ICN 2022

The Twenty-First International Conference on Networks

ISBN: 978-1-61208-940-9

April 24 - 28, 2022

Barcelona, Spain

ICN 2022 Editors

Eugen Borcoci, University Politehnica of Bucharest, Romania

ICN 2022

Forward

The Twenty-First International Conference on Networks (ICN 2022) continued a series of events organized by and for academic, research and industrial partners.

We solicited 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.

We take here the opportunity to warmly thank all the members of the ICN 2022 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 2022. 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 2022 organizing committee for their help in handling the logistics of this event.

ICN 2022 Chairs

ICN 2022 Steering Committee
Eugen Borcoci, University Politehnica of Bucharest, Romania
Pascal Lorenz, University of Haute Alsace, France
Nicola Ciulli, Nextworks, Italy
Shintaro Mori, Fukuoka University, Japan
Yantong Wang, Shandong Normal University, China
Muath Obaidat, City University of New York, USA

ICN 2022 Publicity Chairs
Javier Rocher, Universitat Politècnica de València, Spain
Lorena Parra, Universitat Politècnica de València, Spain

ICN 2022
Committee

ICN 2022 Steering Committee
Eugen Borcoci, University Politehnica of Bucharest, Romania
Pascal Lorenz, University of Haute Alsace, France
Nicola Ciulli, Nextworks, Italy
Shintaro Mori, Fukuoka University, Japan
Yantong Wang, Shandong Normal University, China
Muath Obaidat, City University of New York, USA

ICN 2022 Publicity Chairs
Javier Rocher, Universitat Politècnica de València, Spain
Lorena Parra, Universitat Politècnica de València, Spain

ICN 2022 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
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, University Politehnica of Bucharest, Romania
Fernando Boronat Seguí, Universitat Politecnica de Valencia-Campus de Gandia, Spain
Radoslav Bortel, Czech Technical University in Prague, Czech Republic
Christos Bouras, University of Patras, Greece
An Braeken, Vrije Universiteit Brussel, Belgium

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
Markus Hofmann, Nokia Bell Labs, USA
M. Reza Hoseinyfarahabady, University of Sydney, Australia
Md Shafaeat Hossain, Southern Connecticut State University, USA
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
Pasquale Imputato, University of Naples Federico II, Italy
Muhammad Shahid Iqbal, Institute of Space Technology, Pakistan
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
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, Georgia Institute of Technology, USA
Jaime Lloret Mauri, Polytechnic University of Valencia, Spain
Rafael Lopes Gomes, Universidade Estadual do Ceará (UECE), Brazil
Pascal Lorenz, University of Haute Alsace, France
Quang-Trung Luu, Nokia Bell Labs / University of Paris-Sud, 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
Antonio Matencio-Escolar, University of the West of Scotland (UWS), UK

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

Cooperative Communication Scheme using Network Coding and Constructive-Interference Phenomena for Information-Centric Wireless Networks

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 describes a cooperative communication scheme for information-centric wireless networks, focusing on disaster-resilient smart-city applications. The proposed scheme uses a network coding technique and constructive-interference phenomena to enhance data distribution and reduce radio interference among relay nodes. The results of computer simulations demonstrate the recoverability of forwarding data under the cross-interference environment and the improvement of caching-data spread.**

Keywords-Information-centric wireless networks; Network coding (NC); Cooperative communication

I. INTRODUCTION

Internet-of-Things (IoT) applications have become widespread across various domains, such as smart cities, industrial automation, human healthcare, and smart everything, which has spurred an explosive growth in the number of IoT devices. Central city areas are increasingly using information and IoT technologies to resolve problems related to urbanizations, and thus, the overall IoT systems have been widely adopted as a solution for various urban characteristics, social needs, and governmental structures [1]. Smart cities are typically considered a panacea for urban problems, but large-scale natural disasters and global pandemics are significant treats to our daily lives. Therefore, ensuring the bright future of smart cities, i.e., achieving disaster-resilient smart cities, is of greater importance and influence with modern applications in diverse circumstances. Success in this context is dependent on the effective deployment of advanced wireless network technologies.

Smart-city applications are characterized by use of a massive connectivity, known as machine-type communications, which is quite different from traditional human-type communications in terms of efficiency and reliability. The features of these systems include low power, broad coverage, ultra-density, and mobile edge computing [2]. In addition, today's smart-city solutions face unique limitations due to unpredictable and non-uniform traffic, and some areas may be outside the wireless network coverage, such as rural areas or any area after a disaster has occurred [3]. In disaster-resilient smart cities, the deployment of secure and reliable wireless communications is of extreme importance when dealing with users' health records and other sensitive information [4]. For example, to enable public-safety broadcasting, mission-critical control, and emergency calls, the smart-city applications must be resilient and robust, and provide instant communication with various services [5].

One promising element of the solution for the above demands is the use of an Information-Centric Network (ICN), (e.g., a content-centric network or named-data network). This is a promising network architecture that is poised to replace the current IP-based networks [6]. It natively supports features, such as abstraction, naming, and in-network caching, improves delays and reduces network traffic. Moreover, ICN-based systems can decouple data from its original location and adopt individual data-based security at the network level. However, as far as we know, suitable wireless systems have not yet been sufficiently investigated and discussed from the viewpoints of integrating communication, caching, computing, control, sensing, and localization technologies in disaster-affected and communications outage areas [7].

As a physical-layer protocol underpinning ICN-based networks, the ad-hoc wireless networking and multi-hop relay networking techniques function as clues for adopting practical usage. These technologies enhance the domain of autonomous-distributed services at the cost of efficient utilization of system resources [8]. However, they come with several technical concerns, including limited battery power, range between devices, bandwidth, dis-connectivity, network overload, data redundancy, communication overhead, network lifetime, lack of information, and data integration difficulties. Therefore, a new ICN-based network protocol and friendly wireless communications technologies are strongly required.

On the basis of the above background, this paper provides an overall blueprint of our study in progress, including a novel cooperative communication scheme for effective ICN-based wireless networks. Cooperative communications technologies can be used to increase the gains by harnessing the effects of path diversity, i.e., by having a relay node send the same data to a destination node if the data transmission is not successful. In the proposed scheme, in order to improve the performance of such communication, we apply a Network Coding (NC) [9] technique to eliminate the amount of network traffic on relay nodes. In addition, to reduce the radio interference among multiple relay nodes during the data flooding process, the proposed scheme utilizes a constructive-interference phenomena [10].

