodically, once a day, and can be adjusted by an administrator, the system maintains minimal applied ACLs and removes unused ones if a host is no longer connected, thereby leading to higher efficiency [22].

The system also accounts for dynamic authentication events, using an LDAP proxy to monitor authentication activities. This monitoring detects suspicious activities based on failed login events for a specific service. The key idea is that if there are a certain number of failed login attempts, we aim to modify the network layout and ACLs for the particular user being observed, as their actions may be considered suspicious. Therefore, we can also redirect traffic to an Intrusion Detection System to look further since this traffic is probably suspicious or block all traffic or certain parts by adapting the ACLs dynamically. Alert-ID inspires this approach [11] with the extension that we do direct changes in the network when we encounter malicious behavior.

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

This comprehensive approach to Port Access Control, supported by a robust authentication framework, ensures a secure and efficient network access management system, safeguarding the integrity and security of the network infrastructure. To further enhance the effectiveness of our proposed system, we incorporate a quantitative analysis of Access Control Lists within the network when these settings are applied.

C. Quantitative Analysis of Access Control Lists

We rely on a quantitative understanding of ACLs within the network to understand the system's complexity. The ACL count is determined based on user, group, and port configurations, providing insights into the scale and complexity of the access control mechanisms.

- **Number of ACLs per User** (N_u): This metric quantifies the number of ACL entries associated with each user. It is calculated by summing the ports across all groups a user belongs to, given by

$$N_{u_i} = \sum_{g=1}^{G_i} P_{g_i} \tag{1}$$

where P_{g_i} is the number of ports for group g for user i and G_i is the total number of groups for user i.

- **Global Number of ACLs** (N_{global}): The total number of ACLs across the network reflects the overall complexity of access control. It is computed as

$$N_{global} = \sum_{i=1}^{U} N_{u_i} \tag{2}$$

where U represents the total number of users, and N_{u_i} is the number of ACLs for user i.

- **Number of ACLs per Switch** (N_s): Understanding the ACL count for each switch helps in optimizing access control at the local level. This metric is determined by

$$N_s = \sum_{i \in S} N_{u_i} \tag{3}$$

where S is the set of users connected to the switch, and N_{u_i} represents the number of ACLs for user i in the set S.

The subsequent Sections will examine the implementation details, demonstrating its capabilities and effectiveness.

V. Implementation

Following the conceptual framework outlined in Section IV, the practical implementation of the Port Access Control system integrated various components. The design in Figure 2 presents how different parts work together.

We chose the Faucet SDN Controller [23] for our prototype, which uses Ryu [24] in the backend and Gauge to view events on the switch. It has the considerable advantage that the rules are defined in YAML (YAML Ain't Markup Language). One significant advantage of this architecture is that these files are human-readable and easy to understand. The initial setup of the Openflow Switch contains just the port information

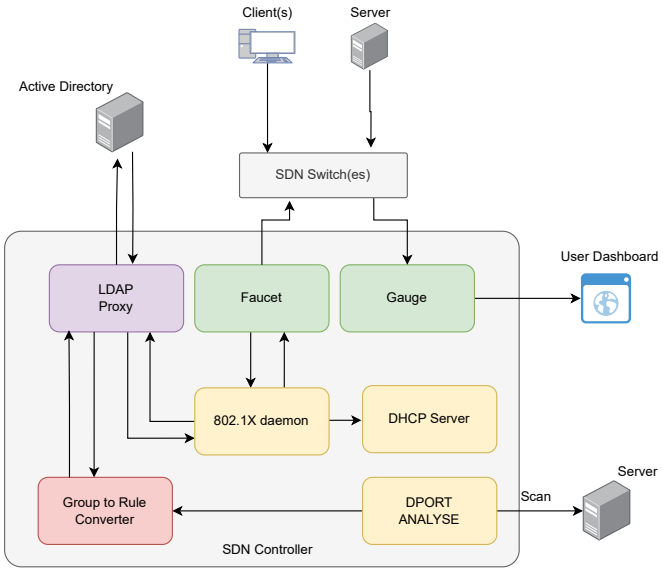

Figure 2. Architecture SDN Controller

and requires authentication before connecting to the Network. Furthermore, specific default rules, such as special treatment for the SDN controller and the Active Directory, were specified beforehand, as these settings are essential when configuring a new network. We used the 802.1X daemon Chewie [23] as a starting point, and then it was heavily adapted to get the user groups via a simple LDAP proxy. A second Service called "Group to Rule" will apply the ACLs as discussed in Section IV. An example rule can be found in Figure 3. It shows the resulting rule with a defined protocol, port, source MAC & IP address, and destination IP address. Since this is directly applied to the port, no other traffic can pass the OpenFlow switch port.

```
acls:
  mac_whitelist_user_ben:
    - rule:
        dl_type: 0x800       # ipv4
        nw_proto: 6          # tcp
        tcp_dst: 80          # port
        eth_src:  32:90:43:57:f2:01
        ipv4_src: 192.168.0.1
        ipv4_dst: 192.168.0.9
        actions:
            allow: 1
    - rule:
        actions:
          allow: False
```

Figure 3. FAUCET ACL CONFIGURATION

A Python script that searches for UDP and TCP ports on the Server provides the open ports needed to craft the ACLs. It then saves this information into a database with the corresponding MAC address and IP address. One problem is that the Server does not directly have an IP address when we try to scan it. We must wait until the IP address is handed

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

over via the DHCP server to start scanning. Therefore, for a server, the ACLs can only be applied later on and not directly, which is not a problem since the default rules still block all access to the Network, and only DHCP is then allowed to obtain an IP address. A simple folder structure was defined for the templates where an administrator can place templates for Groups and specific Ports and the initial Network operations like DHCP and DNS.

The dynamic adaption of the ACLs depending on authentication data is done via the LDAP Proxy. All servers in the Network try to authenticate their users via LDAP Bind requests to the LDAP Proxy, which then forwards this to the Active Directory. This approach allows us to determine whether authentication was successful. We implemented a counter for each user with a threshold of six, which will reset after some time, as determined in Alert-ID [11]. To identify the client who tries to connect, a small script monitors the log file with authentication events and then sends the event from the Server to the controller via a small script. This procedure serves only for demonstration purposes, and, in the future, an SSO Server can perform this task since every client needs to authenticate there when they access a server. For example, blocking users who attempt to authenticate with a username different from the one they initially used for network access would also be possible. Additionally, we could restrict access by blocking only the specific port or server access for such users. Another option would be to reroute all their traffic through an IDS. This approach could alleviate the load on IDS systems by selectively forwarding traffic from suspicious hosts.

The implementation phase reaffirmed the proposed system's potential to enhance network security through fine-grained access controls. However, it highlighted the complexities of managing an extensive rule set, especially in larger networks.

VI. Evaluation

In this evaluation chapter, we begin with a detailed examination of the technical aspects of our experimental setup, laying the groundwork for a thorough assessment. We then delve into the feasibility of the proposed system, followed by a comparative analysis of its efficiency and complexity against existing systems. This analysis sets the stage for a nuanced discussion synthesizing our findings and their implications.

A. Experimental Conditions

Our experimental setup was designed to mirror a realistic laboratory environment consisting of multiple physical PCs and servers to simulate a conventional corporate network infrastructure. The network configuration included five Windows clients alongside a singular Linux client. We assigned the clients to different user groups in the active directory. The configuration of each Client and its connected port can be found in Table I. In setting up our experiment, we went with a mix that one would typically find in an office: a mail server for emails and a GitLab instance for the devs to collaborate on code. This way, we could see how different roles, like developers needing GitLab and managers relying on emails, would interact with the system. It is a practical approach that helps us understand how our setup performs in a real-world scenario.

At the core of our Network was an Active Directory on a Windows Server 2019, connected to a dedicated port at the OpenFlow switch. This switch was a Linux PC running Ubuntu 22.10, with an Intel(R) Core(TM) i7-8700 CPU supporting OpenFlow protocol version 1.3.

This detailed setup provides a solid foundation for evaluating the system's feasibility, performance, and complexity.

TABLE I. CLIENTS IN THE NETWORK

Client Name	Port	Source MAC	Groups
Client1	3	1C:69:7A:6D:C6:27	mail, gitlab
Client2	4	1C:69:7A:43:7C:12	mail
Client3	5	1C:69:7A:6D:C8:B0	mail, gitLab
Client4	6	1C:69:7A:6D:C7:EE	mail
Client5	7	1C:69:7A:6D:C8:16	mail

B. Feasibility

The project aimed to demonstrate the feasibility of such a system and highlight its advantages. Therefore, we conducted multiple experiments to verify the system's operability to achieve this.

1) Experiment 1: Connection to the Network: In our initial experiment, we aimed to verify the functionality of the system's initial configuration and the practical application of Access Control Lists. We began by attempting to connect a server to the Network. Initially, all packets except EAP packets were blocked, preventing any network connection without proper authentication.

To facilitate authentication, we configured the server's wpa_supplicant with EAP after setting up a dedicated user account in the Active Directory for the server, marked by the "server" group identifier, to distinguish it as such. Additionally, the server was assigned to the "mail" group to define its access rights. The authentication process utilized standard Username and Password credentials defined within the Active Directory.

Upon initiating these configurations, we observed successful authentication, followed by the server obtaining an IP address via DHCP. The IP address assignment was managed by the SDN Controller, ensuring the server's connectivity post-authentication. We then proceeded with a port scan, which was feasible only after the OpenFlow switch recognized the server's IP, confirming that the server was operational. The procedure was repeated for the second GitLab server.

Subsequently, we connected a client machine to the Network. Like the server setup, this Client was denied network access until authentication credentials were provided. After authentication, the SDN Controller dynamically generated ACLs based on the Client's group memberships.

For example, the first Client, identified as a developer, was granted access to both the mail server and GitLab, as reflected in the applied ACLs (refer to Table II, lines 1 and 2). This access control was strictly enforced, with all unauthorized traffic being blocked at the port level based on the authenticated

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

TABLE II. ACL CONFIGURATION FOR FIVE CONNECTED CLIENTS

OpenFlow Port	Source MAC	Source IP	Group	Destination IP	Destination Port	Description
3	1C:69:7A:6D:C6:27	192.168.11.11	mail	192.168.11.101	25, 993, 995	Mailserver
3	1C:69:7A:6D:C6:27	192.168.11.11	gitlab	192.168.11.102	22, 80, 443	GitLab
4	1C:69:7A:43:7C:12	192.168.11.12	mail	192.168.11.101	25, 993, 995	Mailserver
5	1C:69:7A:6D:C8:B0	192.168.11.13	mail	192.168.11.101	25, 993, 995	Mailserver
5	1C:69:7A:6D:C8:B0	192.168.11.13	gitlab	192.168.11.102	22, 80, 443	GitLab
6	1C:69:7A:6D:C7:EE	192.168.11.14	mail	192.168.11.101	25, 993, 995	Mailserver
7	1C:69:7A:6D:C8:16	192.168.11.15	mail	192.168.11.101	25, 993, 995	Mailserver

source MAC address and specified port. In contrast, a client identified as a manager, and thus only requiring access to the mail server, demonstrated restricted network access in line with their role (refer to Table II, line 3). Attempts to access GitLab by this Client were blocked, illustrating the ACLs' role-based access control. After connecting all clients, each Client and port results can be found in Table II.

Upon issuing a logoff command to the RADIUS server, all associated ACLs were cleared, reverting the system to its default state of blocking all traffic from the disconnected Client. Logging in with a different username on the same PC triggered a reallocation of ACLs, aligning with the new user's access rights. This experiment demonstrated the feasibility of initially creating and effectively applying ACLs within our network environment.

2) Experiment 2: Failed Logins: In the second experiment, we tested failed login attempts to evaluate the systems' response mechanisms. This test simulated incorrect authentication attempts on the GitLab server to observe the system's reaction.

The experiment began with a series of failed login attempts, with each unsuccessful attempt logged by the SDN Controller. After the sixth failed attempt, the SDN Controller adjusted the ACLs, cutting off the Client's access to the server and other network components. An alert was automatically sent to the network administrator, who could either restore the Client's access after a successful re-authentication or suspend the Client for further investigation.

Additionally, we tested the Network's traffic mirroring feature. In this part of the experiment, despite multiple failed login attempts, the Client was not disconnected from the Network. Instead, the Client's traffic was mirrored to a specific port on an OpenFlow switch. This procedure was verified using tcpdump to confirm that the traffic mirroring was functioning as intended, without the integration of an IDS, since this was not in the scope of the experiment.

C. Complexity and Efficiency

To evaluate our system's complexity and OpenFlow rule management capability, we compared it against other SDN security methods by examining the number of OpenFlow rules in different scenarios. Our analysis included a baseline scenario without ACLs, a basic ACL setup, and scenarios involving VLANs. The scenario with Basic ACLs has only rules for direct IP access. That means we only specify that user X can access server Y without further defining which ports or protocols. The VLAN example does not have any specific ACLs. It splits the users into two groups, usually some kind of department in a corporate network. This option has the disadvantage of allowing clients from the same department to communicate, which does not prevent malware from spreading.

Table III summarizes the OpenFlow rule count for each scenario. As observed, the number of OpenFlow rules directly reflects the count of Faucet ACLs. As discussed in Formula 2, the number of ACLs will increase linearly to the number of users. The dynamic ACL configuration, while more complex, demonstrates the system's flexibility and responsiveness to network changes without significantly impacting performance.

TABLE III. RULE COUNT COMPARISON

Scenario	Faucet ACLs	OpenFlow Rules
No ACLs	0	27
Basic ACLs	8	67
VLANs	N/A	67
Dynamic ACLs	32	91

D. Discussion

In discussing the outcomes and implications of our experiments, it is essential to consider both the implemented system's strengths and potential challenges. To offer a comprehensive understanding of our system's enhanced performance and its innovative approach to network security and management, we performed an extensive comparison across several key metrics, including security level, scalability, and manageability. This comparison, detailed in Table IV, is based on empirical data from ACL number analytics, a comparative analysis of system architectures, and their maintenance needs, highlighting our proposed system's superiority in terms of security, scalability, and ease of management.

The introduction of automated, fine-grained whitelist ACLs represents a significant step forward in network security management. The configuration process substantially decreases administrative overhead and mitigates the risk of human error, which is prevalent in manual configurations. A crucial advantage of this approach is the centralization of security decisions, such as access rights, in a singular user database. This consolidation ensures that modifications to access rights are uniformly applied across the Network and all services utilizing this common user database, thereby enhancing consistency and security within the system. As demonstrated in Experiment 1, the automation of ACL configuration significantly reduced administrative overhead since all needed restrictions are applied individually for each Client without the need for additional

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

TABLE IV. COMPARISON OF NETWORK SECURITY AND MANAGEMENT APPROACHES

Metric	**No ACLs**	**Basic ACLs**	**VLANs**	**Resonance**[12]	**ACL Based on X812**[13]	**Proposed System**
Security Level	Low	Medium	Medium	High	Very High	Very High
Port Security	None	None	None	None	Full	Full
Performance Impact	Low	Medium	Medium	Moderate	Moderate	Moderate
Scalability	High	Moderate	Good	Moderate	Moderate	Excellent
Manageability	Easy	Moderate	Moderate	Moderate	Moderate	Easy
Centralization	None	Low	Low	Medium	Medium	High
Flexibility	Low	Moderate	Low	Very High	Very High	High
Cost Efficiency	High	Moderate	Moderate	Low	Moderate	High
Integration Capability	Seamless	Moderate	Challenging	High	High	Low
Resilience	Low	Medium	Medium	High	High	High
Automation & Dynamic Response	None	None	None	Semi-Automated	Semi-Automated	Fully Automated
ACLs based on Authentication Events	None	None	None	None	None	Supported

adaption by the administrator. Contrarily, this centralized, automated approach ensures that only authenticated users and their associated MAC addresses are actively maintained in the system, limiting access to authorized entities and inherently reducing the risk of unauthorized access. Another significant benefit of our approach is its scalability and ease of integration across multiple switches without additional overhead. Since ACLs and clients are bound to specific ports and not to the physical switches themselves, our system can seamlessly scale to accommodate an extensive network infrastructure with multiple SDN switches. ACLs are applied uniquely to each switch, as delineated in Formula 3, ensuring efficient and tailored security measures are in place, irrespective of the size or complexity of the Network.

However, this automation and simplification come at the cost of increased complexity due to the more significant number of ACLs required to maintain fine-grained control over network access. The number of ACLs does not directly impact the system's performance since it only inspects the TCP header to minimize performance impact, compared to, for example, complex rules that inspect the TCP payload. According to Cabrera et al. [25], the time required to check the payload is, on average, 4.5 times longer than that required for header checks. Therefore, even with many ACL rules, the focus on header information ensures minimal impact on network throughput, as even a single ACL with a TCP header rule necessitates the inspection of every packet. One drawback is that we need to prepare the clients and the server to perform an 802.1X authentication.

One of the more critical considerations is the system's approach to handling failed login attempts, as demonstrated in Experiment 2. Completely blocking access after a series of incorrect credentials can safeguard against brute-force attacks but also pose a risk to business continuity. For instance, automated tools using outdated credentials could inadvertently trigger these security measures, leading to unnecessary disruptions. This aspect of the system necessitates a careful balance between maintaining robust security and ensuring uninterrupted business operations.

Integrating traffic mirroring for suspicious hosts presents a nuanced approach to enhancing security monitoring without overloading the Network or the IDS. By selectively mirroring traffic from potentially compromised hosts, the system can focus on analyzing and responding to genuine threats, improving overall security efficiency. This concept aligns with the approach discussed in [26], which proposes a clustering-based flow grouping scheme that assigns Network flows to various IDSs based on routing information and flow data rates, aiming to optimize the load distribution among IDSs and enhance attack detection capabilities.

VII. Conclusion

In conclusion, the proposed system for Port Access Control presents a straightforward implementation framework that significantly enhances network security by enforcing fine-grained access control rules. By leveraging a common user database, such as Active Directory, and binding access controls to specific MAC addresses, the system ensures that only authenticated users can access network ports, providing a robust security posture. Moreover, the approach of mirroring traffic from suspicious hosts, particularly those with repeated failed login attempts, suggests a promising avenue for optimizing IDS performance. Optimizing algorithms for handling multiple concurrent access requests, managing extensive rule sets without compromising performance, and assessing the impact of selective traffic mirroring on IDS efficiency are critical future research areas to enhance scalability and maintain security in complex network architectures.

Acknowledgment

The authors would like to thank Prof. Hiroki Takakura for useful advice. This work was partially supported by JSPS KAKENHI Grant Number JP19K20268 and JP23H03396.

References

[1] H. Zhou, C. Wu, M. Jiang, B. Zhou, W. Gao, T. Pan, and M. Huang, "Evolving defense mechanism for future network security," *IEEE Communications Magazine*, vol. 53, no. 4, pp. 45–51, 2015.

[2] S. T. Ali, V. Sivaraman, A. Radford, and S. Jha, "A survey of securing networks using software defined networking," *IEEE transactions on reliability*, vol. 64, no. 3, pp. 1086–1097, 2015.

[3] F. Bannour, S. Souihi, and A. Mellouk, "Distributed SDN control: Survey, taxonomy, and challenges," *IEEE*

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

Communications Surveys & Tutorials, vol. 20, no. 1, pp. 333–354, 2017.

[4] Global Market Insights, "Software Defined Networking (SDN) Market," 2024, report ID: GMI2395, Accessed: 2024-04-09. [Online]. Available: https://www.gminsights.com/industry-analysis/software-defined-networking-sdn-market

[5] F. S. Ali, R. Amin, M. Majeed, and M. M. Iqbal, "Dynamic ACL Policy Implementation in Software Defined Networks," in *2022 International Conference on IT and Industrial Technologies (ICIT)*, Oct 2022, pp. 01–07.