 ISBN: 978-1-61208-940-9

The combination of NC and ICN has attracted significant interest in recent studies. Montpetit et al. [11] utilized them in the internetworking layer, by applying the NC technique to enhance the performance of forwarding data in ICNs. Mekbungwan et al. [12] proposed an NC-based data dissemination system made up of bulk data, such as photos, maps, and databases for situational awareness in post-disaster areas. It was designed on the basis of delay-torrent networking's store-carry-and-forward method in order to reduce the amount of network traffic on relay nodes. For the next-generation cellular networks, the packet duplication method is being introduced to meet the 99.999% reliability requirement, where the original packet and its duplicate are transmitted via two different physical paths, which is the same concept as the path-diversity technique [13]. However, the radio resource consumption is proportional to the number of data copies, and this duplication of data caching leads to a significant waste of radio resources. To tackle this problem, Zhu et al. [14] proposed a new task-oriented communication technology in which the waveform superposition property of a wireless channel is exploited to achieve over-the-air aggregation of data simultaneously transmitted by devices. The idea of overlaying signals can be seen as a kind of NC in the physical layer.

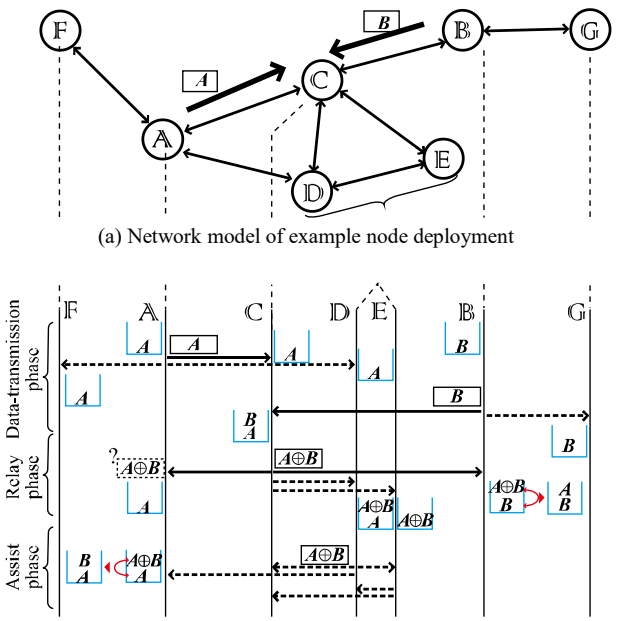
The remainder of this paper is organized as follows. In Section II, we go over the basic principle of cooperative communications. Section III describes the proposed scheme, and Section IV presents the numerical results. We conclude in Section V with a brief summary and mention of future work.

II. COOPERATIVE COMMUNICATIONS IN WIRELESS NETWORKS

Communication between source and destination takes place through different paths by means of cooperating entities called relays. Among the relay techniques in wireless (multi-hop and ad-hoc) networks, the decode-and-forward relaying method is used to decode the data that reaches the relay node and then re-encode and forward them. Another technique, the amplify-and-forward relaying method, can be selected as a simple forwarding mechanism. In the example shown in Figure 1 (a), we focus on nodes $\mathbb{A}$, $\mathbb{B}$, and $\mathbb{C}$ and presume that $\mathbb{A}$ and $\mathbb{B}$ send $\mathbb{A}$'s data of $\boldsymbol{A}$ and $\mathbb{B}$'s data of $\boldsymbol{B}$, and $\mathbb{C}$ exchanges them as relay nodes. In this case, the data transmission is completed in four steps: sending $\boldsymbol{A}$ from $\mathbb{A}$ to $\mathbb{C}$, sending $\boldsymbol{B}$ from $\mathbb{B}$ to $\mathbb{C}$, forwarding $\boldsymbol{A}$ from $\mathbb{C}$ to $\mathbb{B}$, and forwarding $\boldsymbol{B}$ from $\mathbb{C}$ to $\mathbb{A}$. The NC technique is used here with the aim of improving throughput. When $\mathbb{C}$ transfers the bit-by-bit mixed data of $\boldsymbol{A}$ and $\boldsymbol{B}$ by utilizing an Exclusive OR (XOR) operation, the data transmission procedure can be reduced to three steps: sending $\boldsymbol{A}$ from $\mathbb{A}$ to $\mathbb{C}$, sending $\boldsymbol{B}$ from $\mathbb{B}$ to $\mathbb{C}$, and forwarding $\boldsymbol{A} \oplus \boldsymbol{B}$ from $\mathbb{C}$ to $\mathbb{A}$ and $\mathbb{B}$ during broadcasting. After receiving $\boldsymbol{A} \oplus \boldsymbol{B}$, $\mathbb{A}$ can restore $\boldsymbol{B}$ by $(\boldsymbol{A} \oplus \boldsymbol{B}) \oplus \boldsymbol{A}$, and $\mathbb{B}$ can restore $\boldsymbol{A}$ in the same manner. Note that $\oplus$ denotes the XOR operator.

III. PROPOSED SCHEME

ICN decouples the data from its original location using a name-based data-centric network scheme, which enables the network layer to cache and deliver named data regardless of the availability of the original (source) publisher. Moreover, ICN can provide content-based security, i.e., all security-sensitive information can be exchanged via the wireless channel. In this section, we provide an overview of the proposed scheme, including the proposed cooperative scheme, Media Access Control (MAC) protocol, and wireless communications protocol.



(a) Network model of example node deployment

(b) Message and data flow of proposed data-forwarding processing

Figure 1. Proposed cooperative communication scheme

A. Proposed cooperative communications

In-network caching—where each node duplicates the frequently used data by leveraging their embedded cache memory—helps to decrease the end-to-end delay and reduce the network traffic. To accelerate the effect of the caching processing, the nodes should actively accumulate the caching data. One of the key features of a wireless communication system is that it is generally able to overhear what neighbor nodes can receive whether they desire it or not. For example, in Figures 1 (a), when $\mathbb{A}$ sends $\boldsymbol{A}$ to $\mathbb{C}$, $\mathbb{F}$ and $\mathbb{D}$ can also receive $\boldsymbol{A}$; similarly, when $\mathbb{B}$ sends $\boldsymbol{B}$ to $\mathbb{C}$, $\mathbb{G}$ can also receive $\boldsymbol{B}$, which is essentially an off-path caching mechanism. Similarly, $\boldsymbol{A} \oplus \boldsymbol{B}$ from $\mathbb{C}$ can be received not only from $\mathbb{A}$ and $\mathbb{B}$ but also from $\mathbb{D}$ and $\mathbb{E}$.

For the NC-encoded data, in the proposed scheme, $\mathbb{D}$ and $\mathbb{E}$ also send $\boldsymbol{A} \oplus \boldsymbol{B}$ as a helper with $\mathbb{C}$ by performing multiplexing in the assist phase, as shown in Figure 1 (b). As a result, if $\mathbb{A}$ fails to receive $\boldsymbol{A} \oplus \boldsymbol{B}$ from $\mathbb{C}$, it can recover it by utilizing $\boldsymbol{A} \oplus \boldsymbol{B}$ from $\mathbb{D}$ thanks to the benefit of path diversity afforded through the different wireless channels. By using this mechanism, the nodes located around $\mathbb{D}$ and $\mathbb{E}$ but outside the coverage area of $\mathbb{C}$ can be additionally off-path cached, which expands the number of new cashable nodes.

 ISBN: 978-1-61208-940-9

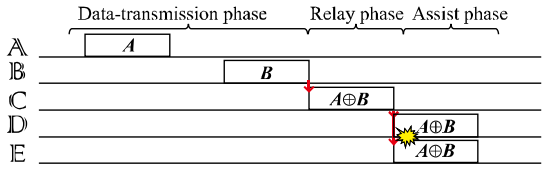


Figure 2. Baseline scenario of proposed scheme in MAC layer

B. Proposed MAC protocol

The pure (unslotted) Aloha method has been adopted as a channel access protocol in commercial low-power wide-area networks, such as SigFox and long-range alliance. In these systems, to eliminate automatic repeat-request messages, the data are iteratively transmitted. The motivation behind using an uncomplicated protocol is to simplify the device implementation (including low-energy consumption), and because the synchronization among nodes is not practically available.

For the above scenario, the proposed scheme is similarly based on current systems. To support the cooperative mechanism, each node has four states: standby, transmission, relay, and assist. Every node regularly maintains the standby state (e.g., $\mathbb{F}$ and $\mathbb{G}$), and the status is changed to the transmission state when it makes a data transmission request (e.g., $\mathbb{A}$ and $\mathbb{B}$). While receiving surrounding (overheard) data, if the node receives two different data and those data should be forwarded, the status moves to the relay state (e.g., $\mathbb{C}$). On the other hand, if the received data is NC-encoded data, the status switches to the assist state (e.g., $\mathbb{D}$ and $\mathbb{E}$). We assume that every node knows whether it needs to relay the data, that the NC-encoded data's number of multi-hops is predefined, and that the nodes can determine their upper limitation of forwarding in order to avoid unlimited hops.

The current wireless communication systems using the pure-Aloha method presuppose that the data transmission has a sufficiently long interval, implying that collision or interference among nodes will not be fatal issues. However, as shown in Figure 2, in the proposed scheme, since the relay nodes and assist nodes forward the NC-encoded data immediately (in the relay and assist phases), collision and interference in a regional area are inevitable.

C. Proposed physical protocol

To tackle the issue of collision and interference caused by forwarding multiple NC-encoded based packets, the proposed scheme adopts the constructive-interference phenomena—if receiver-side nodes can detect a superposition of baseband signals from multiple transmitter-side nodes, the interference can be ignored. In wireless sensor networks, constructive interference has not been extensively exploited because of the difficulty of achieving sufficiently accurate synchronization and the requirement of highly predictable software delays. However, this approach is suitable for the scenarios in which the proposed scheme is applicable (i.e., in the relay and assist phases). Note that, in cases where different data are in conflicts with each other, the proposed scheme cannot be applied, which is beyond the scope of our present study.

TABLE I. SIMULATION PARAMETERS

Terms	Values
Frame length	1,000 bit
Error-control coding	N/A
Modulation method	Binay phase shift keying with Gray mapping rule
Detector's decision type	Hard-decision
Carrier-signal filter	Raised cosine (square root) Rolloff factor: 0.22, Span: 12 symbol
Sampling rate of waveform	4 samples/symbol
Channel model	Additive white Gaussian noise
Signal-to-noise ratio	Relay node: 20 dB, Assist node: 20 dB

IV. NUMERICAL RESULT

In our initial evaluation of the proposed scheme, we investigated the restorability of the baseband signal by using the constructive-interference phenomena and the improvement in data caching among nodes.

Assuming an experimental network composed of a relay node, an assist node, and a receiver node, we implemented a scenario in which the relay node and the assist node send the same data packet to the receiver node. In other words, it is the same as the case where $\mathbb{C}$ and $\mathbb{D}$ forward the NC-encoded data and $\mathbb{A}$ receives them (Figure 1). The computer simulation is conducted using the Matlab simulator and the simulation parameters are listed in Table I. The waveforms of the radio signal arriving from the relay and assist nodes are generated using the same data and system, but they have a time gap of φ. Figure 3 shows the detector's performance for the received signal. Let T denotes the time period required to transmit one symbol of the modulation method. Due to space limitations, we do not illustrate the cases where $\varphi = 0$ and $\varphi = T$ but the former had a clearly separated constellation of received signals and a clear eye pattern, while the latter had the opposite result and thus the detector could not demodulate. According to the results in Figure 3, when $\varphi = 1/4T$, $2/4T$, and $3/4T$, we could achieve the separate construction and obtain a clear eye pattern, and the receiver node could correctly decode.