[6] S. T. Yakasai and C. G. Guy, "FlowIdentity: Software-defined network access control," in *2015 IEEE Conference on Network Function Virtualization and Software Defined Network (NFV-SDN)*. IEEE, 2015, pp. 115–120.

[7] M. Casado, M. J. Freedman, J. Pettit, J. Luo, N. McKeown, and S. Shenker, "Ethane: taking control of the enterprise," in *Proceedings of the 2007 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*, ser. SIGCOMM '07. New York, NY, USA: Association for Computing Machinery, 2007, p. 1–12.

[8] D. M. Ferrazani Mattos and O. C. M. B. Duarte, "AuthFlow: authentication and access control mechanism for software defined networking," *annals of telecommunications*, vol. 71, pp. 607–615, 2016.

[9] T. Xing, D. Huang, L. Xu, C.-J. Chung, and P. Khatkar, "SnortFlow: A OpenFlow-Based Intrusion Prevention System in Cloud Environment," in *2013 Second GENI Research and Educational Experiment Workshop*, March 2013, pp. 89–92.

[10] A. Le, P. Dinh, H. Le, and N. C. Tran, "Flexible Network-Based Intrusion Detection and Prevention System on Software-Defined Networks," in *2015 International Conference on Advanced Computing and Applications (ACOMP)*, Nov 2015, pp. 106–111.

[11] J. Chu, Z. Ge, R. Huber, P. Ji, J. Yates, and Y.-C. Yu, "ALERT-ID: analyze logs of the network element in real time for intrusion detection," in *Research in Attacks, Intrusions, and Defenses: 15th International Symposium, RAID 2012, Amsterdam, The Netherlands, September 12-14, 2012. Proceedings 15*. Springer, 2012, pp. 294–313.

[12] A. K. Nayak, A. Reimers, N. Feamster, and R. Clark, "Resonance: dynamic access control for enterprise networks," in *Proceedings of the 1st ACM Workshop on Research on Enterprise Networking*, ser. WREN '09. New York, NY, USA: Association for Computing Machinery, 2009, p. 11–18.

[13] B. J. C. de A. Martins, D. M. Mattos, N. C. Fernandes, D. Muchaluat-Saade, A. B. Vieira, and E. F. Silva, "An Extensible Access Control Architecture for Software Defined Networks based on X.812," in *2019 IEEE Latin-American Conference on Communications (LATINCOM)*, 2019, pp. 1–6.

[14] N. McKeown, T. Anderson, H. Balakrishnan, G. Parulkar, L. Peterson, J. Rexford, S. Shenker, and J. Turner, "OpenFlow: enabling innovation in campus networks," *ACM SIGCOMM computer communication review*, vol. 38, no. 2, pp. 69–74, 2008.

[15] M. S. Farooq, S. Riaz, and A. Alvi, "Security and Privacy Issues in Software-Defined Networking (SDN): A Systematic Literature Review," *Electronics*, vol. 12, no. 14, 2023.

[16] "IEEE Standard for Local and Metropolitan Area Networks–Port-Based Network Access Control," *IEEE Std 802.1X-2020 (Revision of IEEE Std 802.1X-2010 Incorporating IEEE Std 802.1Xbx-2014 and IEEE Std 802.1Xck-2018)*, pp. 1–289, 2020.

[17] J. Vollbrecht, J. D. Carlson, L. Blunk, D. B. D. Aboba, and H. Levkowetz, "Extensible Authentication Protocol (EAP)," RFC 3748, Jun. 2004.

[18] B. Desmond, J. Richards, R. Allen, and A. G. Lowe-Norris, *Active Directory: Designing, Deploying, and Running Active Directory*. " O'Reilly Media, Inc.", 2008.

[19] J. Sermersheim, "Lightweight Directory Access Protocol (LDAP): The Protocol," RFC 4511, Jun. 2006.

[20] H. Hasegawa, Y. Sato, and H. Takakura, "Construction of Secure Internal Network with Communication Classifying System Using Multiple Judgment Methods," *International Journal on Advances in Telecommunications*, vol. 13, no. 3 & 4, 2020.

[21] Y. Sato, H. Hasegawa, and H. Takakura, "Construction of Secure Internal Networks with Communication Classifying System," in *ICISSP*, 2019, pp. 552–557.

[22] M. Ali, N. Shah, and M. A. Khan Khattak, "DAI: Dynamic ACL Policy Implementation for Software-Defined Networking," in *2020 IEEE 17th International Conference on Smart Communities: Improving Quality of Life Using ICT, IoT and AI (HONET)*, Dec 2020, pp. 138–142.

[23] FaucetSDN, "Faucet," 2024, accessed: 2024-04-09. [Online]. Available: https://github.com/faucetsdn/faucet

[24] Ryu SDN Framework Community, "Ryu sdn framework," 2024, accessed: 2024-04-09. [Online]. Available: https://ryu-sdn.org

[25] J. B. Cabrera, J. Gosar, W. Lee, and R. K. Mehra, "On the statistical distribution of processing times in network intrusion detection," in *2004 43rd IEEE Conference on Decision and Control (CDC)(IEEE Cat. No. 04CH37601)*, vol. 1. IEEE, 2004, pp. 75–80.

[26] T. Ha, S. Yoon, A. C. Risdianto, J. Kim, and H. Lim, "Suspicious flow forwarding for multiple intrusion detection systems on software-defined networks," *IEEE Network*, vol. 30, no. 6, pp. 22–27, 2016.

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

Application of a Deep Reinforcement Learning Algorithm to Virtual Machine Migration Control in Multi-Stage Information Processing Systems

Yukinobu Fukushima
Okayama University
Okayama, Japan
e-mail: fukusima@okayama-u.ac.jp

Yuki Koujitani
Okayama University
Okayama, Japan

Kazutoshi Nakane
Nagoya University
Nagoya, Japan

Yuya Tarutani
Okayama University
Okayama, Japan

Celimuge Wu
The Univ. of Electro-Commun.
Tokyo, Japan

Yusheng Ji
National Institute of Informatics
Tokyo, Japan

Tokumi Yokohira
Okayama University
Okayama, Japan

Tutomu Murase
Nagoya University
Nagoya, Japan

***Abstract*—This paper tackles a Virtual Machine (VM) migration control problem to maximize the progress (accuracy) of information processing tasks in multi-stage information processing systems. The conventional methods for this problem (e.g., VM sweeping method and VM number averaging method) are effective only for specific situations, such as when the system load is high. In this paper, in order to achieve high accuracy in various situations, we propose a VM migration method using a Deep Reinforcement Learning (DRL) algorithm. It is difficult to directly apply a DRL algorithm to the VM migration control problem because the size of the solution space of the problem dynamically changes according to the number of VMs staying in the system while the size of the agent's action space is fixed in DRL algorithms. Therefore, the proposed method divides the VM migration control problem into two problems: the problem of determining only the VM distribution (i.e., the proportion of the number of VMs deployed on each edge server) and the problem of determining the locations of all the VMs so that it follows the determined VM distribution. The former problem is solved by a DRL algorithm, and the latter problem is solved by a heuristic method. The simulation results confirm that our proposed method can select quasi-optimal VM locations in various situations with different link delays.**

Keywords-Multi-stage information processing system, VM migration control, Deep reinforcement learning, Deep Deterministic Policy Gradient (DDPG)

I. INTRODUCTION

In recent years, ultra-real-time services, such as Cross Reality (XR) and automated driving, are expected to appear. In these services, information processing tasks requested by clients need to be executed immediately (e.g., on the order of milliseconds) and the processing results should be as accurate as possible.

A multi-stage information processing system [1] [2] is one of the promising candidates for the edge computing infrastructures for ultra-real-time services. In the system, information processing tasks requested by clients are executed in parallel by an edge server and a data center. The edge server prioritizes responsiveness over accuracy; it returns the highly responsive but low accurate processing results to the clients while the data center prioritizes accuracy over responsiveness; it return the highly accurate but low responsive processing results to the clients. When operating ultra-real-time services in a multi-stage information processing system, it is important to maximize the accuracy of information processing tasks executed by the edge servers that satisfy the responsiveness requested by clients.

Previous researches on multi-stage information processing systems focused on improving the accuracy of information processing tasks executed by edge servers through Virtual Machine (VM) migration control [1] [2]. VM migration control dynamically migrates VMs, which execute the information processing tasks requested by clients on edge servers, among multiple edge servers, which leads to effective use of CPU resources on edge servers and reducing the communication delay between clients and VMs, thereby improving the accuracy of the information processing tasks. In the previous researches, as heuristic methods for VM migration control, VM sweeping method [2], VM number averaging method [2], early-blooming type priority processing method [1], and late-blooming type priority processing method [1] were proposed and their effectiveness were confirmed. These methods are, however, effective only in specific situations, such as when the system load is high and the type of information processing tasks is late-blooming type. Since the system load and the type of information processing tasks change dynamically, VM migration control that can achieve high accuracy in a wide variety of situations is needed.

In this paper, in order to achieve high accuracy in a variety of situations, we propose a VM migration method using a Deep Reinforcement Learning (DRL) algorithm. DRL algorithms are expected to achieve a quasi-optimal performance in a variety of situations through interactions between a learning agent and a dynamically changing environment. On the other hand, it is difficult to directly apply a DRL algorithm to the VM migration control problem because, in the problem, the size of the solution space dynamically changes according to the dynamic changes in the number of VMs staying in the system while the size of the agent's action space is fixed in DRL algorithms. Therefore, in this paper, we divide the VM migration control problem into two problems: the problem of determining only the VM distribution (i.e., the proportion of the number of VMs deployed on each edge server) and the

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

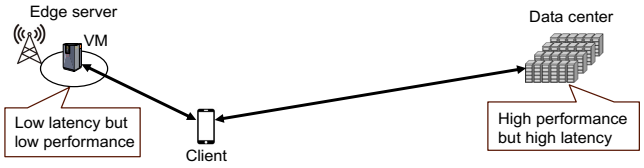

Figure 1. Multi-stage information processing systems.

problem of determining the locations of all the VMs so that it follows the determined VM distribution. The former problem is solved by a DRL algorithm, and the latter problem is solved by a heuristic method. This approach makes it possible to apply a DRL algorithm with a fixed action space size to the VM migration control problem.

The rest of this paper is organized as follows. Section 2 introduces related work on VM migration control. Section 3 describes the multi-stage information processing system and the VM migration control problem. In Section 4, we propose a VM migration method using a DRL algorithm. In Section 5, we evaluate the effectiveness of our proposed method with computer simulations. In Section 6, we summarize the paper and describe our future works.

II. Related Work

The work in [3]–[10] tackle VM migration control problems in server migration services and propose heuristic methods [3] [5], mathematical programming methods [4], [6]–[8], [10], and Q-learning methods [9]. These methods, however, aim at improving the communication quality between clients and VMs and reducing network power consumption, and do not consider the accuracy of information processing tasks.

The research in [1] [2] tackle VM migration control problems in multi-stage information processing systems, and propose the heuristic methods; VM sweeping method [2], VM number averaging method [2], early-blooming type priority processing method [1], and late-blooming type priority processing method [1]. These methods are, however, effective only in specific situations. For example, the VM sweeping method is shown to be effective only in situations where the system load is high and the type of information processing tasks is late-blooming type. Since the system load and the type of information processing tasks change dynamically, VM migration control that can achieve high accuracy in a wide variety of situations is needed.

The work in [11] [12] tackle VM migration control problems in mobile edge computing, and propose VM migration methods using Deep Q-Network (DQN) [13], which is a kind of DRL algorithms. These methods, however, can only be applied to VM migration control problems with a single VM because the size of an agent's action space is fixed in DQN, and cannot be applied to VM migration control problems with multiple VMs.

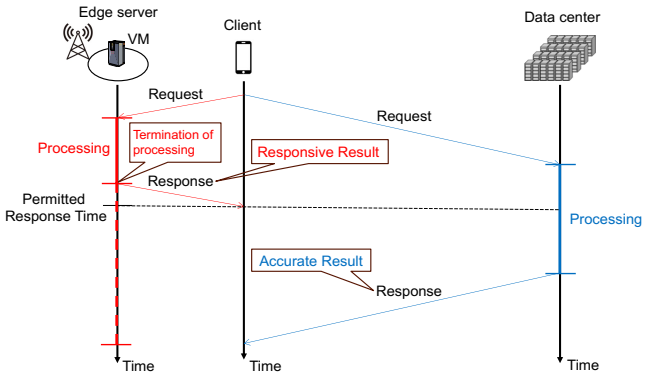

Figure 2. Flow of information processing in a multi-stage information processing system.

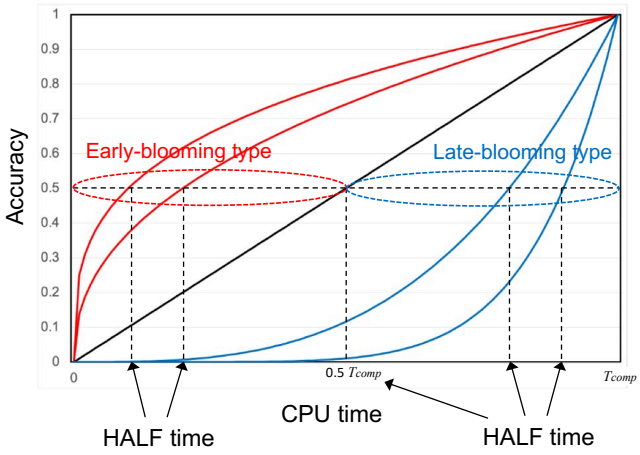

Figure 3. Relationship between CPU time allocated to a task and accuracy of the task.

III. Multi-Stage Information Processing Systems

As shown in Figure 1, a multi-stage information processing system consists of edge servers located proximate (e.g., base stations) to clients and data centers located distant from them. The system provides clients with both highly responsive and highly accurate processing results by executing information processing tasks in parallel at the edge servers and the data centers.

Figure 2 shows the flow of information processing in a multi-stage information processing system. A client requests both an edge server and a data center to process its task in parallel. When the response time permitted by the client approaches, the edge server terminates its processing to meet the permitted response time and returns the highly responsive processing result to the client. The data center, on the other hand, accomplishes its processing and returns the highly accurate processing result to the client.

In this paper, we assume the accuracy model (i.e., the relationship between the CPU time (t_{CPU}) allocated to a task

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

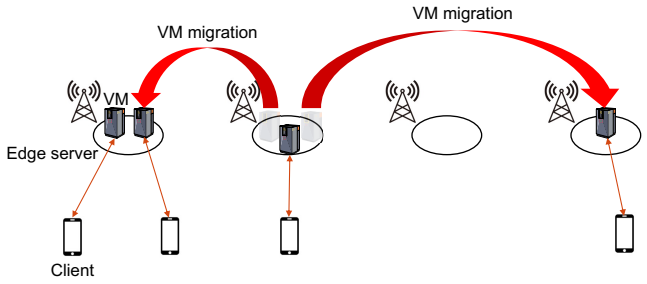

Figure 4. VM migration control in a multi-stage information processing system.

and the accuracy ($f(t_{CPU})$) of the task) that is adopted in [2]. Figure 3 shows the accuracy model. In the accuracy model, the accuracy of the task is calculated as follows.

$$f(t_{CPU}) = (\frac{t_{CPU}}{T_{comp}})^{\frac{log(0.5)}{log(\frac{HALFtime}{T_{comp}})}} \quad (1)$$

where T_{comp} represents the time for the task to be completed (i.e., accuracy reaches 1.0) and HALF time represents the time for the task to reach accuracy of 0.5. Tasks are classified based on their HALF time. The tasks with HALF time shorter than 0.5 T_{comp} are classified into early-blooming type while those with HALF time longer than 0.5 T_{comp} are classified into late-blooming type.

In this paper, we tackle a VM migration control problem among multiple edge servers for maximizing the accuracy of information processing tasks returned by edge servers within the permitted response times (Figure 4). VM migration control enables effective use of CPU resources on edge servers and reducing the communication delay between clients and VMs, thereby improving the accuracy of information processing tasks.

IV. Proposed Method

In this paper, in order to achieve high accuracy in a variety of situations, we propose a VM migration method using a DRL algorithm. With regard to applying a DRL algorithm to a VM migration control problem, it should be noted that the size of the solution space (i.e., the total number of all possible solutions) of the problem dynamically changes according to the dynamic changes in the number of VMs staying in the system. As shown in Figure 5, the size of the solution space is E^K where E is the number of edge servers and K is the number of VMs, and the size of the solution space E^K dynamically changes according to the number of VMs K. On the other hand, the size of the agent's action space in DRL algorithms is fixed. For example, an agent in Deep Deterministic Policy Gradient (DDPG) [14] outputs a vector with a fixed number of dimensions. Therefore, It is difficult to directly apply a DRL algorithm to the VM migration control problem.

To cope with the dynamic change in the size of solution space, we divide the VM migration control problem into two

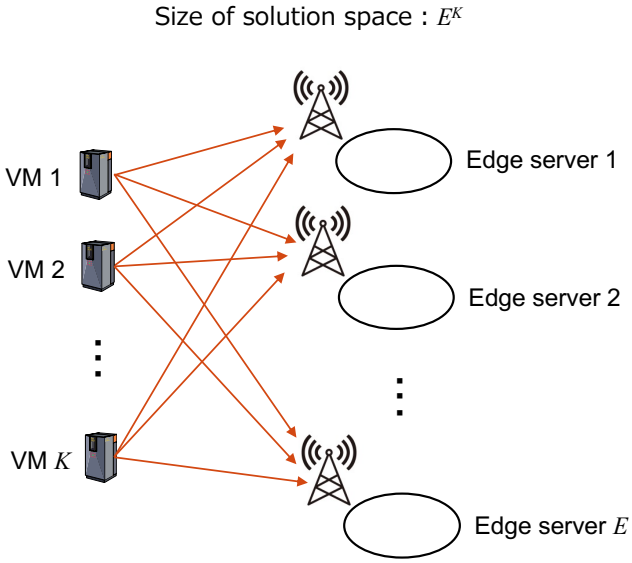

Figure 5. Size of solution space in VM migration control problem.

problems (Figure 6): the problem of determining only the VM distribution (i.e., the proportion of the number of VMs deployed on each edge server) and the problem of determining the locations of all the VMs so that it follows the determined VM distribution. The former problem is solved by a DRL algorithm, and the latter problem is solved by a heuristic method. This approach makes it possible to apply a DRL algorithm with a fixed action space size to the VM migration control problem because the VM distribution can be expressed by a vector with a fixed number of dimensions.

We adopt DDPG [14] as a DRL algorithm. DDPG approximates both a policy function $\mu(s|\theta)$ (Actor), which maps a given state to an action to be taken, and an action-value function $Q(s, a|\phi)$ (Critic) with deep neural networks. In DDPG, the Actor can output the VM distribution (i.e., the proportion of the number of VMs deployed on each edge server) as an action because it can operate over continuous action space. As well as DQN [13], DDPG adopts experience replay and target network techniques in order to learn Actor and Critic in a stable and robust way.

Figure 7 depicts an interaction between a DDPG agent and an environment, which corresponds to the VM migration control problem. When applying a DRL algorithm to the VM migration control problem, we need to define action, state, and reward in accordance with the problem. Action a_t of the agent is defined as the VM distribution (i.e., the proportion of the number of VMs deployed on each edge server), and is expressed with the following equation.

$$a_t = (p_1, p_2, \ldots, p_E) \quad (2)$$

where p_i is the proportion of the number of VMs deployed on edge server i. State s_t of the environment is defined as the

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

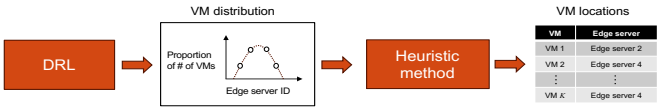

Figure 6. Outline of our proposed method.