To illustrate the benefit of the proposed assist nodes, we performed another evaluation using computer simulation implemented in C++ language. In this simulation, 10,000 nodes were deployed in a 1-km^2 area, the communication range of the nodes was set to 100 m, and the unreachable probability of the data (i.e., packet error ratio) was set to 5%. In the conventional scheme, the relay node forwards the NC-encoded data three times, whereas, in the proposed scheme, the assist nodes that receive the NC-encoded data forward at the same time as the relay node. Since the assist nodes that receive the first-forwarded data from the relay station will transmit them twice, and the nodes that receive the second-forwarded data will transmit them once, the end of the assist phase can keep in step with the end of the relay phase, and the proposed scheme can prevent infinite data flooding. As we can see in Figure 4 (a), the number of successfully cached nodes per 10,000 was improved by 43.5% thanks to the assist nodes. As for spreading the NC-encoded data, as shown in

 ISBN: 978-1-61208-940-9

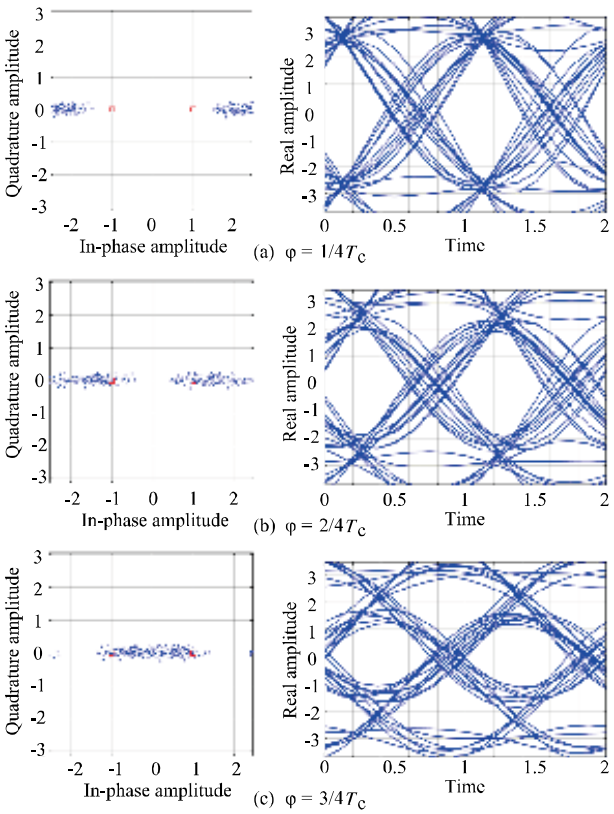


Figure 3. Receiver-side detector's performance, including constellation diagram and eye diagram, for received signals with a time lag.

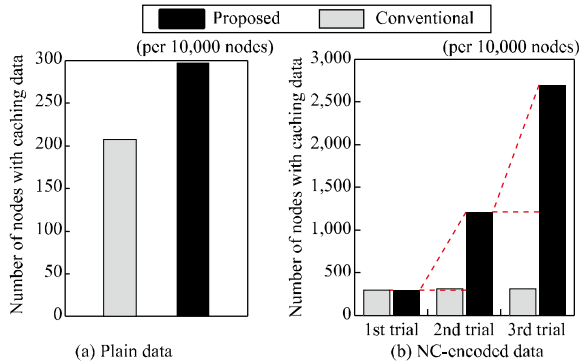


Figure 4. Number of nodes that successfully received and cached

Figure 4 (b), when the forwarding trials were increased, the nodes were enhanced by 306% and 123% for the proposed scheme compared to just 4.98% and 0.234% for the conventional scheme. At the end of the relay and assist phases, the proposed scheme could cache 8.61 times as many nodes as the conventional one. Note that, for decoding the NC-encoded data, plain data is required, e.g., either $\boldsymbol{A}$ or $\boldsymbol{B}$ for $\boldsymbol{A} \oplus \boldsymbol{B}$.

V. CONCLUSION

This paper proposed a novel cooperative communication scheme using the NC technique with constructive-interference phenomena for information-centric wireless networks. Numerical results of our initial evaluation of the scheme were reported. As future work, we will expand the practical scenarios from the quality-of-service perspective to investigate long-lifetime and robustness characteristics.

ACKNOWLEDGMENT

A part of this work was supported by JSPS KAKENHI Grant Number JP21H03436.

REFERENCES

[1] N. Chen, T. Qiu, L. Zhao, X. Zhou, and H. Ning, "Edge intelligent networking optimization for Internet of things in smart city," *IEEE Wireless Commun.*, vol. 28, no. 2, pp. 26–31, Apr. 2021.

[2] F. Guo, et al., "Enabling massive IoT toward 6G: A comprehensive survey," *IEEE Internet of Things J.*, vol. 8, no. 15, pp. 11891–11915, Aug. 2021.

[3] M. El-Tanab and W. Hamouda, "An overview of uplink access techniques in machine-type communications," *IEEE Network*, vol. 35, no. 3, pp. 246–251, May/June 2021.

[4] Y. Li, Y. Yu, W. Susilo, Z. Hong, and M. Guizani, "Security and privacy for edge intelligence in 5G and beyond networks: Challenges and solutions," *IEEE Wireless Commun.*, vol. 28, no. 2, pp. 63–69, Apr. 2021.

[5] Y. Boujelben, "Scalable and QoS-aware resource allocation to heterogeneous traffic flows in 5G," *IEEE Internet of Things J.*, vol. 8, no. 20, pp. 15568–15581, Oct. 2021.

[6] O. Serhane, K. Yahyaoui, B. Nour, and H. Moungla, "A survey of ICN: Content naming and in-network caching in 5G and beyond networks," *IEEE Internet of Things J.*, vol. 8, no. 6, pp. 4081–4104, Mar. 2021.

[7] B. Ji et al., "Several key technologies for 6G: Challenges and opportunities," *IEEE Commun. Std. Mag.*, vol. 5, no. 2, pp. 44–51, June 2021.

[8] O. Hayat, Z. Kaleem, M. Zafarullah, R. Ngah, and S. Z. M. Hashim, "Signaling overhead reduction techniques in device-to-device communications: Paradigm for 5G and beyond," *IEEE Access*, vol. 9, pp. 11037–11050, 2021.

[9] D. Umehara, T. Hirano, S. Denno, M. Morikura, and T. Sugiyama, "Wireless network coding in slotted aloha with two-hop unbalanced traffic," *IEEE J. Sel. Areas in Commun.*, vol. 27, no. 5, pp. 647–661, June 2009.

[10] F. Ferrari, M. Zimmerling, L. Thiele, and O. Saukh, "Efficient network flooding and time synchronization with Glossy," *Proc. ACM/IEEE Int. Conf. Info. Process. Sensor Networks (IPSN)*, Apr. 2011, pp. 73–84.

[11] M. Montpetit, C. Westphal, and D. Trossen, "Network coding meets information-centric networking: An architectural case for information dispersion through native network coding," *Proc. ACM WS Emerging Name-Oriented Mobile Networking Design-Architecture, Algorithms, and App. (NOM)*, June 2012, pp. 31–36, doi: 10.1145/2248361.2248370.