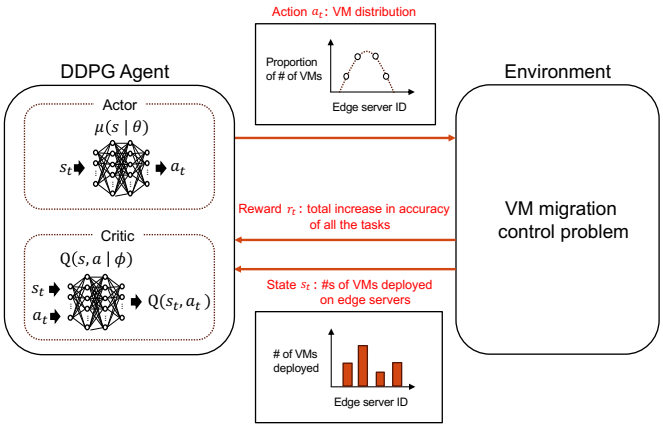

Figure 7. Interaction between the DDPG agent and the environment.

numbers of VMs deployed on edge servers, and is expressed with the following equation.

$$s_t = (d_1, d_2, \ldots, d_E) \tag{3}$$

where d_i is the number of VMs deployed on edge server i. Reward r_t is defined as the total increase in accuracy of all the tasks during the period from the last VM migration control to the current one. Algorithm 1 in Figure 8 shows the procedure of our proposed method.

After determining the VM distribution, we determine the locations of all the VMs by a heuristic method so that it follows the determined VM distribution. In this paper, we adopt a minimum client-VM delay method as the heuristic method. The minimum client-VM delay method selects the VM location with the minimum sum of the delays between clients and VMs in a brute force manner among the VM locations that follow the VM distribution determined by the DDPG agent.

V. Performance Evaluation

In this section, we evaluate our proposed method with computer simulations. Section V.A explains the simulation model. Section V.B shows the evaluation results.

A. Simulation Model

We developed the VM migration control simulator and the DDPG agent with OpenAI Gym [15] and Keras-rl [16], respectively. Table I summarizes the parameter settings as to the DDPG agent. We adopt the same parameter values as those used by the DDPG agent in Keras-rl [16] because the

Algorithm 1 Procedure of our proposed method

1: Randomly initialize weights θ of Actor $\mu(s|\theta)$ and weights ϕ of Critic $Q(s, a|\phi)$
2: Initialize weights of Actor's target network $\mu'(s|\theta')$ and Critic's target network $Q'(s, a|\phi')$: $\theta' \leftarrow \theta$, $\phi' \leftarrow \phi$
3: Initialize replay buffer R
4: **for** episode = 1, M **do**
5: Initialize a random noise $\mathcal{N}$ for action exploration
6: Observe initial state s_1 from the environment
7: **for** t = 1, T **do**
8: Select VM distribution $a_t = \mu(s_t|\theta) + \mathcal{N}_t$ as action
9: Determine locations of all the VMs by the heuristic method among the VM locations that follow the determined VM distribution a_t, and migrates the VMs
10: Observe reward r_t and the next state s_{t+1}
11: Store experience (s_t, a_t, r_t, s_{t+1}) in R
12: Sample a random minibatch of N experiences (s_i, a_i, r_i, s_{i+1}) from R
13: Learning of Critic:
Calculate target value y_i:
$y_i = r_i + \gamma Q'(s_{i+1}, \mu'(s_{i+1}|\theta')|\phi')$
Update weights ϕ with a gradient descent method so that loss $L = \frac{1}{N}\sum_i (y_i - Q(s_i, a_i|\phi))^2$ is minimized
14: Learning of Actor:
Calculate policy gradient $\nabla_\theta J$:
$\nabla_\theta J \propto \frac{1}{N}\sum_i \nabla_a Q(s_i, \mu(s_i)|\phi) \nabla_\theta \mu(s_i|\theta)$
Update weights θ with a gradient ascent method so that performance of Actor J is maximized
15: Update weights of target networks:
$\theta' \leftarrow \tau\theta + (1-\tau)\theta'$
$\phi' \leftarrow \tau\phi + (1-\tau)\phi'$
16: **end for**
17: **end for**

Figure 8. Procedure of our proposed method.

work [14] reports that a DDPG agent with the same parameter setting successfully solved various physics tasks.

The left side of Figure 9 shows the network model. This paper tackles the early stage of the performance evaluation of our proposed method; we focus on the case where four clients join and leave the multi-stage information processing system in a specific pattern on the small-scale network. The network consist of four edge servers, which are connected in a full mesh manner. We assume that the delays of all the links are identical. In order to evaluate whether our proposed method can cope with various situations with different link delay, we set the delay of each link to one of the following values: 1, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 [ms]. An edge server equally allocates its CPU time to all the VMs located on it. A VM is individually generated for each client. We set the response time permitted by a client to 110 [ms], the completion time

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

TABLE I
PARAMETER SETTINGS

Parameter	Value
Number of training episodes (M)	10,000
Discount rate (γ)	0.99
Number of hidden layers	Actor : 2, Critic : 5
Number of neurons in a hidden layer	Actor : 256, 256, Critic : 16, 32, 32, 256, 256
Activation function of hidden layers	Actor : relu, Critic : relu
Learning rate (α)	Actor : 0.001, Critic : 0.002
Noise process for action exploration ($\mathcal{N}$)	Ornstein-Uhlenbeck process
Size of replay buffer	10,000
Minibatch size (N)	64
Weights of updated parameters when updating the weights of target networks (τ)	0.005

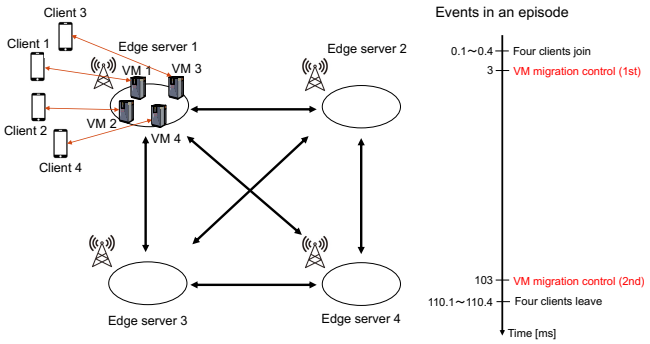

Figure 9. Network model and events in an episode.

of an information processing task (T_{comp}) to 110 [ms], and HALF time to 11 [ms] (= 0.1 $\times T_{comp}$) assuming the task type is the early-blooming type.

During an episode of the simulation, the following events occur (right side of Figure 9). When an episode starts, four clients join the system in turn every 0.1 [ms] from time 0.1 [ms] to time 0.4 [ms]. The locations of all the clients are fixed at edge server 1 during the episode. The initial locations of all the VMs are set to edge server 1. At time 3 [ms], we perform the first VM migration control. Then, at time 103 [ms], we perform the second VM migration control. Lastly, the four clients leave the system in turn every 0.1 [ms] from time 110.1 [ms] to time 110.4 [ms]. The first VM migration control aims at determining the locations of the VMs during the episode and the second VM migration control aims at obtaining the reward and the experience for learning the DDPG agent.

We compare our proposed method with the following methods.

- VM sweeping method [2]
  It migrates a VM with higher accuracy increase rate to an idle edge server so that the VM occupies the CPU time on it.
- VM number averaging method [2]
  It equally distributes all the VMs to all the edge servers for load balancing.
- Non-migration method
  It fixes all the VMs at their initial locations.
- Minimum client-VM delay method
  It locates each of the VMs to the location most proximate to its client.

B. Evaluation Results

Figure 10 shows the average accuracy among the information processing tasks executed by the four VMs for all the VM migration methods. The accuracy of our proposed method (DDPG + Minimum client-VM delay method) is plotted with 95% confidence interval because it varies depending on the initial weights of Actor and Critic, and the noises for action exploration.

Both non-migration method and minimum client-VM delay method show the constant accuracy of about 0.65 regardless

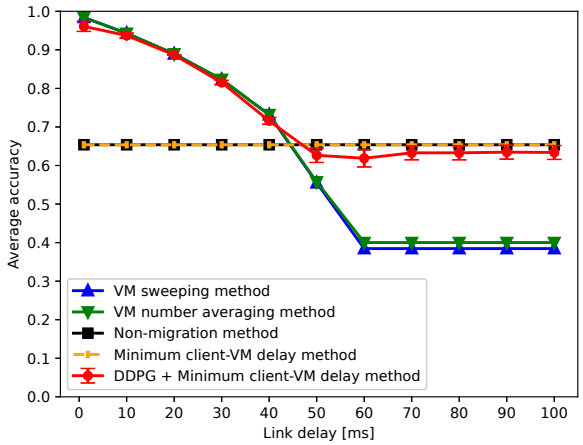

Figure 10. Average accuracy as a function of link delay.

of the link delay. This is because these methods always fix all the VMs at their initial locations (edge server 1) regardless of the link delay.

Both VM sweeping method and VM number averaging method achieve the maximum accuracy of about 0.98 when the link delay is 1 [ms], and the accuracy decreases as the link delay increases. This is explained as follows. These methods always distribute the VMs to all edge servers so that a VM is located at an edge server regardless of the link delay. As the link delay increases, the VM migration time and the communication delay between the client and the VM increases, and consequently the CPU time allocated to the task at the VM decreases after VM migration.

We compare the performances of non-migration method, minimum client-VM delay method, VM sweeping method, and VM number averaging method. When the link delay is lower than or equal to 40 [ms], VM sweeping method and VM number averaging method achieve 12 to 50% higher accuracy than non-migration method and minimum client-VM delay method. Therefore, in this case, it is desirable to distribute all the VMs to different edge servers. When the link delay

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

is higher than or equal to 50 [ms], non-migration method and minimum client-VM delay method achieve 17 to 70 % higher accuracy than VM sweeping method and VM number averaging method. Therefore, in this case, it is desirable to fix all the VMs at their initial locations.

Lastly, we focus on the performance of our proposed method. When the link delay is lower than or equal to 40 [ms], our proposed method 1) achieves 9 to 47% higher accuracy than non-migration method and minimum client-VM delay method, and 2) achieves almost as high accuracy (at most 2% lower accuracy) as VM sweeping method and VM number averaging method. When the link delay is higher than or equal to 50 [ms], our proposed method 1) achieves 12 to 65% higher accuracy than VM sweeping method and VM number averaging method, and 2) achieves almost as high accuracy (at most 5% lower accuracy) as non-migration method and minimum client-VM delay method. Therefore, our proposed method can select quasi-optimal VM locations in various situations with different link delays.

VI. Conclusions

In this paper, we proposed a VM migration method using a DRL algorithm in order to achieve high accuracy of information processing tasks in various situations for multi-stage information processing systems. Our proposed method divides the VM migration control problem into two problems: the problem of determining only the VM distribution and the problem of determining the locations of all the VMs so that it follows the determined VM distribution. Our proposed method solves the former problem by a DRL algorithm and the latter problem by the minimum client-VM delay method. The simulation results confirm that our proposed method can select quasi-optimal VM locations in various situations with different link delays.

In our future work, we plan to evaluate the performance of our proposed method 1) when the number of clients and VMs, and the type of information processing tasks dynamically change and 2) when different heuristic methods are adopted for the VM location decision problem in our proposed method.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Number JP23K11065.

References

[1] K. Nakane, T. Anjiki, J. Xie, Y. Fukushima, and T. Murase, "VM Migration Considering Downtime for Accuracy Improvement in Multi-Stage Information Processing System," in *Proc. of IEEE ICCE*, Jan. 2022, pp. 335–336.

[2] T. Anjiki, K. Nakane, and T. Murase, "Performance Improvement by Controlling VM Migration between Edge Nodes in a Multi-Stage Information Processing System," in *Proc. of WSCE*, Sept. 2022, pp. 53–58.

[3] A. Yamanaka, Y. Fukushima, T. Murase, T. Yokohira, and T. Suda, "Destination selection algorithm in a server migration service," in *Proc. of CFI*, Sept. 2012, pp. 15–20.

[4] Y. Fukushima, T. Murase, T. Yokohira, and T. Suda, "Optimization of server locations in server migration service," in *Proc. of ICNS*, March 2013, pp. 200–206.

[5] Y. Hoshino, Y. Fukushima, T. Murase, T. Yokohira, and T. Suda, "An online algorithm to determine the location of the server in a server migration service," in *Proc. of IEEE CCNC*, Jan. 2015, pp. 740–745.

[6] Y. Fukushima, T. Murase, T. Yokohira, and T. Suda, "Power-Aware Server Location Decision in Server Migration Service," in *Proc of ICTC*, pp. 150–155, Oct. 2016.

[7] Y. Fukushima, T. Murase, G. Motoyoshi, T. Yokohira, and T. Suda, "Determining Server Locations in Server Migration Service to Minimize Monetary Penalty of Dynamic Server Migration," Springer Journal of Network and Systems Management, Vol. 26, Iss. 4, pp. 993–1033, Oct. 2018.

[8] Y. Fukushima, T. Murase, and T. Yokohira, "Link Capacity Provisioning and Server Location Decision in Server Migration Service," in *Proc of IEEE CloudNet*, pp. 1–3, Oct. 2018.

[9] R. Urimoto, Y. Fukushima, Y. Tarutani, T. Murase, and T. Yokohira, "A Server Migration Method Using Q-Learning with Dimension Reduction in Edge Computing," in *Proc of ICOIN*, pp. 301–304, Jan. 2021.

[10] Y. Fukushima, T. Suda, T. Murase, Y. Tarutani, and T. Yokohira, "Minimizing the Monetary Penalty and Energy Cost of Server Migration Service," Wiley Transactions on Emerging Telecommunications Technologies, Vol. 33, Iss. 9, pp. 1–34, Sept. 2022.

[11] D. Zeng, L. Gu, S. Pan, J. Cai, and S. Guo, "Resource Management at the Network Edge: A Deep Reinforcement Learning Approach," IEEE Network, Vol. 33, Iss. 3, pp. 26–33, May/June 2019.

[12] C. Zhang and Z. Zheng, "Task Migration for Mobile Edge Computing Using Deep Reinforcement Learning," Future Generation Computer Systems, Vol. 96, pp. 111–118, 2019.

[13] M. Volodymyr, et al. "Playing Atari with Deep Reinforcement Learning," arXiv preprint arXiv:1312.5602, 2013.

[14] T. P. Lillicrap, et al. "Continuous Control with Deep Reinforcement Learning," arXiv preprint arXiv:1509.02971, 2015.

[15] G. Brockman, et al. "OpenAI Gym," arXiv preprint arXiv:1606.01540, 2016.

[16] M. Plappert, "Keras-rl," GitHub repository, 2016.

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

Measurement of the Per-flow Burst Ratio Parameter in IP Networks

Dominik Samociuk
Department of Computer Networks and Systems
Silesian University of Technology
Gliwice, Poland
email: dsamociuk@polsl.pl

***Abstract*—In this study, we investigate the burst ratio, a significant parameter in networking that characterizes the propensity of packet losses to occur in extended, consecutive sequences. The presence of prolonged sequences of packet losses can adversely affect multimedia streams, especially those of real-time nature. In packet networks, the burst ratio is often escalated due to packet buffer overflows that occur in routers and switches, which are inherent in these systems. To analyze the burst ratio, we focus on a per-flow approach, wherein we assess the burst ratio individually for each flow of packets passing through a network node. Additionally, we compare all the per-flow burst ratios with one another and with the burst ratio calculated for the multiplexed traffic. The study explores the impact of various factors on burst ratios. Specifically, we examine the influence of flow rates and their relative proportions, the standard deviation of interarrival times, buffer capacity, the load imposed on the buffer, and the distribution of service time. Particular emphasis is placed on models with non-Poisson flows, as they are not analytically tractable. To conduct this investigation, we utilize real-world data from networking scenarios rather than simulations, allowing us to draw more accurate and robust conclusions. The obtained burst ratio measurements provide valuable insights into the behavior of packet loss processes in complex network environments and aid in enhancing the performance of multimedia streams, particularly real-time applications.**

***Keywords*—burst ratio, packet networks, IP networks, packet loss**

I. Introduction

The concept of burst ratio, denoted as B, was initially introduced in reference [1]. It is determined by the ratio of the average length of observed loss sequences within a given stream of interest, denoted as G, to the hypothetical average length of loss sequences in a Bernoulli process. In the Bernoulli process, all losses are considered to be entirely random, uncorrelated, and share a common probability of occurrence denoted as L.

Mathematically, this relationship can be expressed as follows:

Where:

$$B = G/K, \tag{1}$$

- B represents the burst ratio characteristic,
- G denotes the average length of observed loss sequences in the stream of interest, and
- K denotes the hypothetical average length of loss sequences in the Bernoulli process.

This formulation enables the quantification of the extent to which losses within a specific stream deviate from the behavior expected in an idealized Bernoulli process, facilitating a comparative analysis of the burstiness of different data streams in networking scenarios.

Through a straightforward verification, it is evident that in the case of the Bernoulli process, the value of K can be represented as:

$$K = (1 - L)^{-1}, \tag{2}$$

Consequently, as an alternative to the previously mentioned equation (1), the burst ratio (B) can be calculated using the following expression:

$$B = (1 - L)G. \tag{3}$$

To provide an illustrative example, let's consider a stream of 20 packets:

$$SSDSSSSDDDSSSSSSDSSS. \tag{4}$$

In this sequence, 'S' represents a successful packet transmission, and 'D' indicates a packet loss, packet dropped. By calculating the overall loss probability (loss ratio), denoted as L, we find:

$$L = 5/20 = 1/4 = 0.25, \tag{5}$$

Accordingly, the hypothetical mean length of the sequence of 'D's in the Bernoulli process with L=0.25 should be:

$$K = (1 - L)^{-1} = 4/3 = 1.33(recurring), \tag{6}$$

However, in our stream, we identify three sequences of 'D's, with first and third sequences containing only one dropped packet and second sequence containing three lost packets. Thus, the mean length of the sequence of 'D's in our stream, denoted as G, is equal to 1.66 (recurring).

Consequently, the burst ratio, B, is computed as:

$$B = G/K = 1.66/1.33 = 1.25, \tag{7}$$

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

In contemporary packet networks, empirical studies reveal that the parameter G is often 1.5 to 2.0 times greater than the parameter K [2] [3]. This discrepancy suggests an underlying mechanism influencing consecutive packet losses, whereby the probability of losing a packet is heightened following the loss of its predecessor. This phenomenon primarily arises from buffer overflows in routers and switches. During an overflow event, all incoming packets are discarded until sufficient buffer space becomes available. Consequently, this leads to a series of packet losses.

The increased burst ratio observed in modern networks, particularly the Internet, detrimentally affects the Quality of Experience (QoE) in multimedia streaming, especially in real-time applications. The loss of lengthy sequences of video or audio packets can significantly degrade QoE more than frequent losses of shorter sequences. This impact is attributable to the nature of human perception and the codecs employed in real-time multimedia transmission. For certain transmission types, this effect is quantifiable, which models the decline in voice transmission quality in Internet telephony as a function of the burst ratio.

In this study, we adopt a per-flow approach to investigate the burst ratio, assessing it for each individual packet flow passing through a network node, as opposed to evaluating a global burst ratio for aggregated traffic. Our methodology involves comparing the burst ratios of individual flows against each other and against the burst ratio of multiplexed traffic. We also examine the impact of various factors on these ratios, including flow rates and their relative proportions, the standard deviation of packet interarrival times, buffer capacity, load presented to the buffer, and service time distribution. Notably, we focus on non-Poisson flow models due to their relevance in real-world scenarios.

Previous research on burst ratios has predominantly focused on multiplexed traffic. However, our analysis underscores the importance of understanding burst ratios at the individual flow level. Indeed, the QoE for an end user engaging with multimedia content is more directly influenced by the burst ratio within the specific multimedia flow they are accessing, rather than the burst ratio in the aggregated network traffic.

It is challenging to ascertain whether the burst ratio for a specific flow will be lesser, greater, or equivalent to the burst ratio of multiplexed traffic, due to the interplay of two opposing phenomena. On one hand, a loss sequence within multiplexed traffic might comprise losses from multiple flows, potentially resulting in shorter loss subsequences within individual flows. Conversely, it is feasible for a flow to experience a sequence of packet losses even in the absence of such a sequence in the multiplexed traffic. For illustration, consider a sequence of eight packets from two distinct flows, alternating as 0 1 0 1 0 1 0 1. In this scenario, the first flow (odd packets) exhibits a sequence of four consecutive losses, which is not evident in the multiplexed traffic.

Addressing these complexities, our research into individual burst ratios is carried out using real-life network equipment and traffic in a controlled laboratory environment. This approach enables the direct observation and analysis of network behavior under various conditions, providing empirical data that closely represents real-world scenarios. Our experimental setup processes vast quantities of network packets, ensuring robustness in our results by significantly reducing the impact of random variations. This methodology also allows for the exploration of different types of interarrival time distributions within each flow, ensuring a comprehensive and realistic examination of network dynamics.

The structure of this paper is organized into six distinct sections. Section 2 presents an overview of the relevant literature and previous studies. Section 3 introduces the laboratory equipment employed. Section 4 revisits course of the experiments. Subsequently, Section 5 unveils and examines new findings pertaining to per-flow burst ratios under various conditions. The paper concludes with Section 6, summarizing the key conclusions drawn from the research.

II. Related work

In this section, we provide an overview of the existing literature and methodologies pertinent to the measurement of the per-flow burst ratio parameter in IP networks.

A. Early Theoretical Models and Their Limitations

Initial studies in the domain of packet loss and burst ratio heavily relied on stochastic processes, particularly Markov chains, to model packet loss without directly accounting for queuing mechanisms [4]–[11]. These ”black box” models, while foundational, lacked the direct representation of a queue with a limited buffer, a critical element in real-world network environments. The two-state Markov model [4] and its extensions, such as the Gilbert model [5]–[7] and the Gilbert-Elliott model [8] [9], provided a basic framework but fell short in accurately mimicking the statistical structure of consecutive losses caused by queuing and buffer overflow. The limitations of these early models, particularly in capturing the variance in loss sequences, are discussed in [12]. These models could only approximate the average length of loss sequences, not their large variances, which are crucial in realistic network scenarios.

B. Transition to Empirical Measurement-Based Approaches

Recognizing the inadequacies of purely theoretical models, recent research has shifted towards empirical measurements and more realistic modeling of packet loss causes, such as queuing and buffer overflows. This transition is evident in studies that have extended the analysis to various Poisson stream configurations [13]–[15] and empirical investigations in controlled lab settings and real-world network environments [2] [3]. The loss ratio (L), a fundamental characteristic of the packet loss process, has been extensively studied through both direct network measurements [16]–[20] and analytical formulas in queue models with finite buffers [13]–[15]. These studies have provided a more nuanced understanding of the loss process in network traffic.

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

C. Broader Perspectives and Advanced Models

In addition to the burst ratio, other metrics for loss characterization have been explored. Studies like [21] and [22] have investigated the probability distributions of loss sequences and the lengths of initial loss sequences, offering alternative perspectives on packet loss in networks. Advanced models, such as the four-state Markov chain [10] and the general k-state Markov chain model [4] [7] [11], have been developed to allow for a more detailed analysis of loss patterns, including the loss of k consecutive packets in specific states.

D. Active Queue Management and Burst Ratio Mitigation

To address the burst ratio in TCP/IP networks, research has focused on Active Queue Management (AQM). A variety of AQM methods have been proposed [23] [24]. Particularly, algorithms based on the dropping function [25] have shown promise in reducing the burst ratio, as demonstrated in experimental studies [26]–[28].

E. Burst Ratio in Aggregated Traffic

The study of burst ratio extends beyond individual packet flows to encompass aggregated traffic, a critical aspect in understanding network behavior on a larger scale. In aggregated traffic scenarios, the burst ratio provides insights into the collective behavior of multiple data streams converging within a network. This approach is essential for comprehending the dynamics of network congestion and the resultant packet loss patterns. Research in this area often employs complex models that simulate the interaction of multiple traffic flows, thereby offering a more comprehensive view of network performance under varying load conditions. Key studies in this domain have focused on the impact of burst ratio in aggregated traffic [29]–[32].

F. Analytical Approaches to Burst Ratio in Individual Network Flows

In contrast to empirical measurement-based studies, analytical approaches to understanding the burst ratio in individual network flows have also been prominent. These approaches typically involve developing mathematical models that can predict packet loss behavior in a single flow, considering various factors such as traffic intensity, buffer size, and service policies. Such models are invaluable for theoretical analysis and for designing network systems that can efficiently handle expected traffic patterns. Significant contributions in this area have included the development of formulas and algorithms that accurately predict the burst ratio in individual flows, taking into account the unique characteristics of each flow [33] [34].

III. Laboratory equipment

In order to conduct measurements of the burst ratio per flow in a network laboratory, a test network comprising three main components was established:

1. **Spirent SPT-N4U Traffic Generator**: This device, equipped with the MX2-10G-S12 module, features 12 fiber optic ports, each capable of generating traffic at maximum speeds of 1 Gb/s or 10 Gb/s. It is adept at generating substantial volumes of stateful and stateless traffic (TCP and UDP) across all ports independently, without degrading the internal performance of the device.

2. **Test Device - Cisco 3750X Layer 3 Router/Switch**: This device, fitted with 12 ports of 1 Gb/s and two ports of 10 Gb/s, is utilized for observing actual packet losses. Its role is critical in the precise measurement of the burst ratio per flow in various network conditions.

3. **High-Performance HP Enterprise ProLiant DL380 Gen9 Servers**: These servers are employed for receiving and storing both incoming and outgoing traffic in databases, and for executing packet loss analysis, with a specific focus on per-flow burst ratio measurements.

The Spirent SPT-N4U, functioning as both a traffic generator and analyzer, is the primary device used in this setup. This generator proved indispensable for generating high traffic loads, which are challenging to achieve using standard computers. It offers a complete, portable environment for network operations and performance testing of complex topologies. The generator's chassis facilitates the connection of various load modules, allowing potential traffic generation up to 400 Gb/s, with precision up to 10 ns per generated or analyzed frame. It supports an integrated test controller (client-server type application) that manages the entire system.

The load modules facilitate processing of necessary signals for testing both standard transmission and voice/video transmissions, as well as applications from layers 2 to 7. The load module used in the experiments, the MX2-10G-S12, has 12 fiber optic ports operating at 1 Gb/s or 10 Gb/s speeds as required. Each port supports packet generation and analysis at layer 2 and 3 at cable speed (maximum throughput of the used physical layer), along with efficient emulation of routing and switching protocols. Each port contains a separate RISC processor under Linux OS control, with an optimized TCP/IP test stack.

In the experiments, the ports of the generator configured in the Test Center application are treated as separate instances of the Linux OS. This ensures that each interface is tested independently, with the status of the respective instance relayed to the supervising machine. Each port configured in the application can operate in dual mode - as both generator and analyzer. During traffic analysis, one can monitor the total throughput of received and generated traffic (in frames/packets or bits), packet contents, average delay (as well as minimum and maximum values), analyze sequencing mechanisms, and construct histograms based on different parameters of the experiment and apply various filters for analyzing only specific packets or their groups.

The test device used in the experiments, a Cisco 3750X layer 3 switch/router, was equipped with 12 ports of 1 Gb/s using multimode SFP transceiver modules and two multimode SFP+ transceiver modules operating at 10 Gb/s. Its switching matrix performance of 160 Gb/s and 35.7 million packets per second ensured that no additional delay would be introduced into queuing mechanisms. The Cisco 3750X supports both

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

standard packet sizes and Jumbo frames up to 9216B. Additionally, it allows for stacking (combining several physical devices into one logical device) as needed, providing up to 468 total ports.

Direct measurement of the burst ratio per flow using the Spirent SPT-N4U traffic generator and analyzer is not feasible. Network traffic must be captured, transmitted to servers, and stored in databases for subsequent calculation of packet loss characteristics. SQLite was chosen for data storage due to its support on the Spirent SPT-N4U device, ease of use, and simple installation on computational servers. SQLite is a well-known database management system and library written in C, implementing a SQL (Structured Query Language) supporting engine. It implements an SQL engine that allows database use without a separate process of a Relational Database Management System (RDBMS). Due to the requirement to store traffic records from each experiment in a separate database, SQLite was the most practical solution.

During the experiments, three variants of the database were created. The first database, created on the first server, stored a table with source packets. It contained integers describing the following elements of packet headers: timestamps, source and destination IP addresses, IP protocol numbers, IP identifiers, IP packet lengths, source and destination ports, and sequence and acknowledgment numbers for TCP flows.

The second database, created on the second server, stored a table with destination packets with the same fields as in the first database. The third database was created on the third server. It was a combination of the two tables of both previous databases. It was theoretically possible to use only 2 servers, e.g., by copying the first table to the second server, but the use of a third server allowed for speeding up the measurement collection process by performing calculations in parallel for the previous scenario and capturing traffic for the next one. While the third server was still calculating the burst ratio per flow for the previous scenario, the first and second servers were already collecting new traffic in the next scenario.

IV. Course of the Experiment

The test network established for measuring the burst ratio is depicted in Figure 1. Traffic was generated on ports 1 to 4 of the Spirent SPT-N4U generator and received on 1 Gb/s ports numbered 1, 3, 5, and 7 of the Cisco 3750X device. The generated traffic was based on the most common protocol stack: Ethernet at Layer 2, IPv4 at Layer 3, and TCP/UDP at Layer 4. The default TCP congestion control of Spirent, namely New Reno with a maximum window size of 32768B, was employed. At Layer 7, HTTP traffic was emulated, with an Apache WWW server on port 80 on the server side and a Mozilla browser on the client side.

Within the Cisco 3750X device, all incoming traffic was duplicated using the SPAN function to output ports 9 and 13. Port 13 was used to send a copy of the incoming traffic to server 1 for further analysis, while port 9 was the test port — this port experienced actual packet losses (specifically in its buffer). Subsequently, the outgoing traffic from this port was looped back to port 8 and duplicated within the test device (again using the SPAN function) to ports 10 and 14. Finally, traffic from port 10 was sent back to the Spirent analyzer, operating as the packet destination (and stateful client for each TCP connection), while traffic from port 14 was sent to server 2 for further analysis.

On the first two servers, the DPDK-dump application [35] was used for data capture. The data from each packet was removed, and a header with a timestamp and index was stored in a SQLite database. The preliminarily processed packets were then copied to a third server, where databases containing the input and output packets from the test device were merged. If possible, each outgoing packet from the Layer 3 switch was paired with its corresponding input packet. If an input packet did not have a corresponding output packet, it was considered to have been removed from the overflowed buffer — lost in network transmission.

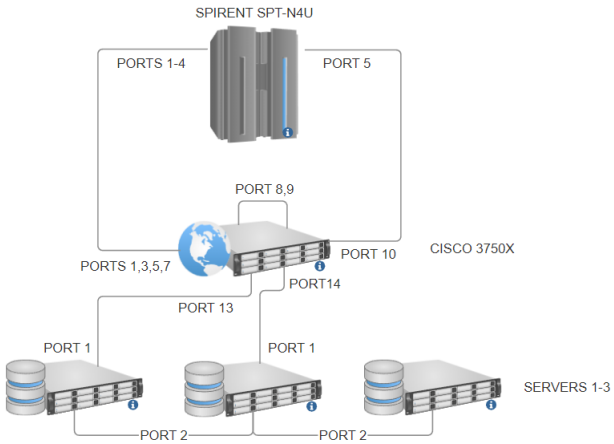

Figure 1. Test network.

As shown in Figure 1, the test traffic was transmitted on six links with up to 1 Gb/s capacity, while additional data for calculating the burst ratio and other network traffic parameters were transmitted through a 10 Gb/s LAN network.

V. Results and Discussion

In the experiment, the burst ratio was measured across multiple scenarios featuring variable traffic characteristics, different number of TCP flows and different buffer sizes. In each scenario, approximately two million packets were generated. UDP traffic was employed as background traffic, constituting 5% of the total link load, unless specified otherwise in the scenario. It is important to note that in all scenarios incorporating at least one TCP connection, it was not feasible to manually set any arbitrary value for the test link load (ρ). The load is automatically adjusted by the New Reno algorithm operating in each TCP sender.

In the initial scenarios, an identical distribution of flows was generated, with the total bandwidth divided among a varying number of flows, specifically: 24, 32, 48, 64, 96, 128, and 192 flows. ρ in each scenario stabilized at a value

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

of 0.944 Mbps. As hypothesized, the burst ratio exhibited a slight increase with the rising number of TCP flows. In the latter scenarios, the measured values of the loss ratio were comparable, leading to the conclusion that beyond a certain threshold of TCP connections, further increases do not significantly affect the burst ratio. Similarly, the per-flow burst ratios were consistently lower than B in all cases. Furthermore, as the number of flows increased, each individual burst ratio (B_i) decreased, indicating a dilutive effect of increased flow counts on the individual burst ratio metrics.

To investigate the impact of interarrival time on both the global burst ratio and the per-flow burst ratios, a varying percentage of UDP traffic within the total link bandwidth was utilized (from 0 to 20%). The increasing proportion of small datagrams notably influenced the rise in variance of the packet service time distribution in the queue. This, in turn, led to an increase in the global burst ratio (B), as well as an enhancement of the per-flow burst ratio (B_i). However, due to the identical characteristics of the flows within a given scenario, B_i remained consistent across different scenarios. Again, the burst ratio for individual flows (B_i) was less than the multiplexed burst ratio (B) for each i.

In the third group of scenarios, the buffer size of the examined port was varied from 50 to 1000 packets, while maintaining 128 TCP connections and 50 Mb/s of UDP traffic. The variable buffer size on the tested port did not result in significant changes in the burst ratio, leading to the conclusion that the burst ratio is weakly dependent on the size of the queue buffer in network transmissions. Similarly, the variable buffer size in selected scenarios did not affect the change in the per-flow burst ratio in relation to the baseline scenario.

In all previous scenario groups, despite differences in aggregated traffic, very similar per-flow traffic characteristics were created. To introduce diversity at the level of flow characteristics, the background traffic (UDP) was modified for only a portion of the flows (25%) while it remained original (5%) for the others. This adjustment enabled the observation of scenarios where flows, even within connection-oriented protocols (TCP used in the scenario), could be compared with varying characteristics, specifically different interarrival times (simulated by varying proportions of small datagrams). Subsequently, in the following scenarios, the UDP traffic was altered for 50% and 75% of the flows. These scenarios yielded interesting results: for most of the scenarios, all per-flow burst ratios were significantly smaller than the global burst ratio (B), and not very different from each other, despite the variances of the flows differing between groups of flows. However, for scenarios with a significant variance in packet interarrival times, which also constituted a substantial portion of the total network load, instances of B_i exceeding B were recorded, highlighting the impact of varied flow characteristics on the burst ratio parameter.

The conclusions drawn from each group of scenarios are summarized in Table I.

TABLE I
SCENARIOS CONCLUSIONS

Scenario	Aggregated burst ratio	Per-flow burst ratios
24 TCP Flows	1.30	All 1.06
32 TCP Flows	1.30	All 1.06
48 TCP Flows	1.34	All 1.06
64 TCP Flows	1.35	All 1.05
96 TCP Flows	1.35	All 1.04
128 TCP Flows	1.35	All 1.04
192 TCP Flows	1.36	All 1.03
UDP 0%	1.23	All 1.02
UDP 5%	1.35	All 1.04
UDP 10%	1.43	All 1.05
UDP 15%	1.46	All 1.05
UDP 20%	1.52	All 1.06
Buffer 50	1.35	All 1.04
Buffer 100	1.35	All 1.04
Buffer 200	1.35	All 1.04
Buffer 500	1.34	All 1.03
Buffer 1000	1.35	All 1.04
UDP 5%, 25% flows	1.35	With UDP 1.07, Without UDP 1.02
UDP 10%, 25% flows	1.43	With UDP 1.07, Without UDP 1.02
UDP 15%, 25% flows	1.46	With UDP 1.08, Without UDP 1.02
UDP 20%, 25% flows	1.52	With UDP 1.09, Without UDP 1.02
UDP 5%, 50% flows	1.35	With UDP 1.10, Without UDP 1.01
UDP 10%, 50% flows	1.43	With UDP 1.11, Without UDP 1.01
UDP 15%, 50% flows	1.46	With UDP 1.14, Without UDP 1.01
UDP 20%, 50% flows	1.52	With UDP 1.15, Without UDP 1.01
UDP 5%, 75% flows	1.35	With UDP 1.27, Without UDP 1.01
UDP 10%, 75% flows	1.43	With UDP 1.35, Without UDP 1.01
UDP 15%, 75% flows	1.46	With UDP 1.47, Without UDP 1.01
UDP 20%, 75% flows	1.52	With UDP 1.59, Without UDP 1.01

VI. CONCLUSIONS AND FUTURE WORK

In this study, we conducted a detailed analysis of the burst ratio on a per-flow basis within a network environment, where each packet flow through a network node was evaluated independently. The burst ratio is a critical metric that quantifies the tendency of packet losses to occur in extended, consecutive sequences, which can adversely affect the quality of service for multimedia streams, especially those requiring real-time transmission.

Empirical measurements were carried out using sophisticated laboratory equipment to ascertain the burst ratios of individual flows. These measurements revealed that the burst ratio for a specific flow could be lower, higher, or equivalent to the aggregate burst ratio, contingent upon the characteristics of the flows involved.

Increasing TCP Flows: The global burst ratio slightly increased with the number of TCP flows, indicating a minor

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

degradation in the network's ability to handle packet losses in a bursty manner as the number of flows increases. Notably, the per-flow burst ratios also exhibited variations, but the impact plateaued beyond a certain number of TCP connections, suggesting a threshold beyond which additional TCP flows do not exacerbate burst behavior significantly.

UDP Traffic Proportion Influence: Adjusting the UDP traffic percentage within the total bandwidth affected both the global and per-flow burst ratios. The presence of small datagrams increased the service time variance, elevating both global and per-flow burst ratios. This highlights the sensitivity of burst ratios to the composition and characteristics of network traffic.

Buffer Size Variation Effects: Changes in buffer size from 50 to 1000 packets showed minimal impact on both global and per-flow burst ratios, indicating a low dependency of burst loss behavior on queue buffer size. This suggests that other factors, beyond buffer capacity, play a more significant role in influencing burst loss patterns.

Differentiated Background Traffic: Introducing variations in background traffic for certain flows altered the burst ratio landscape, with modifications affecting both the global and per-flow metrics. The approach allowed for distinguishing the effects of flow-specific characteristics, particularly interarrival times, on burst loss behavior, underscoring the importance of individual flow properties in network performance analysis.

Varied Interarrival Times: The scenarios with adjusted background traffic for varying proportions of flows highlighted the differential impact on global and per-flow burst ratios. Significant variances in packet interarrival times, especially in flows that constituted a large portion of the network load, were associated with instances of per-flow burst ratios exceeding the global burst ratio. This observation illustrates the complex relationship between traffic diversity, flow characteristics, and their collective impact on network burst loss dynamics.