[12] P. Mekbungwan, A. Tunpan, and K. Kanchanasut, "An NC-DTN framework for many-to-many bulk data dissemination in OLSR MANET," *Proc. Int. Wireless Commun. and Mobile Comp. Conf. (IWCMC)*, Aug. 2015, pp. 964–969, doi: 10.1109/IWCMC.2015.7289213.

[13] S. Baek, D. Kim, M. Tesanovic, and A. Agiwal, "3GPP new radio release 16: Evolution of 5G for industrial Internet of things," *IEEE Commun. Mag.*, vol. 59, no. 1, pp. 41–47, Jan. 2021.

[14] G. Zhu, J. Xu, K. Huang, and S. Cui, "Over-the-air computing for wireless data aggregation in massive IoT," *IEEE Wireless Commun.*, vol. 28, no. 4, pp. 57–65, Aug. 2021.

 ISBN: 978-1-61208-940-9

Resource Allocation in 5G and Beyond Networks for Mu-MIMO Systems

Eirini Barri
Department of Computer Engineering and Informatics
University of Patras
Patras, Greece
Email: ebarri@ceid.upatras.gr

Christos Bouras
Department of Computer Engineering and Informatics
University of Patras
Patras, Greece
Email: bouras@cti.gr

Apostolos Gkamas
Department of Ecclesiastical Academy
University of Vella
Ioannina, Greece
Email: gkamas@aeavellas.gr

Vasileios Kokkinos
Department of Computer Engineering and Informatics
University of Patras
Patras, Greece
Email: kokkinos@cti.gr

Aspasia Koukouvela
Department of Computer Engineering and Informatics
University of Patras
Patras, Greece
Email: st1059617@ceid.upatras.gr

***Abstract*—The resource allocation problem focuses on finding an optimal allocation of a specific number of resources to maintain Quality of Service (QoS). That is where Multiple-Input Multiple-Output (MIMO) come in as a radio antenna technology. In order to optimize data speed, minimize errors and improve the radio transmission capacity, MIMO utilizes multiple antennas at both the transmitter and the receiver. This technology uses a variety of antennas and paths that carry the data. Each antenna uses different paths. Multi-user MIMO (mu-MIMO) stands for a technology that allows routers to communicate simultaneously with multiple endpoint devices and it is the next evolutionary step of single user MIMO. Over the past decade, significant advances have been made to improve the performance of mu-MIMO. Although non-negligible progress has been achieved so far, optimal algorithms for Resource Allocation (RA) will help better the performance of mu-MIMO by increasing the system's performance in terms of throughput, fairness, and QoS. The purpose of this paper is to provide a comparison of different beamforming techniques used for resource allocation and list them in ascending order depending on their efficiency.**

***Index Terms*—5G; MIMO; Mu-MIMO; Bandwidth; Spectral Efficiency; Networks.**

I. Introduction

Radio Resource Allocation (RRA) uses a frequency reuse planning of first-generation cellular systems aiming to increase spectrum efficiency. The above utilization plays a significant role since the spectrum is a widely shared and scarce resource. Resource Allocation (RA) as a discipline of its own encompasses a variety of techniques such as dynamic channel allocation, frequency hopping, and power control. More advanced multiantenna concepts such as Multiple-Input Multiple-Output (MIMO) solutions and beamforming also integrate RA.

Significant advances and remarkable research activities have occurred over the past decade in Multi-user MIMO (mu-MIMO) systems. MIMO aims to provide practical solutions and techniques to send and receive more than one data signal simultaneously at the same radio. There is a separation or isolation that helps prevent them from interfering with each other. That allows the wireless devices and Access Points (AP) to send and receive multiple streams of data simultaneously. That eventually increases the transmission speed of the connection. The transmitter and the receiver have more than one antenna/radio chain aiming to support MIMO connectivity. Each spatial stream is transmitted at the same frequency for both the transmitter and the receiver, but using a different radio/antenna chain. The receiver reconstructs the original stream as it knows the phase offsets of its antennas. While multiple streams can be transmitted, only one device can be served and all the other streams are wasted. The total number of antennas that are transmitting the data must be equal to the quantity of the receiving ones at any moment. MIMO can multiply the network capacity by using the horizontal and vertical polarity of a radio wave. All of the above make scaling challenging and this is where mu-MIMO provides many benefits.

In [1], the authors show how MIMO can increase the capacity of the communication system while also improving the reliability of the link that uses a variety of schemes beyond the spatial diversity. Applications involving multiple-cell networks with multiple access channels are presented in [2] where possible coordination between base stations has set the foundation for research.

Mu-MIMO can isolate the traffic of each subscriber al-

 ISBN: 978-1-61208-940-9

lowing them to transmit and receive concurrently between multiple subscribers. MIMO refers to a range of technologies used to multiply the capacity of a wireless connection without requiring additional spectrum by using the horizontal and vertical polarities of the radio wave. A system's Spectral Efficiency (SE) is mainly based on Signal-to-Noise Ratio (SNR), channel estimated accuracy [3], spatial correlation modeled by the propagation environment that is considered in [4] and is also limited by the theoretic capacity [5].

Early surveys report that RA enhances when tracing the momentary fluctuations of the channel in scenarios using a single transmitter, as presented in [6]. During the last years, various techniques are developing for many heterogeneous and diverse MIMO scenarios. This paper's purpose is to classify the state-of-the-art of the already existing algorithms used for single and multiple transmitter scenarios in regards to mu-MIMO.

Furthermore, an overview of different methodologies used for RA is presented in mu-MIMO systems. Moreover, we compare them, explaining the way that they are defined. We aim to provide guidelines for efficient design of RA algorithms, point out practical challenges, and comment on and compare the processing techniques of mu-MIMO's state of the art.

The paper is organized as follows. In Section II, we are presenting RA in mu-MIMO analyzing both transmission and precoding techniques. Furthermore, in Section III, we present RA algorithms, compare them and propose changes to the most promising one in order to make it more efficient. Finally, the conclusions and our future work are provided in Section IV.