These refined conclusions emphasize the importance of considering both global and per-flow perspectives to fully understand the intricacies of network traffic management and its implications on performance metrics like burst ratio.

Our forthcoming research endeavors will extend our experimental setup to encompass real multimedia streaming traffic and IPv6, aligning with current and progressive networking contexts. We also plan to explore the effects of ON-OFF patterns on Quality of Experience and provide a comprehensive description of traffic assumptions and flow characteristics to facilitate the replication of our experiments. Additionally, we will focus on measuring higher-order statistics from traffic using a methodology similar to that employed for burst ratio analysis. These efforts aim to robustly test the impact of various network conditions and expand the scope and applicability of our research, ensuring its relevance in the study of network behavior and its influence on multimedia content delivery.

ACKNOWLEDGMENT

This research was funded by National Science Centre, Poland, grant number 2020/39/B/ST6/00224.

REFERENCES

[1] J. W. McGowan, "Burst ratio: a measure of bursty loss on packet-based networks", 16 2005. US Patent 6,931,017, 2005.

[2] D. Samociuk, A. Chydziński, and M. Barczyk, "Experimental measurements of the packet burst ratio parameter", Proc. BDAS 2018, pp. 455-466, 2018.

[3] D. Samociuk, M. Barczyk, and A. Chydziński, "Measuring and analyzing the burst ratio in IP traffic", Proc.BROADNETS, pp. 86-101, 2019.

[4] M. Yajnik, S. Moon, J. Kurose, and D. Towsley, "Measurement and modelling of the temporal dependence in packet loss", Proc. of IEEE-INFOCOM, vol. 1, pp. 345-352, 1999.

[5] E. N. Gilbert, "Capacity of a Burst-Noise Channel", Bell system technical journal, vol. 39(5), pp. 1253–1265, 1960.

[6] M. Veeraraghavan, N. Cocker, and T. Moors, "Support of voice services in IEEE 802.11 wireless LANs", Proc. of IEEE INFOCOM, vol. 1, pp. 488-497, 2001.

[7] W. Jiang and H. Schulzrinne, "Modeling of packet loss and delay and their effect on real-time multimedia service quality", Proc. of NOSSDAV, pp. 1–10, 2000.

[8] E. O. Elliott, "Estimates of Error Rates for Codes on Burst-Noise Channels", Bell system technical journal, vol. 42(5), pp. 1977-1997, 1963.

[9] G. Haßlinger and O. Hohlfeld, "The Gilbert-Elliott model for packet loss in real time services on the Internet", Proc. of Measuring, Modelling and Evaluation of Computer and Communication Systems Conference, pp. 1-15, 2008.

[10] A. Clark, "Modeling the effects of burst packet loss and recency on subjective voice quality", Proc. if Internet Telephony Workshop, pp. 123- 127, 2001.

[11] H. A. Sanneck and G. Carle, "Framework model for packet loss metrics based on loss runlengths", Proc. SPIE 3969, Multimedia Computing and Networking 2000, pp. 1-11, 2000.

[12] X. Yu, J.W. Modestino, and X. Tian, "The accuracy of Gilbert models in predicting packet-loss statistics for a single-multiplexer network model", Proc. of IEEE INFOCOM'05, pp. 2602–2612, 2005.

[13] H. Takagi, "Queueing analysis - Finite Systems", North-Holland Amsterdam, 1993.

[14] A. Chydzinski, R. Wojcicki, and G. Hryn, "On the Number of Losses in an MMPP Queue", Lecture Notes in Computer Science, vol. 4712, pp. 38-48, 2007.

[15] A. Chydzinski and B. Adamczyk, "Transient and stationary losses in a finite-buffer queue with batch arrivals", Mathematical Problems in Engineering, vol. 2012, ID 326830, pp. 1-17, 2012.

[16] P. Benko and A. Veres, "A passive method for estimating end-to-end TCP packet loss", Proc. of IEEE GLOBECOM'02, pp. 2609–2613, 2002.

[17] Bolot J. "End-to-end packet delay and loss behavior in the Internet", Proc of ACM SIGCOMM'93, pp. 289–298, 1993.

[18] M. Coates and R. Nowak, "Network loss inference using unicast end-toend measurement", Proc. of ITC Conference on IP Traffic, Measurement and Modeling, pp. 282–289, 2000.

[19] N. Duffield, F. L. Presti, V. Paxson, and D. Towsley, "Inferring link loss using striped unicast probes", Proc. of IEEE INFOCOM'01, pp. 915–923, 2001.

[20] J. Sommers, P. Barford, N. Duffield, and A. Ron, "Improving accuracy in end-to-end packet loss measurement", ACM SIGCOMM Computer Communication Review, vol. 35(4), pp. 157–168, 2005.

[21] M. Bratiychuk and A. Chydzinski, "On the loss process in a batch arrival queue", Applied Mathematical Modelling, Vol. 33, Iss. 9, pp. 3565-3577, 2009.

[22] A. Chydzinski, "On the remaining service time upon reaching a target level in M/G/1 queues", Queueing Systems, vol. 47, issue 1/2, pp. 71–80, 2004.

[23] A. Chydzinski, "Buffer Overflow Period in an MAP Queue. Mathematical", Problems in Engineering, vol. 2007, pp. 1–18, 2007.

[24] A. Chydzinski, "Buffer overflow period in a batch-arrival queue with autocorrelated arrivals", Applied Mathematics Information Sciences 7, no. 4, pp. 1633–1641, 2013.

[25] A. Chydzinski, M. Barczyk, and D. Samociuk, "The Single-Server Queue with the Dropping Function and Infinite Buffer", Mathematical Problems in Engineering, vol. 2018, ID 3260428, pp. 1–12, 2018.

[26] M.Barczyk and A. Chydzinski, "Experimental testing of the performance of packet dropping schemes", Proc. IEEE ISCC, pp. 1-7, 2020.

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

[27] M. Barczyk and A. Chydzinski, ”AQM based on the queue length: A real-network study”, PLoS ONE 17(2): e0263407, 2022.
[28] D. Samociuk and A. Chydzinski, ”On the impact of the dropping function on the packet queueing performance”, Proc. of International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2018, pp. 473-478, 2018.
[29] A. Chydzinski and D. Samociuk, ”Burst ratio in a single-server queue”, Telecommunication Systems, DOI: 10.1007/s11235-018-0476-7, 2018.
[30] A. Chydzinski, D. Samociuk, and B. Adamczyk, ”Burst ratio in the finite-buffer queue with batch Poisson arrivals”, Applied Mathematics and Computation 330, pp. 225-238, DOI: 10.1016/j.amc.2018.02.021, 2018.
[31] D. Samociuk, A. Chydzinski, and M. Barczyk, ”Experimental measurements of the packet burst ratio parameter”, Communications in Computer and Information Science, Springer, Volume 928, pp. 455-466, 2018.
[32] D. Samociuk, M. Barczyk, and A. Chydzinski, ”Measuring and analyzing the burst ratio in IP traffic”, Communications in Computer and Information Science, Springer, Volume 928, 2018.
[33] A. Chydzinski, ”Per-flow structure of losses in a finite-buffer queue”, Applied Mathematics and Computation 428, 127215, pp. 1-15, May 2022.
[34] A. Chydzinski and B. Adamczyk, ”Burst ratio of packet losses in individual network flows”, INFORMATICA, Volume 34, pp. 35–52 February 2023.
[35] DPDK documentation website, URL: https://www.dpdk.org/, visited on 01.02.2024.

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

Localization of Mobile Devices in Future Wireless Networks

Caleb Ludinga Lodi
Department of Computer Science and Software Engineering
Laval University
Quebec, Canada
e-mail: caleb.lodi.1@ulaval.ca

Ronald Beaubrun
Department of Computer Science and Software Engineering
Laval University
Quebec, Canada
e-mail: ronald.beaubrun@ift.ulaval.ca

***Abstract*— High-precision localization is essential for fully realizing the potential of future mobile networks (6G). A critical factor in achieving such localization accuracy is the incorporation of Terahertz (THz) bands into an efficient localization technique. In this context, we propose an approach for localizing mobile devices in 6G mobile networks. Such an approach is based on the Angle of Departure (AoD) localization technique in order to provide high localization accuracy. Using NYUSIM 4.0 for implementing it, we perform simulations for two different scenarios: the first one emulates a 5G mobile network in the mmWave bands, whereas the second one emulates a 6G mobile network in the THz frequency bands. Results show that 6G mobile networks outperform 5G mobile networks in terms of localization accuracy. This improvement is attributed to finer channel estimation facilitated by THz frequencies, which results in enhanced positioning accuracy.**

***Keywords*— 6G; accuracy; Angle of Departure (AoD); future mobile networks; localization technique.**

I. INTRODUCTION

Future wireless technology will offer intelligent and ubiquitous connectivity with Terabits-per-second (Tbps) data rates and sub-millisecond (sub-ms) latency over three-Dimensional (3D) network coverage [1]. In order to achieve such goals, accurate localization information of mobile devices will be essential [2]. In fact, determining the position of mobile devices with high degree of precision is not only crucial for navigation and location-based services, but also for enabling a plethora of emerging applications, such as intelligent telesurgery, holographic teleportation, connected robotics, Augmented Reality (AR) and more [3].

Moreover, higher positioning accuracy is in high demand in many vertical and industrial applications, particularly in indoor environments where satellite-based positioning systems are ineffective. In this context, the introduction of 6G mobile networks will improve high precision positioning within a home or an office. According to [4], with the application of THz bands, such networks are expected to achieve a 3D localization precision of 1 cm. Figure 1 provides a visual representation of how localization will be used in 6G mobile networks.

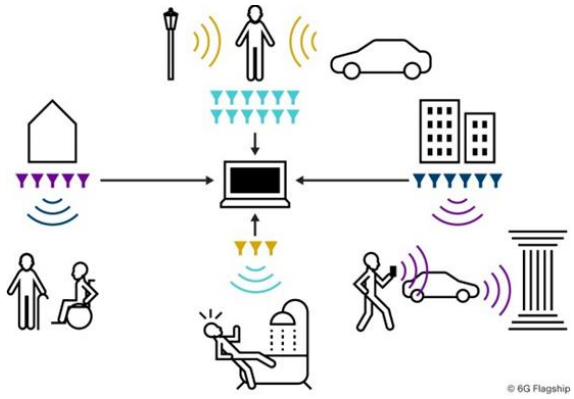

Figure 1. Localization illustration in future wireless networks [5].

The importance of accurate localization gained more attention after the U.S. Federal Communications Commission implemented the enhanced 911 (FCC-E911) rules [6]. Additionally, with the advent of the Global Positioning System (GPS) and the standardization of cellular communication systems, it is now possible to achieve an accuracy of 10 cm in rural areas and 1 m in outdoor urban environments approximately [7]. However, accurately determining localization in indoor environments using cellular networks and GPS can be challenging, as signals are reflected and dispersed by various objects [8]. To overcome such challenging, WiFi-based and Bluetooth-based localization methods have been developed [6]. In complex indoor environments, multi-path signal components and signal blockage can negatively impact localization accuracy, but the use of ultra-wide bandwidth (such as terahertz bands) can make multi-path signals resolvable and improve performance [8].

A number of papers on 6G localization have recently been published [2][9]-[14]. More specifically, the works from [2][13], and [14] were primarily concerned with the vision and technology requirements of 6G localization. Meanwhile, Chen et al. [6] summarize the most recent literature findings related to the use of THz wireless systems for accurate localization. Although this significant work reveals key concepts about THz communication systems and localization, it does not propose transformative solutions required to effectively deploy realistic THz localization systems.

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

Furthermore, while some of these works [11][12] acknowledge the importance of THz localization capability, they fail to highlight the challenges and prospective techniques required to deploy THz systems capable of performing accurate localization in a real-world scenario.

In this paper, we propose an approach for localization of mobile devices in future wireless networks. It is organized as follows. Section II analyses various localization techniques. Section III provides an overview of localization systems and services that are currently in use. Section IV provides an overview of 6G-compatible simulation tools. Section V presents the basic principles of the proposed approach for localizing mobile devices in future mobile networks. Section VI explains the simulation environment, as well as the parameters needed to perform the simulations, and presents simulation results, whereas Section VII presents concluding remarks.

II. LOCALIZATION TECHNIQUES

Localization techniques are employed to estimate locations of mobile devices using readings from reference points. This section provides an overview of well-known positioning techniques.

A. Received Signal Strength Indicator (RSSI)

The Received Signal Strength (RSS)-based approach is a simple and widely used method for indoor localization [15]. It relies on the measurement of RSS, which represents the power of the signal received by the receiver in decibel-milliwatts (dBm) or milliwatts (mW). By analyzing the RSS, the distance between a Transmitter (TX) and a Receiver (RX) can be estimated, where a higher RSS value corresponds to a shorter distance between TX and RX. The absolute distance can be determined using various signal propagation models, provided that the transmission power is known. RSSI, which is a relative measurement of the RSS with arbitrary units, can be expressed as follows [16]:

$$RSSI = -10n \log_{10}(d) + A \quad (1)$$

where d is the distance between TX and RX, n the path loss exponent of the signal attenuation and A represents the RSSI value at a reference distance from RX.

From (1) and a simple path-loss propagation model, the separation distance d between TX and RX can be expressed as follows:

$$d = 10^{\frac{(A-RSSI)}{10n}} \quad (2)$$

where n is the path loss, and A the RSSI value at a predefined distance from RX. Figure 2 illustrates a network topology consisting of three Root Nodes (RN_1, RN_2, RN_3) which are commonly referred to as base stations. The primary objective of this configuration is to demonstrate the process of localizing a mobile device within the network, using RSSI and applying the triangulation technique. Each of the three base stations serves as a reference point with known geographical coordinates.

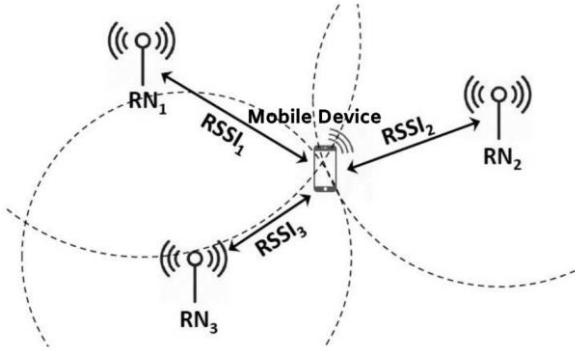

Figure 2. Localization based on RSSI [16].

The mobile device location is to be determined based on the signal strength measurements received from these base stations ($RSSI_1$, $RSSI_2$, $RSSI_3$). The RSSI values from each base station provide an indication of the signal's attenuation due to distance and obstacles. By analyzing the RSSI values received at the mobile device, the network can estimate the relative distances between the mobile device and each base station.

Triangulation is then employed to calculate the mobile device position. This technique involves drawing circles centered at each base station, with the radius of each circle determined by the estimated distance to the mobile device. The intersection points of these circles represent the potential mobile device locations, as the technique uses the known positions of the base stations (RN_1, RN_2, RN_3) and the estimated distances from each base station to refine the mobile device location.

B. Fingerprinting

Fingerprinting, also known as scene analysis-based localization techniques, is used to determine the location of a mobile device by comparing received signal characteristics, known as fingerprints, to a pre-existing database of fingerprints associated with known locations [17]. This method relies on the fact that signal propagation and behavior vary in different physical environments, creating unique patterns or fingerprints for each location.

To create a fingerprint database, measurements of signal characteristics are collected from various points within the environment. For each location where measurements are taken, a fingerprint is generated. This fingerprint represents a unique signature of the wireless signal behavior in that specific location. The fingerprint includes a set of signal parameter values that can be used to identify that particular location. The collected fingerprints, along with their corresponding known locations, are stored in a database. When a user needs to be localized, it measures the same signal characteristics (RSSI) at its current location. These measurements are compared with the fingerprints stored in the database. The system identifies the stored fingerprint that closely matches the measured values. Once a matching fingerprint is found, the associated known location from the

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

database is used to estimate the device position. This can be done through probabilistic methods, artificial neural networks, k-Nearest Neighbor, and Support Vector Machine [16].

The accuracy of fingerprinting depends on the grid size defined during a training period, with a tradeoff between accuracy and complexity. Finer grids offer higher accuracy, but require longer training periods [18]. Fingerprinting localization can provide high accuracy in scenarios with well-defined and stable environments. While fingerprinting has been traditionally associated with Wi-Fi due to the widespread availability of signals in typical indoor environments [19], it has also been utilized in 4G [20] and 5G [21] mobile networks. In the future, the increased densification of base stations in 6G networks will enhance the resolution of fingerprinting-based localization [22].

Despite its advantages in terms of accuracy and low infrastructure investment, fingerprinting has some disadvantages. It is sensitive to environmental changes [23]. A major limitation is the requirement for complex training procedures, which limits its applicability in scenarios where prior exploration is not possible or when covering large areas [22].

C. Time of Arrival (ToA)

Time of Arrival (ToA) or *Time of Flight* (ToF) is a localization technique that uses signal propagation time to calculate the distance between a transmitter TX and a receiver RX [24]. Figure 3 illustrates the application of ToA technique for determining the position of a mobile device within a mobile network. The depicted scenario involves three root nodes, also referred to as base stations (BS), labeled as RN_1, RN_2 and RN_3. As the signal reaches each base station, it is received at slightly different times due to the varying distances between the mobile device and the base stations. Figure 3 indicates these arrival times as t_1, t_2, and t_3 for RN_1, RN_2, and RN_3 respectively. The distances between the mobile device and each base station are labeled as d_1, d_2, and d_3.

To illustrate the calculation of distances using ToA, consider a transmitter TX_i and a receiver RX_j. Suppose TX_i sends a message at time t_i, which is received by RX_j at time t_j. The time taken for the signal to travel from TX_i to RX_j is denoted as t_p, where $t_j = t_i + t_p$.

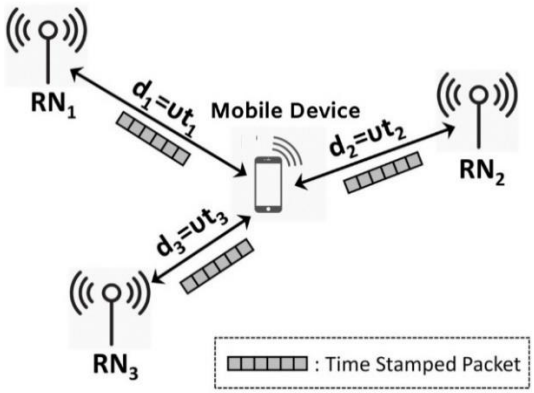

Figure 3. Localization based on ToA [16].

Therefore, the distance between TX_i and RX_j can be calculated as follows:

$$D_{ij} = (t_j - t_i) * v \quad (3)$$

where v represents the signal velocity.

D. Return Time of Arrival (RToA)

Return Time of Arrival (RToA), also known as *Return Time of Flight* (RToF), uses the time taken for a signal to propagate from TX to RX, and back to TX, to estimate the distance [24]. Similar to ToA, RToA involves a two-way signal exchange where the receiver responds to the transmitter's signal, allowing the calculation of the total round-trip time. One advantage of RToA is that it requires relatively moderate clock synchronization between TX and RX compared to ToA [16]. However, the accuracy of RToA estimation is influenced by the same factors as ToA, such as sampling rate and signal bandwidth, which can have a more significant impact since the signal is transmitted and received twice [16].

One challenge with RToA-based systems is the response delay at the receiver, which is dependent on the receiver electronics and protocol overhead. Although this delay can be disregarded in long-range systems where the propagation time outweighs the response time, it becomes significant in short-range applications, like indoor localization.