II. Resource Allocation in Mu-MIMO

During the last years, mu-MIMO systems have been studied from a practical as well as a theoretical point of view. Mu-MIMO systems come without propagation limitations. Channel rank loss or antenna correlation are some of them and make the mu-MIMO system a perfect candidate for the upcoming wireless systems and standards [7]. Massive MIMO or large-scale MIMO employs a few hundred antennas at the Base Station (BS) that are responsible for sending simultaneously data streams to many users. These are evolutions of the mu-MIMO technology. Massive MIMO is an example of a very promising technology, able to get by with the continuous, growing, capacity needs arising with 5G networks. RA management in wireless communications includes various network functions, such as power control, transmission rate control, call admission control, scheduling, handover, transmitter assignment, and bandwidth reservation. Figure 1 illustrates the components mentioned above, the RA policy, and the connection between them. As displayed in Figure 1, each RA technique can either implement in optimal or suboptimal way. A mu-MIMO system uses various RA techniques that are based on user scheduling and signal to process, as mentioned, and presented in Figure 1, and depend on the SNR value, quantity of users, antennas and coordinated transmitters.

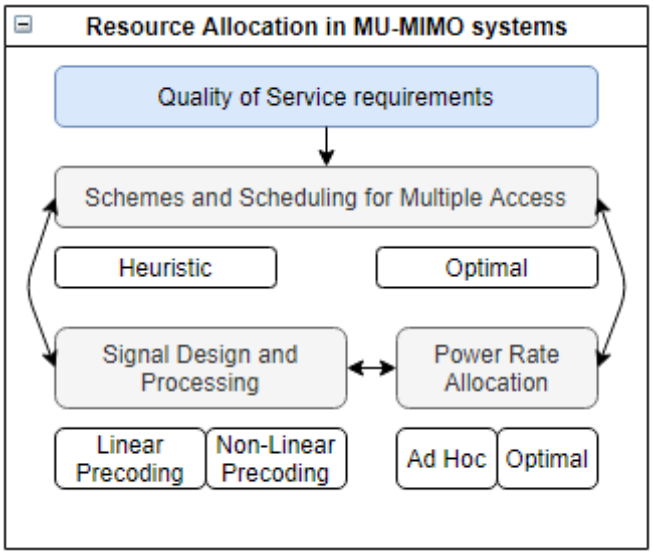


Fig. 1. Coordinated Network in mu-MIMO

In order to maintain the QoS and reduce potential interference that may occur to the users, various components are used. Scheme and Scheduling components as well as a multiple access technique are used to distribute the resources to the users. Mu-MIMO also uses signal design and processing components that are responsible for the transmission of data to the scheduled users simultaneously. Last but not least, a power rate allocation guarantees QoS.

Mu-MIMO allows the same data channels to distribute messages for different users. After the distribution, follows a classification of the individual users that takes place when the data reaches their mobile devices [8]. Serving multiple users with the same transmission ensures increased capacity and better utilization of resources. The latter increases the ability to stream or download with improved user experience even if the area is crowded. Shared data can also provide a faster and more efficient system for all users and can furthermore be switched between one or multiple users.

Each antenna receives both direct and indirect components. The direct ones are intended for this particular antenna and the indirect ones are not. That happens because these two antennas use the same channel. So, it is imperative to divide the transmitted data into multiple independent data streams. The number of streams is always less or equal to the number of antennas and this is further explained in the following section.

The literature on MIMO communications is very rich, and this paper's goal is to provide and compare the different aspects considering RA schemes and mechanisms in Mu-MIMO systems. Users with independent channels can increase the overall performance of a system. However, there are systems that provide orthogonal resource that can be accessed by each user of the cell as stated in [9] and in [10], but the real diversity in MU-MIMO systems comes when many users access the same resource simultaneously.

 ISBN: 978-1-61208-940-9

A. Transmission

Surveys over the years have pointed out that the increase of RA is possible if the instant channel inconstancies are tracked for scenarios including a single transmitter. RA decisions are affected because users have heterogeneous long-term gains because of the wide coverage of areas in wireless communications. Mu-MIMO transmissions are organized into two types, partial and full bandwidth. The latter mode is when Mu-MIMO streams are transmitted and occupy the entire channel of the system in 802.11ac. On the other hand, partial bandwidth mode is the simultaneous use of Orthogonal Frequency-Division Multiplexing Access (OFDMA) and Mu-MIMO. That means that users multiplex in both time and frequency. In order to transmit independently and separately coded data signals from each of the multiple transmit antennas, MIMO uses multiplexing or spatial streaming. The number of mu-MIMO Space-Time Streams (STS) supported by the Access Point (AP) in the Downlink and Uplink depends on the number of transmitter antennas. In general, an AP with N antennas should be able to support up to N STS in both downlink and uplink. When discussing clients, the maximum number of STS for one user can be up to four despite the amount of antennas. Furthermore, the total number of STS is less than or equal to eight. Mu-MIMO aims to accommodate as many users as possible per resource. That is the reason why RA techniques are thoroughly examined at the basic resource unit, e.g., time-slot, code, frequency-time resource block, or single-carrier. Since the RA strategy applies to overall resources, the global system model (single-carrier, OFDM, or Code-division Multiple Access (CDMA)) is not taken into account.

B. Precoding

The term precoding indicates the rotation and scaling of the set of beams having their spatial properties and power adjusted towards one particular purpose. Multi-antenna transmitters provide spatial dimensions and can create autonomous channelization schemes, allowing the user transmitting to serve multiple ones concurrently using the same frequency band and time slot. The above is also known as Space-Division Multiple Access (SDMA). Taking into account the constraints, different techniques can be considered optimal. Precoding can be either linear or non-linear. The latter techniques can provide better performance while the former is computationally less expensive and requires no prior signaling. In mu-MIMO wireless technology, the term beamforming indicates the signal steering needed to achieve SDMA using beams.

III. Comparison

In [11], the authors propose an optimized algorithm for Zero-Forcing (ZF) precoding for a downlink massive MIMO system. To maximize Energy Efficiency (EE), the authors are proposing an iteration algorithm having a sensible power consumption model. Aiming at optimal EE, this particular algorithm uses the perfect Channel State Information (CSI) scenario and uniform rates for each user. In [12], the authors investigate an efficient algorithm achieving the same goal in a massive MIMO system having imperfect CSI. The proposed RA scheme has low complexity and concurrently optimizes the number of antennas, power allocation, and user selection. The EE optimization problem is intractable to solve since it is a mixed integer and non-convex problem and was therefore divided into two subproblems to be solved efficiently. In [13], the authors compared different schemes based on energy efficiency of RA. In Table I, all those schemes are presented and compared based on their results according to their techniques. The first scheme is based on designing prefilters at the Base Transceiver Station (BTS). The above condition meets the ZF criterion and efficiently removes Multi-User Interference (MUI). Having MUI eliminated throughput multiplication can also be achieved since every subcarrier could be assigned concurrently to all the user ends.