E. Angle of Arrival (AoA)

Angle of Arrival (*AoA*)-based methods use antenna arrays to estimate the angle at which the transmitted signal arrives at the receiver [6][16]. This is achieved by calculating the time difference of arrival at individual elements of the antenna array [25]. The primary advantage of AoA is its ability to estimate the location of a mobile device with only two monitors in a 2D environment or three monitors in a 3D environment [16]. Figure 4 illustrates how AoA can be used to estimate a mobile device location based on the angles at which signals are received by the antenna array. Transmitter TX emits a signal that propagates through the air towards receiver RX, equipped with an antenna array. From Figure 4, we can determine d, which represents the separation distance between TX and RX. Such a distance plays a critical role in determining the AoA at which the signal arrives at RX. This angle is denoted as θ.

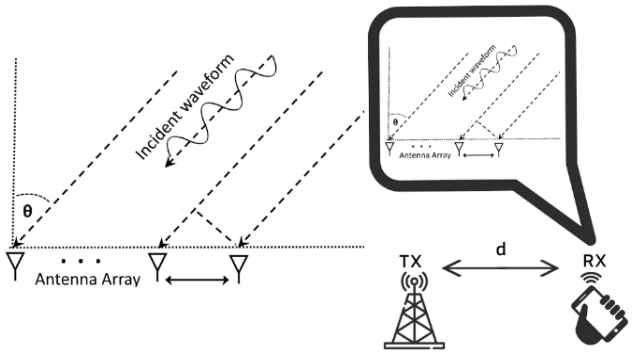

Figure 4. Localization based on AoA [16].

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

F. Phase of Arrival (PoA)

Phase of Arrival (PoA) based approaches estimate distance using the phase of the carrier signal [24]. They can be combined with other techniques for better localization accuracy [16]. However, they rely on line-of-sight, which is often unavailable indoors [16]. Figure 5 demonstrates the application of PoA technique in estimating the user location by analyzing the phase difference of signals received by an antenna array. It shows that when incident signals reach different antennas in an antenna array, they exhibit a phase difference. This phase difference can be used to derive the user location.

Signals emitted by the transmitters propagate through the mobile network medium and reach the mobile device. Due to the varying distances between the transmitters and the mobile device, the signals experience different phase shifts. The receiver measures the phase difference between the received signals from different transmitters. This phase difference reflects the relative time delay that the signals experienced due to their different propagation distances. As the phase difference increases or decreases, the corresponding distance between the transmitter and the mobile device changes. By comparing the measured phase differences with a reference phase, the system can estimate the distances between the mobile device and each transmitter (TX). Using the estimated distances from multiple transmitters, the user position can be determined through multilateration techniques.

G. Angle of Departure (AoD)

The Angle of Departure (AoD) localization technique is a method used to estimate the position of a mobile device by analyzing the angles at which signals are transmitted from the mobile device [26]. AoD is particularly relevant in scenarios where a mobile device is equipped with an array of antennas capable of transmitting signals in different directions [6].

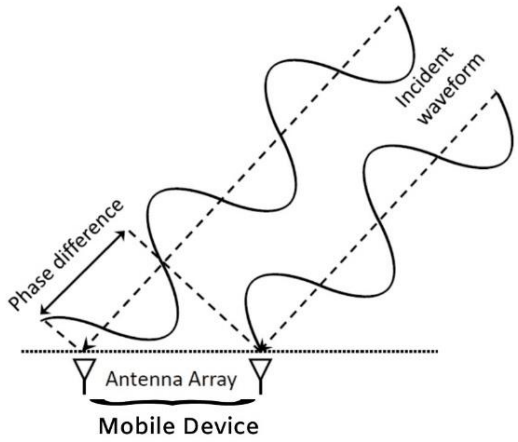

Figure 5. Localization based on PoA [16].

The fundamental principle behind AoD localization is based on the concept that the angle at which a signal departs from the mobile device transmitting antenna provides valuable information about its spatial location relative to the base stations. This information is then utilized to triangulate the position of the mobile device [6]. This involves calculating the intersection point of lines originating from the base stations at the estimated angles. The intersection point obtained through triangulation represents the estimated position of the mobile device. This position is typically provided in terms of coordinates within the coverage area [26].

AoD localization offers several advantages, including its suitability for scenarios where Line-of-Sight (LOS) conditions are prevalent, such as open outdoor environments. Additionally, it can provide high accuracy positioning in both outdoor and indoor settings. However, it may also be sensitive to obstacles and multipath propagation that can alter the angle estimation accuracy [27]. In the context of future mobile networks (*e.g.*, 6G), the AoD technique holds significant potential for improving localization accuracy [26].

H. Angle Difference of Departure (ADoD)

The *Angle-Difference-of-Departure (ADoD)* localization technique is a method used to determine the position of a mobile device based on the angles at which the signals from the mobile device are transmitted from multiple antennas. ADoD-based localization utilizes an array of antennas at the receiver side to measure the angles at which the signals arrive. By comparing the angle differences of arrival at these antennas, the system can triangulate the position of the mobile device [6]. ADoD-based localization can offer improved accuracy compared to single angle localization techniques. However, AoD information can be used to assist ADoD-based localization and estimate orientation alongside estimated positions [6].

III. Current Localization Systems and Services

In modern telecommunications, localization systems and services play a crucial role in enabling a wide array of applications that rely on precise positioning information. This section provides an overview of localization systems and services that are currently in use.

A. Ultra Wideband (UWB)

UWB technology has gained significant attention for indoor localization due to its immunity to interference from other signals [16]. Currently, this technology is primarily employed in short-range communication systems, such as PC peripherals, and various indoor applications [16]. It operates by transmitting ultra-short pulses with a duration of less than 1 nanosecond (ns) over a broad frequency range of 3.1 to 10.6 GHz. It utilizes a low duty cycle, resulting in reduced power consumption [16].

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

UWB signals have the ability to penetrate various materials, including walls, although metals and liquids can potentially interfere with them [16]. Additionally, the short duration of UWB pulses minimizes the impact of multipath effects, enabling the identification of the primary signal path and providing accurate Time of Flight estimation. Research has demonstrated that UWB localization can achieve accuracy levels as precise as 10 cm [28]. However, the development of UWB standards has progressed slowly, particularly in terms of its adoption as a standard in consumer products and portable devices [16].

B. Visible light positioning systems (400-790 THz)

Visible Light Positioning (VLP) systems leverage *Visible Light Communication* (VLC) technology to estimate the location of receivers by utilizing Light-Emitting Diode (LED) transmitters with known positions [6]. In indoor environments, VLP systems, coupled with the capabilities of LED technology, enable precise and reliable localization, making them an affordable solution for applications that demand compact and efficient positioning systems [29]. The potential benefits of VLP in terms of precise localization make it an attractive option for specific applications and environments where high accuracy is crucial [29]. Unlike traditional wireless systems that rely on congested, limited, and costly RF spectrum, VLP systems make use of the visible light portion of the electromagnetic spectrum, which is free from licensing and regulations [30]. This visible light spectrum enables high-speed data transmission and reduces operational costs for operators [29].

While VLP systems offer the advantage of high accuracy localization, one of their main limitations is to design a system, which like GPS, can combine the functions of data receiver and position estimation in order to provide accurate position information without accessing external databases [31]. This means that implementing VLP systems may involve significant upfront investments, including the installation of LED-transmitters and receivers, as well as the development of compatible communication protocols and algorithms.

C. mmWave Systems

With the increasing demand for higher data rates and the growing number of connected devices requiring high-precision localization, traditional frequency bands have become crowded and limited in capacity [3]. This has led to the exploration of Millimeter waves (mmWaves) as a solution to meet the ever-increasing demands. Millimeter waves are electromagnetic waves with frequencies ranging from 30 to 100 GHz [6]. These waves have wavelengths ranging from 1 to 10 millimeters [32], hence the name "millimeter waves". They use a relatively high-frequency band compared to traditional radio waves (below 30 GHz) used in wireless communications.

mmWave frequency bands were first introduced in 5G mobile networks [33]. This integration played a crucial role in achieving higher data rates, reduced latency, and improved localization accuracy [6]. It enabled a localization precision of 10 cm [1], and with the inclusion of antenna arrays at the mobile device, it becomes possible to estimate the orientation of the mobile device [34]. Moreover, by using the NLOS paths and reconfigurable intelligent surfaces (RIS), localization tasks can be accomplished using a single base station [35]. These advantages make mmWave systems highly attractive in communication networks [6].

While mmWave systems offer sufficient localization precision for many applications supported by current communication systems, they fall short in meeting the requirements of applications, such as telesurgery, extended reality, holography, connected vehicles [2]. These applications demand a higher level of location accuracy, typically within the range of 1 cm [3]. To fully realize the potential of these applications and pave the way for new ones to emerge, the utilization of higher frequency bands becomes necessary.

D. THz Localization

With the advancement of highly precise positioning capabilities, there is a belief that THz frequency bands (0.1-10 THz) are gaining attention for high-speed transmissions and high accuracy positioning [3]. The integration of these bands has the potential to fulfill the demands of applications that require both high data rates and high localization accuracy, surpassing the capabilities of current communication systems. The IEEE 802.15.3d standard has already proposed the use of sub-THz frequencies, pushing the signal frequency from 73 GHz to 300 GHz, and the bandwidth from 2 GHz to 69 GHz [6].

While THz systems offer advantages, such as higher frequencies, larger bandwidths, and improved localization performance, they also present challenges in hardware design, coverage, overheads, and computational complexity. As compared to 5G mmWave, 6G THz systems are expected to deliver enhanced localization performance.

IV. 6G-Compatible Simulation Tools

This section provides an overview of the following 6G-compatible simulators: *Aff3ct* [36], *CloudRT* [37], *Matlab* [38], Terasim [39], *NYUSIM* [40][41]. We will highlight their strengths and limitations in the context of 6G localization. Based on this evaluation, we will select the most suitable simulator for the upcoming results.

A. Aff3ct

AFF3CT is a free open-source toolbox for Forward Error Correction (FEC) that includes a simulator and a library written in C++ [36]. The toolbox can effectively emulate physical layer behavior for simulation purposes. However, we

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

have observed that AFF3CT does not simulate the entire THz system. Specifically, AFF3CT is more focused on the physical layer, only implementing digital channels without any waveform generation. Therefore, the ability to analyze signal and waveform-related effects on data transmission is limited.

Additionally, AFF3CT fails to consider a range of propagation effects, such as the ground nature, weather, polarization, human blockage, barometric pressure, humidity, and foliage attenuation. Although AFF3CT provides high throughput simulations with multi-node, multi-threaded, and vectorization paradigms, the results obtained are limited since the simulator does not consider the complete communication system scenario.

B. CloudRT

CloudRT is an open-source Ray Tracer simulator that can be accessed via the platform (www.raytracer.cloud). This platform has three main libraries: the Environment library, Material library, and Antenna library [37]. The Environment library contains 3D models of environments for Ray Tracing simulation, and users can upload and manage their models through the platform's web user interface. The Material library stores different material parameters required for different propagation models, including dielectric parameters, transmission loss, scattering coefficients, and equivalent roughness. The Antenna library provides information on various types of antennas, and users can create and upload their antenna radiation patterns to the library.

However, CloudRT has some limitations. It does not consider essential parameters, such as human blockages, polarization, rain rate, and barometric pressure, which could affect the simulation results. Therefore, the provided scenario may be insufficient for localization simulation.

C. Matlab

Matlab is a robust and useful tool that has been widely utilized in communication systems for various simulations due to its communication toolbox. The toolbox offers an extensive range of signal processing functions, including standard-compliant waveform filters, multi-carrier systems, statistic channel models, and antenna systems [38]. However, Matlab has limitations, as it does not fully support THz communication systems. This means that Matlab's 5G toolbox needs to be adapted to 6G mobile networks. Moreover, since Matlab is a closed platform with an extra license requirement, users have limited insight into specific realizations of algorithms and applications.

D. Terasim

Terasim is a system-level simulator that fully supports THz frequencies [39]. It is an extension of NS-3, with data transmission modeled at a packet level, and the successful reception of packets is determined by the received power. While Terasim considers some useful parameters, such as molecular absorption, it neglects essential parameters, such as human blockages, temperature, barometric pressure, humidity, polarization, foliage, among others. Additionally, the realization of signals in Terasim is limited to the consideration of power density spectra, which means that the effects of inter-symbol interference or multipath propagation cannot be examined. As a result, Terasim is inadequate for localization purposes.

E. NYUSIM 4.0

NYUSIM is an open source, system level simulator based on Matlab that supports millimeter and Sub-THz bands ranging from 0.5-150 GHz [40][41]. Unlike other simulators mentioned earlier, NYUSIM is advantageous because it considers a wide range of parameters, including 21 channel parameters, 12 antenna properties, 10 spatial consistency parameters and 6 human blockage parameters. These parameters are critical in demonstrating and clarifying the proposed approach for 6G localization. Also, NYUSIM has been developed based on extensive real-world measurements [40] at multiple mmWave and Sub-THz frequencies, over 2 terabytes of measurement data from 28 to 142 GHz in various environments obtained during 2011 and 2022 [42]. As of 2022, NYUSIM is widely used by industry and academia as an alternative to 3GPP Spatial Channel Model. In this context, we have chosen NYUSIM for our simulations.

V. A 6G-BASED APPROACH

The proposed approach is based on AoD localization technique, and is tailored for real-world scenarios, where signal variations are introduced due to environmental factors, such as reflection, diffraction, and scattering. In this section, we frequently use the term Transmitter (TX), which refers to a Base Station (BS), and the term Receiver (RX), which refers to a mobile device. To determine the localization coordinates of RX, we rely on the concept explained in [26]. To initiate this process, the knowledge of three key parameters is required: the distance d_{ML} between TX and RX at a particular point, the AoD β at the TX side, as well as the x and y location coordinates of TX, represented as x_{TX} and y_{TX} respectively, as illustrated in Figure 6.

Since TX remains stationary and acts as the reference point positioned at coordinates (0, 0), we need the values of d_{ML} and β to obtain coordinates x and y of RX. In this scenario, the following formulas can be used:

$$x_{RX} = d_{ML} cos(\beta) \quad (4)$$

$$y_{RX} = d_{ML} sin(\beta) \quad (5)$$

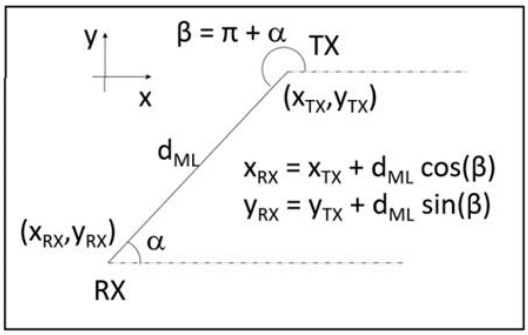

Figure 6. Calculation of mobile device localization coordinates [26].

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

where x_{RX} is the x localization coordinate of RX, y_{RX} is the y localization coordinate of RX, d_{ML} is the distance between TX and RX at a particular point, β is the AoD at the TX side.

There is an alternative method to calculate x_{RX} and y_{RX} from α, when α is provided. This can be achieved by using the following formulas:

$$x_{RX} = d_{ML} cos(180° + \alpha) \quad (6)$$

$$y_{RX} = d_{ML} sin(180° + \alpha) \quad (7)$$

where α is the AoA at RX and d_{ML} is the distance between TX and RX at a particular point.

Since 6G communication systems will operate at THz frequencies [1], the proposed approach integrates AoD localization technique into *THz localization* in order to achieve high-precision user terminal localization.

VI. Implementation and Results

A. Simulation setup

By using NYUSIM 4.0 [40][41] for setting up the simulation environment, parameters, such as building layouts, street configurations, and environmental conditions, are considered. Also, based on real weather data, parameters, such as rain rate, humidity, and temperature, will be set to accurately model the effects of weather on the localization performance. The placement of BS, mobile device, and other relevant elements is carefully determined to mimic the actual deployment scenarios in Downtown Quebec City. Additionally, the mobility patterns of users, such as walking speed and trajectories, are adjusted to match typical pedestrian movements in the chosen area. This ensures that the simulations accurately reflect real-world scenarios and enable to draw meaningful insights about the performance of 6G localization systems in a location-specific context.

Using NYUSIM 4.0 [40][41], we perform simulations for two different scenarios: Simulation 1 (which implements scenario 1) emulates a 5G mobile network using a frequency of 28 GHz and a bandwidth of 800 MHz, whereas simulation 2 (which implements scenario 2) emulates a 6G mobile network using the frequency of 142 GHz and the bandwidth of 1000 MHz to replicate a 6G mobile network. To ensure that the simulations will produce meaningful and applicable results, we carefully tailor each scenario to match the conditions and characteristics of a specific area in Downtown Quebec city.

As shown in Figure 7, the simulations take place for each scenario along two streets, Cook Street (located in front of the ministry of municipal affairs and housing) and Dauphine Street, each covering a distance of 40 meters. More specifically, the simulations replicate a pedestrian's walk from Cook Street (Point A to B) and then, upon reaching point B, make a left turn onto Dauphine Street (Point B to C).

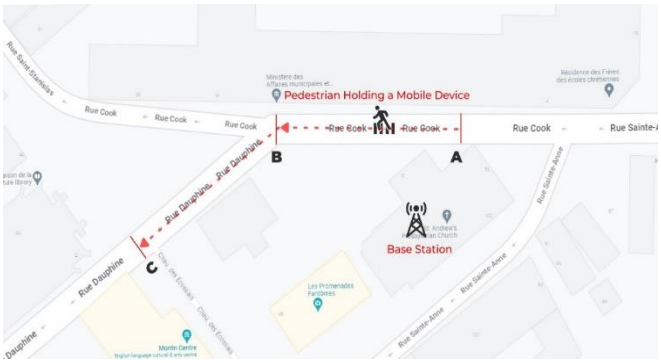

Figure 7. Trajectory of a pedestrian moving from A to C with a mobile device.

The moving distance is set to 80 meters, representing the total distance the user can travel along the path. Throughout the walk, we continuously track and locate the pedestrian, gathering valuable data for analysis. In particular, the corresponding x and y coordinates along this path are recorded.

For each scenario, the stationary BS is located at point (0, 0), serving as the reference point for the measurements. More specifically, it is placed on the rooftop of St. Andrew's Church, with a total height of 35 meters, considering the presence of trees and buildings that create NLOS scenarios. Since the mobile device will be held by a moving pedestrian, we set the mobile device height to 1.5 m from the ground, assuming the pedestrian will use his mobile phone while traveling at a velocity of 1 m/s along the path. Since only one pedestrian is involved in the simulation at a time, the number of RX locations was set to 1. Other environmental factors, such as barometric pressure (1013.25 mbar), humidity (50%), and temperature (20 degrees Celsius), are also set to realistic values. As there is no rainfall in this scenario, the rain rate is set to 0 mm/hr. These carefully selected simulation parameters ensure a realistic representation of the chosen environment and facilitate the collection of relevant localization data for our analysis.

B. AoD Measurement

Since the proposed approach is based on AoD localization technique, it is important to measure this angle in order to get precise localization values. Figures 8 and 9 demonstrate the procedure for extracting the AoD for simulation 1 and simulation 2, respectively. Such a procedure involves using wepik, an online graphic editor, to trace a line between the BS and a specific point to determine the rotation angle of the line. To obtain the AoD, we subtract the rotation angle from 180 degrees. Moreover, when the resulting rotation angle exceeded 180 degrees, we obtained the AoD by subtracting 180 degrees from the obtained rotation angle. This approach improved the accuracy of angle extraction process.

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

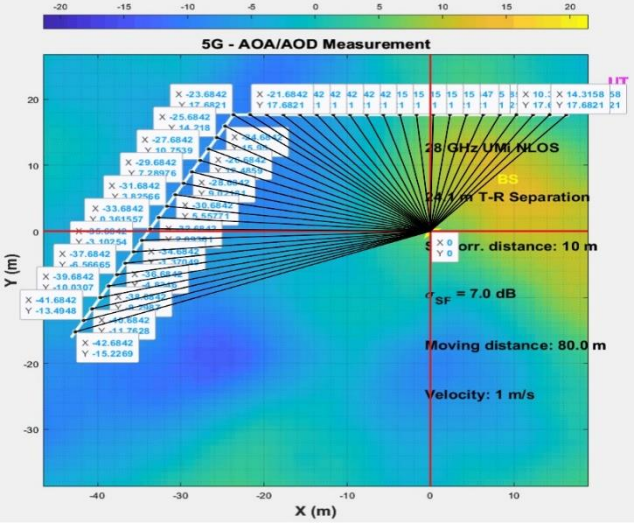

Figure 8. Measurement of AoD for 5G mobile network.

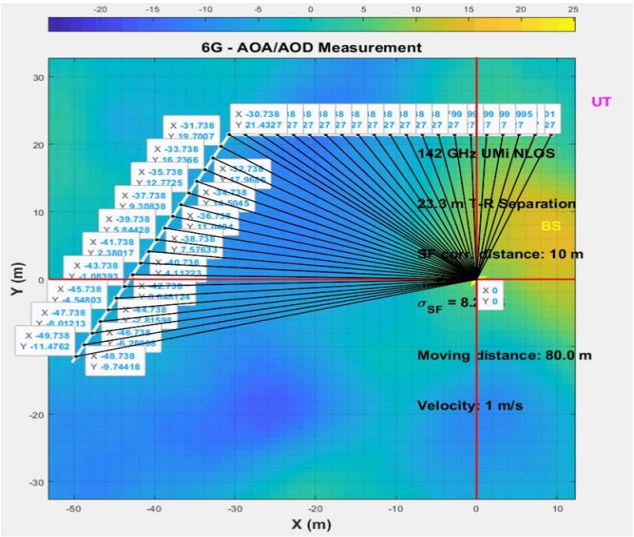

Figure 9. Measurement of AoD for 6G mobile network.

Figures 8 and 9 display various x and y localization coordinates of the mobile device, generated by NYUSIM during the simulations.

C. Evaluation of localization error

The localization error represents how accurately the NYUSIM localization algorithm can track and estimate the pedestrian position while they walk along the trajectory. It measures the discrepancy between the position obtained from the proposed approach and the position obtained from the simulation tool. As a result, it provides valuable insights into the accuracy and performance of the localization system, and it reflects the level of deviation between the actual path taken by the pedestrian and the path estimated by the simulations. A lower localization error indicates a higher accuracy in positioning, while a higher error suggests a less accurate estimation of the pedestrian location.

To calculate the localization error at a specific point, we take the (x, y) coordinates obtained from the simulation tool, and compare them with localization coordinates (x_{RX}, y_{RX}) obtained from the proposed approach. More specifically, we calculate the localization error, first on the X-axis, then on the Y-axis. This decision enables to separately examine and analyze the localization error on both X and Y axes, which helps better understand the factors contributing to the deviation in the x and y coordinates and gain a more detailed understanding of the localization accuracy.

The localization error x_E on the X-axis can be calculated as follows:

$$x_E = x - x_{RX} \tag{8}$$

where x is the localization coordinate on the X-axis obtained from the NYUSIM simulation tool, and x_{RX} is the localization coordinate on the X-axis obtained from the proposed approach. Similarly, the localization error on the Y-axis y_E can be calculated as follows:

$$y_E = y - y_{RX} \tag{9}$$

where y is the localization coordinate on the Y-axis obtained from the simulation tool, and y_{RX} is the localization coordinate on the Y-axis obtained from the proposed approach.

By comparing the localization errors between different simulations (5G vs 6G scenario), we can evaluate the performance and effectiveness of the localization techniques in each scenario, in terms of localization accuracy at different points along the trajectory. We observed that the mean localization error for 6G mobile network is smaller, indicating that the estimated positions from the 6G simulation is closer to actual positions of the pedestrian than the positions estimated in the 5G simulation. More specifically, the mean localization error along the X-axis for the 5G mobile network is 4.71 cm, while 6G mobile network achieves a significantly reduced mean localization error of 0.27 cm. Similarly, along the Y-axis, the 5G mobile network has a mean localization error of 4.57 cm, while the 6G mobile network excels with a small mean localization error of 0.64 cm. Such results are summarized in Table 1 and show that the incorporation of THz frequency bands in 6G mobile networks has the potential to significantly improve localization accuracy, and opens up new possibilities for various applications and services that rely on precise positioning information.

TABLE I. SUMMARY OF THE MEAN LOCALIZATION ERRORS FOR 5G AND 6G MOBILE NETWORKS

Mean localization error	5G mobile network	6G mobile network
X-Axis	4.71 cm	0.27 cm
Y-Axis	4.57 cm	0.64 cm

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

VII. CONCLUSION

In this paper, we propose an approach which combines the AoD technique with THz localization for localizing mobile devices in future wireless networks. The simulation environment is set up by using NYUSIM 4.0. Simulation results show that, for a given trajectory, 6G mobile networks outperform 5G mobile networks in terms of mean localization errors. By providing such accurate positioning and navigation capabilities, the proposed approach enables to realize the full potential of future mobile networks while addressing an aspect that may significantly impact various sectors, including healthcare, industrial automation, agriculture, emergency response, public safety, as well as disaster management.

REFERENCES

[1] B. Ji et al., "Several Key Technologies for 6G: Challenges and Opportunities," in *IEEE Communications Standards Magazine*, vol. 5, no. 2, pp. 44-51, June 2021, doi: 10.1109/MCOMSTD.001.2000038.

[2] Z. Xiao, and Y. Zeng. "An overview on integrated localization and communication towards 6G." *Science China Information Sciences* 65 (2022): 1-46.

[3] C. Chaccour, M. N. Soorki, W. Saad, M. Bennis, P. Popovski and M. Debbah, "Seven Defining Features of Terahertz (THz) Wireless Systems: A Fellowship of Communication and Sensing," in *IEEE Communications Surveys & Tutorials*, vol. 24, no. 2, pp. 967-993, Secondquarter 2022, doi: 10.1109/COMST.2022.3143454.

[4] N. Faiza, et al. "A review of vision and challenges of 6G technology." *International Journal of Advanced Computer Science and Applications* 11.2 (2020).

[5] B. Andre, et al. "6G White Paper on Localization and Sensing." *arXiv preprint arXiv:2006.01779* (2020).

[6] H. Chen, H. Sarieddeen, T. Ballal, H. Wymeersch, M. -S. Alouini and T. Y. Al-Naffouri, "A Tutorial on Terahertz-Band Localization for 6G Communication Systems," in *IEEE Communications Surveys & Tutorials*, vol. 24, no. 3, pp. 1780-1815, thirdquarter 2022, doi: 10.1109/COMST.2022.3178209.

[7] J. A. del Peral-Rosado, R. Raulefs, J. A. López-Salcedo and G. Seco-Granados, "Survey of Cellular Mobile Radio Localization Methods: From 1G to 5G," in *IEEE Communications Surveys & Tutorials*, vol. 20, no. 2, pp. 1124-1148, Secondquarter 2018, doi: 10.1109/COMST.2017.2785181.

[8] A. Alarifi, A. Al-Salman, M. Alsaleh, A. Alnafessah, S. Al-Hadhrami, M. A. Al-Ammar, and H. S. Al-Khalifa "Ultra wideband indoor positioning technologies: Analysis and recent advances." Sensors 16, no. 5.2016: 707.

[9] S. Ju, O. Kanhere, Y. Xing and T. S. Rappaport, "A Millimeter-Wave Channel Simulator NYUSIM with Spatial Consistency and Human Blockage," *2019 IEEE Global Communications Conference (GLOBECOM)*, Waikoloa, HI, USA, 2019, pp. 1-6, doi: 10.1109/GLOBECOM38437.2019.9013273.

[10] C. De Lima et al., "Convergent Communication, Sensing and Localization in 6G Systems: An Overview of Technologies, Opportunities and Challenges," in *IEEE Access*, vol. 9, pp. 26902-26925, 2021, doi: 10.1109/ACCESS.2021.3053486.

[11] J. Lee, A. A. Badrudeen and S. Kim, "6G Integrated Sensing and Communication: Recent Results and Future Directions," *2022 13th International Conference on Information and Communication Technology Convergence (ICTC)*, Jeju Island, Korea, Republic of, 2022, pp. 1219-1221, doi: 10.1109/ICTC55196.2022.9952377.

[12] J. Sanusi, O. Oshiga, S. Thomas, S. Idris, S. Adeshina and A. M. Abba, "A Review on 6G Wireless Communication Systems: Localization and Sensing," *2021 1st International Conference on Multidisciplinary Engineering and Applied Science (ICMEAS)*, Abuja, Nigeria, 2021, pp. 1-5, doi: 10.1109/ICMEAS52683.2021.9692415.

[13] C. Laoudias, A. Moreira, S. Kim, S. Lee, L. Wirola and C. Fischione, "A Survey of Enabling Technologies for Network Localization, Tracking, and Navigation," in *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 3607-3644, Fourthquarter 2018, doi: 10.1109/COMST.2018.2855063.

[14] H. Wymeersch et al., "6G Radio Requirements to Support Integrated Communication, Localization, and Sensing," *2022 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*, *Grenoble*, France, 2022, pp. 463-469, doi: 10.1109/EuCNC/6GSummit54941.2022.9815783.

[15] Y. Zheng, Z. Zhou, and Y. Liu. "From RSSI to CSI: Indoor localization via channel response." *ACM Computing Surveys (CSUR)* 46, no. 2 (2013): 1-32.

[16] F. Zafari, A. Gkelias and K. K. Leung, "A Survey of Indoor Localization Systems and Technologies," in *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2568-2599, thirdquarter 2019, doi: 10.1109/COMST.2019.2911558

[17] T. Xiaohua, W. Li, Y. Yang, Z. Zhang, and X. Wang. "Optimization of fingerprints reporting strategy for WLAN indoor localization." *IEEE Transactions on Mobile Computing* 17, no. 2 (2017): 390-403.

[18] C. Giuseppe, L. De Nardis, F. Lemic, V. Handziski, A. Wolisz, and M. Di Benedetto. "ViFi: Virtual fingerprinting WiFi-based indoor positioning via multi-wall multi-floor propagation model." *IEEE Transactions on Mobile Computing* 19, no. 6 (2019): 1478-1491.

[19] S. Shuang, and L. Wang. "Overview of WiFi fingerprinting-based indoor positioning." *IET Communications* 16, no. 7 (2022): 725-733.

[20] Z. Heng, Z. Zhang, S. Zhang, S. Xu, and S. Cao. "Fingerprint-based localization using commercial LTE signals: A field-trial study." *In 2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall)*, pp. 1-5. IEEE, 2019.

[21] B. M. Majid, A. Rao, and D. Yoon. "RF fingerprinting and deep learning assisted UE positioning in 5G." *In 2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, pp. 1-7. IEEE, 2020.

[22] K. Emil, C. S. Alvarez-Merino, H. Q. Luo-Chen, and R. B. Moreno. "Designing a 6g testbed for location: Use cases, challenges, enablers and requirements." *IEEE Access* 11 (2023): 10053-10091.

[23] W. Yutian, X. Tian, X. Wang, and S. Lu. "Fundamental limits of RSS fingerprinting based indoor localization." *In 2015 IEEE conference on computer communications (INFOCOM)*, pp. 2479-2487. IEEE, 2015.

[24] D. Waltenegus, and C. Poellabauer. Fundamentals of wireless sensor networks: theory and practice. John Wiley & Sons, 2010.

[25] J. Xiong, and K. Jamieson. "ArrayTrack: A Fine-Grained indoor location system." In *10th USENIX Symposium on Networked Systems Design and Implementation (NSDI 13)*, pp. 71-84. 2013.

[26] O. Kanhere and T. S. Rappaport, "Position Locationing for Millimeter Wave Systems," *2018 IEEE Global Communications Conference (GLOBECOM)*, Abu Dhabi, United Arab Emirates, 2018, pp. 206-212, doi: 10.1109/GLOCOM.2018.8647983.

[27] D. Zhang, A. Li, M. Shirvanimoghaddam, Y. Li and B. Vucetic, "Exploring AoA/AoD Dynamics in Beam Alignment of Mobile Millimeter Wave MIMO Systems," in *IEEE Transactions on Vehicular Technology*, vol. 68, no. 6, pp. 6172-6176, June 2019, doi: 10.1109/TVT.2019.2910307.

[28] C. Witsarawat, C. Suwatthikul, S. Manatrinon, K. Athikulwongse, K. Kaemarungsi, R. Ranron, and P. Suksompong. "On performance study of UWB real time locating system." *In 2016 7th International Conference of Information and Communication Technology for Embedded Systems (IC-ICTES)*, pp. 19-24. IEEE, 2016.

[29] K. M. Furkan, A. D. Sezer, and S. Gezici. "Localization via visible light systems." *Proceedings of the IEEE* 106, no. 6 (2018): 1063-1088.

[30] P. Parth, X. Feng, P. Hu, and P. Mohapatra. "Visible light communication, networking, and sensing: A survey, potential and challenges." *IEEE communications surveys & tutorials* 17, no. 4 (2015): 2047-2077.

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

[31] C. Stefanie, C. He, A. Neild, and J. Armstrong. "Indoor visible light positioning: Overcoming the practical limitations of the quadrant angular diversity aperture receiver (QADA) by using the two-stage QADA-plus receiver." *Sensors* 19, no. 4 (2019): 956.

[32] M. D. Sheen, D. L. McMakin, and T. E. Hall. "Detection of explosives by millimeter-wave imaging." In *Counterterrorist Detection Techniques of Explosives*, pp. 237-277. Elsevier Science BV, 2007.

[33] K. Sakaguchi, T. Haustein, S. Barbarossa, E. C. Strinati, A. Clemente, G. Destino, A. Pärssinen et al. "Where, when, and how mmWave is used in 5G and beyond." *IEICE Transactions on Electronics 100*, no. 10 (2017): 790-808.

[34] A. Shahmansoori, G. E. Garcia, G. Destino, G. Seco-Granados and H. Wymeersch, "Position and Orientation Estimation Through Millimeter-Wave MIMO in 5G Systems," *in IEEE Transactions on Wireless Communications*, vol. 17, no. 3, pp. 1822-1835, March 2018, doi: 10.1109/TWC.2017.2785788.

[35] J. He, H. Wymeersch, L. Kong, O. Silvén and M. Juntti, "Large Intelligent Surface for Positioning in Millimeter Wave MIMO Systems," *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, Antwerp, Belgium, 2020, pp. 1-5, doi: 10.1109/VTC2020-Spring48590.2020.9129075.

[36] C. Adrien, O. Hartmann, M. Leonardon, K. He, C. Leroux, R. Tajan, O. Aumage et al. "Aff3ct: A fast forward error correction toolbox!." *SoftwareX* 10 (2019): 100345.

[37] H. Danping, B. Ai, K. Guan, L. Wang, Z. Zhong, and T. Kürner. "The design and applications of high-performance ray-tracing simulation platform for 5G and beyond wireless communications: A tutorial." *IEEE communications surveys & tutorials* 21, no. 1 (2018): 10-27.

[38] MathWorks. Bridging wireless communications design and testing with matlab (2019).

[39] H. Zahed, Q. Xia, and J. Miquel Jornet. "TeraSim: An ns-3 extension to simulate terahertz-band communication networks." *Software Impacts* 1 (2019): 100004.

[40] S. Sun, G. R. MacCartney and T. S. Rappaport, "A novel millimeter-wave channel simulator and applications for 5G wireless communications," *2017 IEEE International Conference on Communications (ICC)*, Paris, France, 2017, pp. 1-7, doi: 10.1109/ICC.2017.7996792.

[41] H. Poddar, "NYUSIM Wireless Channel Simulator Extension Above 100 GHz and Implementation in ns-3." *PhD diss*., New York University Tandon School of Engineering, 2023.

[42] X. Yunchou, and T. S. Rappaport. "Millimeter wave and terahertz urban microcell propagation measurements and models." *IEEE Communications Letters* 25, no. 12 (2021): 3755-3759.

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

Applications of Computer Vision to Posture Corrections and Eye Disease Prevention

Fu-Yu Chen, and Jin-Shyan Lee*
Department of Electrical Engineering
National Taipei University of Technology, Taipei, Taiwan
*Corresponding author: jslee@mail.ntut.edu.tw

***Abstract*—The aim of this paper is to present a posture correction system that can be used to address childhood myopia and shoulder-neck and eye problems caused by prolonged poor posture in adults. It uses the camera on a laptop to analyze images through recognition technology. When poor posture is detected, it alerts the user via the screen and records user data. The system can improve occupational health problems caused by prolonged poor posture among many workers and assist children and adolescents during periods of rapid visual changes to reduce the trend of myopia that occurs at a younger age. After further research, this concept can also be applied to other consumer electronics products such as tablets and televisions to promote public health and well-being.**

***Keywords*—Computer Vision; Posture Corrections,; OpenCV.**

I. INTRODUCTION

In today's rapidly advancing technology landscape, most people use electronic devices such as computers daily. However, many overlook their usage patterns and the time spent on these devices. Various abnormal sitting postures and computer use [1] have led to various health problems, particularly eye problems such as myopia. Although some studies suggest that there is no significant correlation between near work and myopia, it is still beneficial to control near work time and spend more time outdoors to care for the eyes.

To address this, we recognize the importance of image recognition technology in reminding users to maintain proper postures when using electronic devices. While existing image recognition excels at basic facial and body part recognition, detecting finer details, especially in posture, remains a challenge due to the need for speed and real-time detection. Therefore, we intend to employ technologies such as deep learning and the MediaPipe [2]-[4] image recognition framework to prompt users to adopt the correct posture. Using joint points within the MediaPipe framework [5][6] and tools like OpenCV [7], our goal is to instantly recognize sitting postures and provide real-time reminders, helping improve posture and reducing the risk of eye and other health problems.

Currently, most research mainly uses fewer feature points such as the eyes, nose, and mouth for facial recognition, which can result in less precise identification. We utilize more facial landmarks recognized by MediaPipe, such as cheeks, eyebrows, etc., to calculate the angles x, y, and z, representing the orientation of the user's face. Through these parameters, we determine whether the user has poor posture.

Considering that environments may differ for users, we have designed a feature that allows users to adjust the distance from the screen and the position of their shoulders themselves. This ensures that experiences are better optimized for different settings. Additionally, when the user maintains poor posture for a certain period, a prompt message will be displayed on the GUI interface. Although this paper uses 10 seconds as an example, users can adjust the time as needed.

In Section II we discuss the system's construction, provide an explanation of the system flowchart, and offer detailed insights into the implementation of the four recognition functions. Section III presents the results of these recognition functions, which allow the GUI interface to indicate whether there are any abnormalities in posture based on these findings. Finally, Section IV provides more technical details regarding future work, encountered failures, and challenges during experimentation.

II. SYSTEM IMPLEMENTATION

This section will discuss the overall architecture of the system. First, we will introduce the entire architecture of the system. Next, we present the facial recognition and shoulder recognition functions. Finally, we will explain the relevant settings to run the GUI interface.