TABLE I
Compared Algorithms

Work	Algorithm	Techniques	Results
[11]	ZF	Design prefilters to remove multi-user interference	Good throughput, EE quite low
[12]	EE optimization	Obtain an asymptotic EE expression utilizing smatrix theory	Target QoS
[13]	TSD	Assign each subcarrier to one user.	Best EE Bad throughput
[13]	SM	Spatial Multiplexing	Good throughput EE quite low
[13]	SDMA	Subcarrier assigned to multiple users concurrently	Improved throughput for low power consumption
[14]	Power allocation	Global optimization ranges depending on the user's high or low SINR.	Total EE increases significantly
[14]	MRT	Beamforming	Low Complexity Processing
[15]	Fully-adaptive RA scheme	Joint EE-SE performance	EE performance exceeds semi-adaptive and non-adaptive RA schemes
[16]	MMSE	Combination of ZF and MRT	Good throughput, EE quite low

*EE: Energy Efficiency

The second scheme presented in this paper bases on Transmit Spatial Diversity (TSD). A subcarrier is assigned by the BTS to a user having the maximum vector channel gain.

The third scheme presented in the same paper utilizes the Spatial Multiplexing (SM) technique. This particular technique employs the MIMO system between a specific RT and BTS. Data rate increases since the MIMO system can potentially

 ISBN: 978-1-61208-940-9

send multiple substreams. Every subcarrier is assigned to a particular user each time for both the TSD and the SM scheme. Lastly in the fourth scheme presented, the subcarrier is designated for every user having spatial signatures. This scheme is also known as SDMA and it manages to increase the throughput. The main disadvantage is that spatial channels are seldom orthogonal in practice.

In [14], the authors investigated a MIMO downlink system. This particular system contains a BS with an immense number of antennas serving many single-antenna users. This kind of system creates a problem that considers the circuit power consumption, CSI, QoS requirements including a minimum required data throughput rate. The power allocation scheme which is proposed is optimized for maximization of the EE of data transmission.

Furthermore, in [14] the authors analyze Maximum Radio Transmission (MRT) and present it as practical precoding that can balance the system's performance and complexity. This algorithm provides excellent performance having low processing complexity in massive MIMO systems. Furthermore, MRT manages to maximize the signal gain of the designated user. MRT appears to be almost optimal at systems where noise is limited and inter-user interference is insignificant compared to the noise.

In [15], a demonstration of a fully adaptive RA scheme is presented. The scheme exceeds the EE performance of semi-adaptive and non-adaptive RA schemes. The authors point out that data streams or Radio Frequency (RF) chains are equivalent to two times the number of receiving antennas. These are adequate for achieving EE in a mmWave massive MIMO network. That concludes in reducing the amount of required RF chains and further decreasing the power consumption needed.

In [16], authors analyze the Minimum Mean Square Error algorithm (MMSE) which employs both the ZF and the MRT algorithms analyzed above and is a linear precoding algorithm. Hence, this creates a balance between them and achieves adequate performance in systems with moderate noise and interference. By utilizing the mean square error, this algorithm manages to minimize the error and filter the already sent symbols transmitted from the BS to the received terminal. MMSE's performance exceeds the one of ZF and MRT.

Having analyzed the algorithms, we will compare them further and place them in descending order proposing the one we think excels.

In our opinion, the best algorithm appears to be the TSD since it has the best EE and high utility, which is the result of low power consumption. This algorithm also has low throughput as it is a single-user RA algorithm. The second best algorithm is the power allocation scheme, implementing a low complexity algorithm that also presents a significant EE increase. The following algorithm is considered to be the EE optimization that achieves target QoS and better performance by having low complexity and good EE. The SDMA scheme presents improved throughput for low power consumption since it is a multi-user protocol. SDMA has increased EE when an antenna array at BTS is employed. The disadvantage of SDMA is that channels are not orthogonal most of the time, and the spatial signatures are designated that way. The SM algorithm is a single-user scheme that presents low EE, but has good throughput. The next scheme is MMSE which is a combination of the ZF and MRT algorithms. Although MMSE achieves adequate throughput, it comprises a matrix inversion throughout the processing, making detection methods computationally ineffective for many antennas. The following algorithm is ZF, and even though it has good throughput, it presents bad EE and utility. That is a result of the significant power consumption needed to separate the transmission of multiple users. The MRT algorithm balances performance and efficiency. Despite that fact, it is almost optimal only when the Inter-User Interference (IUI) is less contrasted to the noise. This comparison lowers the efficiency of the algorithm. Adding to that pilot contamination and the imperfect CSI should be taken into account. The Fully Adaptive RA scheme requires a panoramic view of the traffic demands and the loading of the network for efficient scheduling. Despite that disadvantage, EE increases according to the SE. When EE peaks, it starts degrading even if SE is still increasing.

Since the TSD scheme appears to perform better compared to the other schemes, an analysis of the TSD algorithm explained in [17] is presented. The algorithm consists of the following steps:

1) The first subcarrier is assigned to the channel with maximum gain. That corresponds to the maximum number of bits transmitted among all users and antennas in that subchannel. We set Y as the following: $Y = (a, b)$. The pair $(a*, b*)$ represents the data modulated onto the nth subcarrier. The data is transmitted from the BS to the b*th antenna of the a*th Remote Terminal (RT). Furthermore, we set $m^n_{(a,b)}$ as the maximum number of bits that can potentially be transmitted to ath user using nth subcarrier.
2) Calculate $m^n_{(a,b)}$ and T_{a^*} using the appropriate math equations.
3) Ensure that the user T_{a^*} meets the minimum bit rate constraint stating that the data rate designated for ath user should equal Ra bits per OFDM symbol. If $Ta* < Ra*$, the algorithm has to backtrack and check for the next subcarrier. If $Ta* \geqslant Ra*$, then we temporarily dismiss user a^* and backtrack to the first step distributing the subchannels to those users that fail to meet the above constraint.
4) The algorithm keeps backtracking to step 1 until all the users meet the minimum bit rate constraint and, therefore, all the subcarriers have been allocated to subchannels.

Observing Table I, the TSD algorithm seems to be the best option for resource allocation. It is also visible and noticeable that TSD's disadvantage, bad throughput, it's MMSE's advantage. That means, in an ideal scenario, the combination of both of them will lead up to the most efficient resource allocation

 ISBN: 978-1-61208-940-9

algorithm.