A. System Architecture

Fig. 1 depicts the operational flow diagram of our recognition system. At the beginning of the system, OpenCV will activate the camera to capture user images. Users are first required to confirm whether their shoulders are in a proper sitting position via the GUI interface and then adjust the screen-to-face distance accordingly. The collected data and images are then transmitted to MediaPipe for analysis and processing. Once the analysis is complete, the data is sent back to OpenCV for graphical rendering. The processed image, along with user data and all prompt text, is displayed together on the GUI interface, and this process is repeated iteratively.

B. Facial Recognition

The system processes images through MediaPipe, which utilizes its database to recognize faces and transmits the values of facial landmarks to OpenCV for drawing and display. MediaPipe provides 468 facial landmarks for user calculations. We use these landmarks to determine whether the face is misaligned, focusing on four main aspects: Yaw (horizontal tilt), pitch (vertical tilt), roll (rotation), and screen-to-face distance. Based on these data, the system alerts users to any inappropriate posture.

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

C. Shoulder Recognition

The system analyzes the images using MediaPipe, using its database to recognize the head, hands, limbs, and torso, and marking the corresponding nodes and skeletons. These feature points are then transmitted to OpenCV for drawing and display. We use the feature points at the shoulder joints to identify whether the shoulders are misaligned or if there is poor posture.

D. GUI Interface Operation

Users need to confirm two initial settings. First, they must establish the correct shoulder position, allowing the system to record the data for the correct shoulder position. Verifying the correct shoulder position helps prevent situations where both shoulders are raised simultaneously, but are not recognized as improper posture. Second, users must set the screen-to-face distance. Since the size of the screens varies for each individual, this setting allows the system to record the distance at which the user feels that the screen is too close. When the user's distance from the screen is less than the distance they have set, the system will provide a reminder. This design adds flexibility to the recognition function.

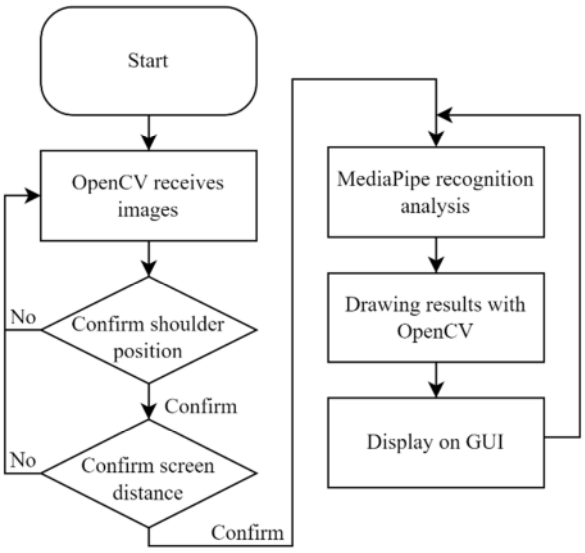

Fig. 1. Recognition system operation flow chart

III. Experiments and Results

The posture correction function in this paper utilizes feature point data computed from MediaPipe's recognition, resulting in four main aspects: recognition of facial yaw and pitch angles, facial roll angle recognition, facial distance from the screen recognition, and shoulder horizontal recognition.

A. Recognition of facial angle of pitch and wrinkle

This recognition function uses the feature points of the eyes, nose, and mouth to calculate the orientation of the face. When the angle of yaw is higher, the absolute value of y increases, as shown in Fig. 2 Similarly, when the pitch angle is higher, the absolute value of x increases, as illustrated in Fig. 3.

B. Recognition of facial roll angle

This recognition function uses the feature points on the left and right cheeks of the face to calculate whether the face has rolled. Calculate the angle by taking the xy values of the feature points on the left and right cheeks. When the rolling angle is greater, the angle itself increases, as shown in Fig. 4.

C. Facial distance from screen recognition

This recognition function utilizes the z-value of the feature point on the nose to calculate the distance of the face from the screen. When the face is closer to the screen, the value becomes smaller, as illustrated in Fig. 5.

D. Horizontal shoulder recognition

This recognition function utilizes the y-value of the feature point on the shoulders to calculate whether the shoulders are tilted horizontally. Calculate the difference in y values between the left and right shoulders as the distance gap. The larger the height difference between shoulders, the greater the value, as shown in Fig. 6.

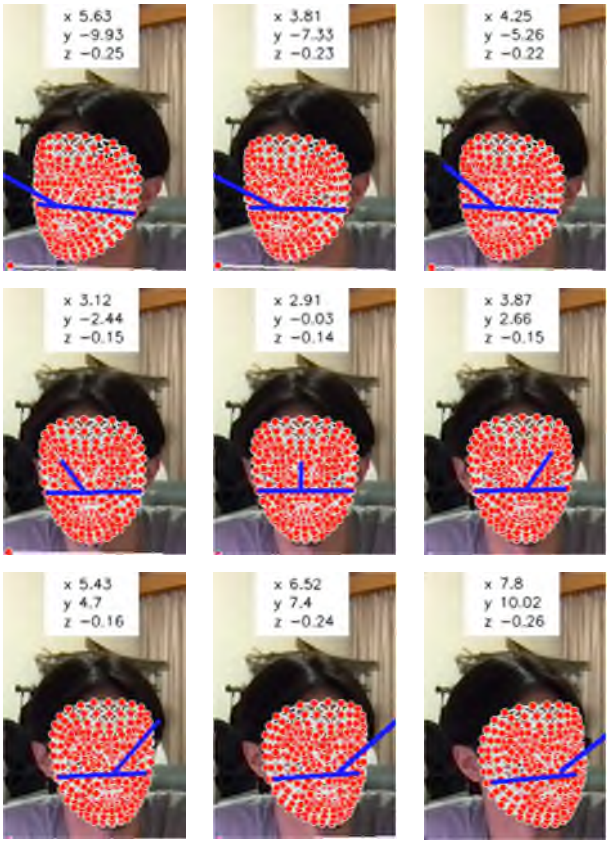

Fig. 2. Recognition of facial yaw angle

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

Copyright (c) IARIA, 2024. ISBN: 978-1-68558-174-9

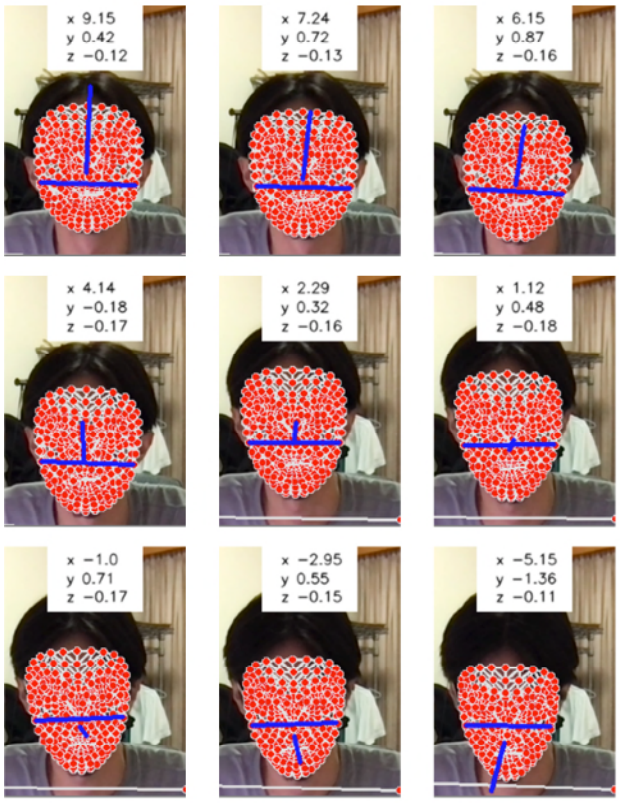

Fig. 3. Recognition of facial pitch angle

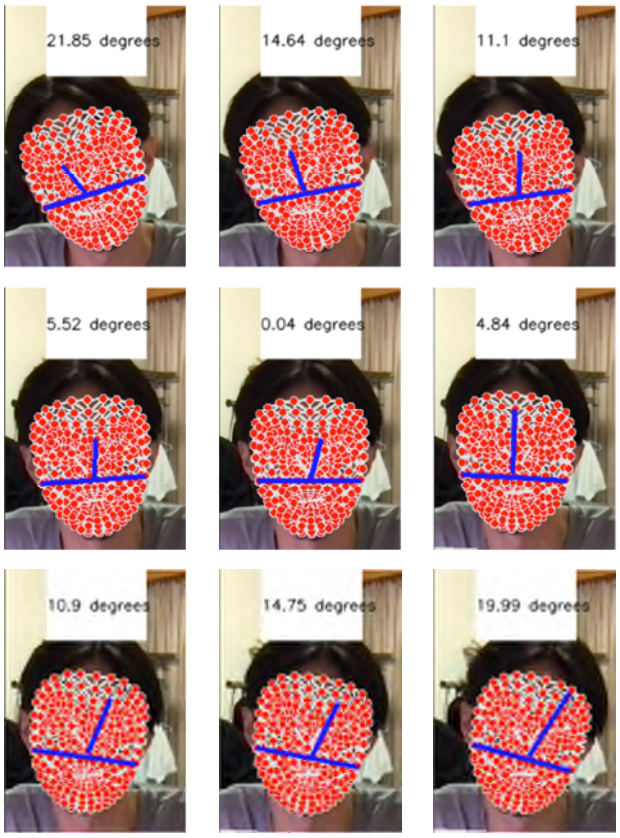

Fig. 4. Recognition of facial row angle

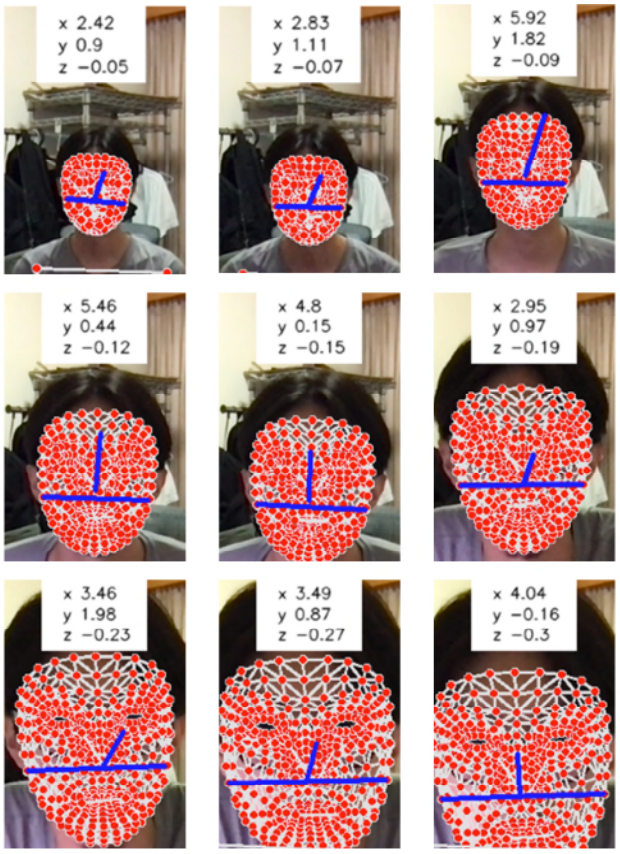

Fig. 5. Screen-to-Face distance recognition

Fig. 6. Horizontal shoulder recognition

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

Copyright (c) IARIA, 2024. ISBN: 978-1-68558-174-9

E. Display of Bad Posture Warning

When a user continuously maintains bad posture for more than 10 seconds, a warning message will appear on the GUI interface. In addition to displaying the word "WARNING", each of the four recognition functions will also individually display the recognition result on the interface, allowing the user to understand the reasons for poor posture and areas that need adjustment. The result messages are as shown in Table I and the result figures are as shown in Fig. 7, Fig. 8, Fig. 9 and Fig. 10.

TABLE I
RECOGNITION FUNCTION RESULT MESSAGES

	Correct Posture	Incorrect Posture
Facial angle of pitch and wrinkle	good position	bad position
Facial roll angle	normal	askew
Facial distance from screen recognition	normal	too close
Horizontal shoulder	normal	askew

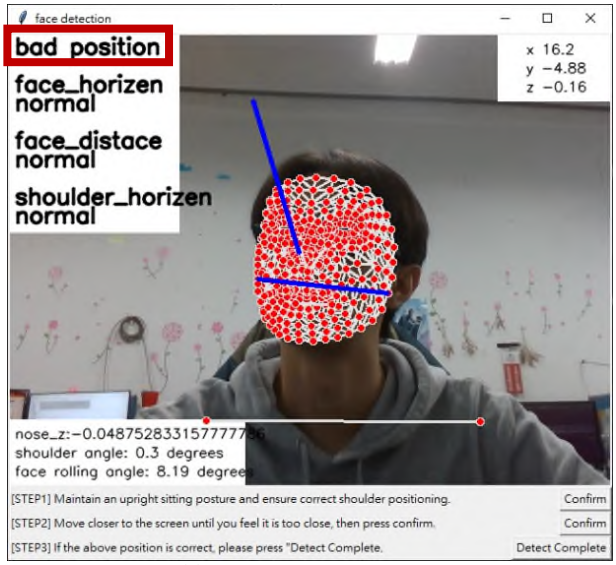

Fig. 7. Results of facial yaw and pitch angle recognition results

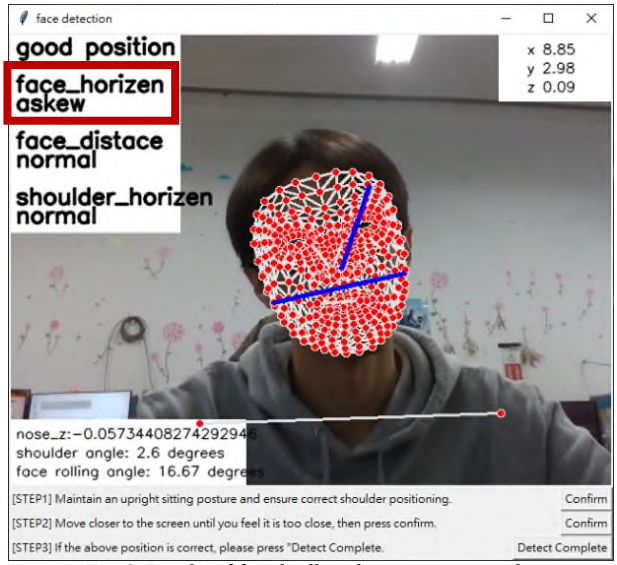

Fig. 8. Results of facial roll angle recognition results

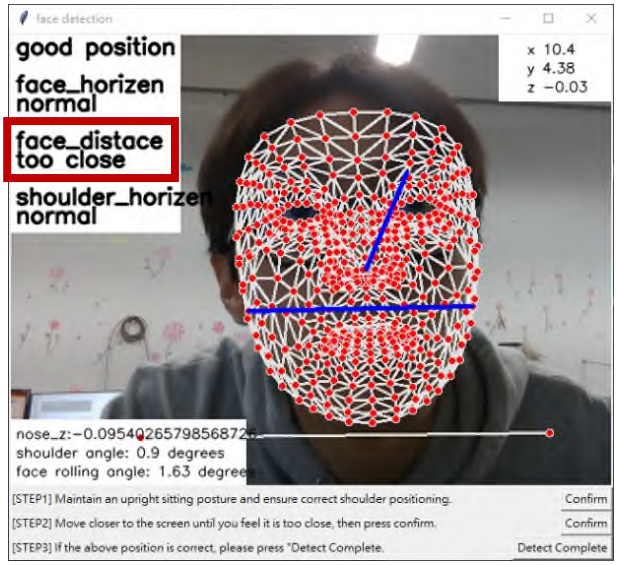

Fig. 9. Facial distance from screen recognition results

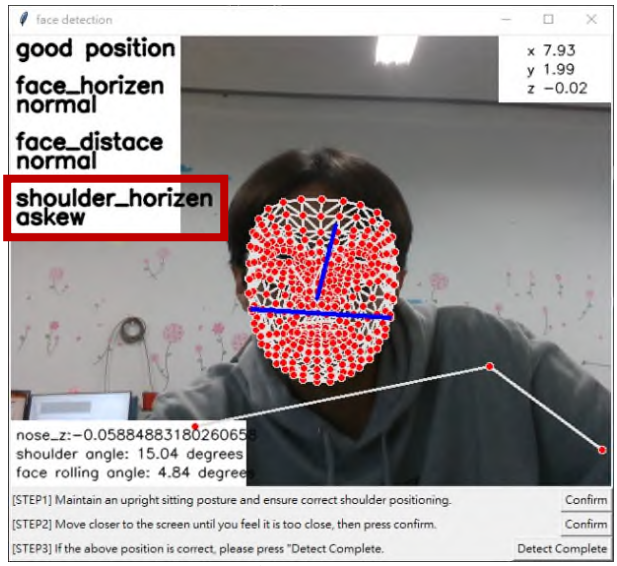

Fig. 10. Results of horizontal recognition results

IV. CONCLUSION AND FUTURE WORK

This paper proposes the use of deep learning, MediaPipe's image recognition framework, and other technologies to remind users to adopt the correct postures when using electronic devices. By leveraging joint points from the MediaPipe framework to design recognition functions, the system can instantaneously recognize sitting postures of users and provide real-time reminders to help them improve poor posture and reduce the risk of eye-related health issues.

The recognition functions and the GUI result display are accurate and smooth, achieving real-time effects. Due to the high accuracy of facial landmarks, the facial recognition function is successful. However, the recognition of shoulder horizontal position sometimes encounters errors and inaccuracies in landmark detection, likely due to the only focus on the upper body. As a result, subtle adjustments in shoulder position may not be properly recognized, leading to occasional failures in shoulder horizontal detection.

In future work, we plan to quantify facial angles and

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

feature point evaluation metrics for comparison with other methods. We also intend to increase the number of subjects to examine whether shoulder level recognition is an individual issue. However, we believe that training a model specifically focused on upper body landmarks can improve the accuracy of shoulder landmark detection, addressing inaccuracies in shoulder horizontal recognition.

ACKNOWLEDGMENTS

This work was supported by National Science and Technology Council of Taiwan under grant NSTC-111-2221-E-027-144.

The authors would like to thank Mr. Hong-Yu Chen and Mr. Bo-Lian Lin from the National Taipei University of Technology for his assistance in the system implementation.

REFERENCES

[1] C. Lance and S. M. Saw, “The association between digital screen time and myopia: A systematic review,” *Ophthalmic and Physiological Optics*, pp. 1-14, Nov. 2019.

[2] C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, F. Zhang, C. L. Chang, M. Yong, J. Lee, and W. T. Chang, ‘‘MediaPipe: A framework for perceiving and processing reality,’’ in *Proc. 3rd Workshop Comput. Vis. AR/VR IEEE Comput. Vis. Pattern Recognit. (CVPR)*, pp. 1-4, Jun. 2019.

[3] C. Lugaresi et al., “MediaPipe: A framework for building perception pipelines,” Jun. 2019.

[4] J. W. Kim, J. Y. Choi, E. J. Ha, and J. H. Choi, “Human Pose Estimation Using MediaPipe Pose and Optimization Method Based on a Humanoid Model,” *Appl. Sci.*, vol. 13, no. 4, pp. 2700, Feb. 2023.

[5] Google, “Pose - MediaPipe,” [Online]. Available: https://google.github.io/mediapipe/solutions/pose.html

[6] Google, “Holistic - Mediapipe,” [Online]. Available: https://google.github.io/mediapipe/solutions/holistic.html

[7] J. W. Kim, J. Y. Choi, E. J. Ha, and J. H. Choi, “Human Pose Estimation Using MediaPipe Pose and Optimization Method Based on a Humanoid Model,” *Appl. Sci.*, vol. 13, no. 4, pp. 2700, Feb. 2023.

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

Copyright (c) IARIA, 2024. **ISBN: 978-1-68558-174-9**