IV. Conclusion

Considering that the technology is constantly evolving, it is difficult to keep up with, especially when it comes to wireless networks. Thus, already existing algorithms and technologies must be reconsidered. This paper's ultimate goal, as mentioned above, is to compare algorithms and techniques that a mu-MIMO system uses, and provide the reader with useful information about how the resource allocation works in mu-MIMO systems and how that can significantly improve in order to increase the system's performance. The comparison results show that the TSD algorithm is the most efficient, having the best EE, as well as low power consumption among the others. A suggestion for possible future work might be a comparison provided with the simulation of each algorithm based on fixed network topology in order to present the results figuratively.

References

[1] N. Johannsen, N. Peitzmeier, P. Hoeher, and D. Manteuffel, "On the feasibility of Multi-Mode antennas in UWB and IoT applications below 10 GHz, toward 6G networks: use cases and technologie," IEEE Communications Magazine, vol. 58, page 55-61, 2020.

[2] M. Ben Zid and K. Raoof, "Multi-User MIMO communication : basic aspects, benefits and challenges," Recent Trends in Multi-user MIMO Communications (2013): 3-24.

[3] F. Rusek et al., "Scaling up MIMO: opportunities and challenges with very large arrays," IEEE Signal Processing Magazine, vol. 30, no. 1, pp. 40-60, 2013.

[4] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," IEEE Transactions on Information Theory, vol. 56, no. 6 , pp. 2845-2866, 2010.

[5] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Uplink power efficiency of multiuser MIMO with very large antenna arrays," in 2011 49th Annual Allerton Conference on Communication, Control, and Computing (Allerton), 2011, pp. 1272-1279.

[6] D. Gesbert, M. Kountouris, R. Heath, C.-B. Chae, and T. Sälzer, "From single user to multiuser communications: shifting the MIMO paradigm," IEEE Signal Process. Mag., vol. 24, January 2007.

[7] P. Aggarwal and V. A. Bohara, "A nonlinear downlink multiuser MIMO-OFDM systems," IEEE Wireless Communications Letters, vol. 6, no. 3 , pp. 414-417, 2017.

[8] J. Zhang, E. Björnson, M. Matthaiou, D. W. K. Ng, H. Yang, and D. J. Love, "Multiple Antenna Technologies for Beyond 5G," ArXiv, vol. abs/1910.00092, 2019.

[9] F. Capozzi, G. Piro, L. Grieco, G. Boggia, and P. Camarda, "Downlink packet scheduling in LTE cellular networks: key design issues and a survey," IEEE Communications Surveys Tutorials, vol. 15, no. 2, pp. 678-700, 2013.

[10] G. Y. Li, J. Niu, D. Lee, J. Fan, and Y. Fu, "Multi-Cell coordinated scheduling and MIMO in LTE," IEEE Communications Surveys Tutorials, vol. 16, no. 2, pp. 761-775, 2014.

[11] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient Multi-User MIMO systems: Is Massive MIMO the answer?," IEEE Transactions on Wireless Communications, vol. 14, no. 6, pp. 3059-3075, 2015.

[12] Y. Zhang, H. Gao, F. Tan, and T. Lv, "Resource allocation of energy-efficient Multi-User Massive MIMO systems," 2016 IEEE Globecom Workshops (GC Wkshps), pp. 1-6, 2016.

[13] W.-C. Wu, "Toward the energy efficiency of resource allocation algorithms for OFDMA downlink MIMO systems," Journal of Electronic Science and Technology, vol. 17, no. 4 , pp. 100007, 2019.

[14] Yingchu Guo, Gang Wu, Zhenzhen Hu and Shaoqian Li, "Energy efficient resource allocation for Massive MIMO cellular systems," American Journal of Engineering Research (AJER), vol. 9, no. 4 , pp. 141-149, 2020.

[15] S. A. Busari, K. Huq, G. Felfel, and J. Rodriguez, "Adaptive resource allocation for energy-efficient millimeter-wave Massive MIMO networks," 2018 IEEE Global Communications Conference (GLOBECOM), pp. 1-6, December 2018.

[16] M. A. Albreem, A. H. A. Habbash, A. M. Abu-Hudrouss, and S. S. Ikki, "Overview of Precoding Techniques for Massive MIMO," IEEE Access, vol. 9, pp. 60764-60801, 2021

[17] C.-Y. Wei-Chiang Wu, "Subcarriers and its allocation algorithms for downlink OFDMA-based MIMO systems," ScienceDirect, April 2017.

 ISBN: 978-1-61208-940-9

Multi-beam and High Gain Antenna Array for Radio Frequency Energy Harvesting Applications in 5G Network

Chemseddine Benkalfate
Electrical and Electronics Engineering Department
Quartz laboratory, ENSEA
Cergy, France
e-mail: benkalfate.chemseddine@ensea.fr

Achour Ouslimani
Electrical and Electronics Engineering Department
Quartz laboratory, ENSEA
Cergy, France
e-mail: achour.ouslimani@ensea.fr

Abeb-Elhak kasbari
Electrical and Electronics Engineering Department
Quartz laboratory, ENSEA
Cergy, France
e-mail: abed-elhak.kasbari@ensea.fr

Mohammed Feham
Telecommunications Department
STIC laboratory, university of Tlemcen
Tlemcen, Algeria
e-mail: mohammed.feham@univ-tlemcen.dz

***Abstract*—In this paper, a new multi-beam high gain antenna array for Radio Frequency Energy Harvesting (RF-EH) applications in 5G network is presented which consists of 4 sub-antenna arrays (2 x 2) in order to increase the gain and directivity. The sub-antenna array covers the two 5G network bands (3.5 GHz and 26/28 GHz) with gains of 3 dB and 6 dB respectively. The proposed antenna array is studied and simulated for the 26/28 GHz band. The gain reaches 12 dB at 26 GHz and 15 dB at 27.5 GHz. The simulated efficiency of this antenna is of 78%. The antenna is designed on Teflon glass substrate with a relative permittivity of 2.1 and 0.67 mm of thickness. It total dimension is of 65 x 55 mm². These gain values satisfy the beam forming technology adopted by 5G network.**

Keywords—Antenna array, 5G network, High gain, Beam forming.

I. INTRODUCTION

In 5G network, the communication consists in focusing the transmitted power in the direction of the receiver to ensure a high quality of data transmission. This technique is known as beam forming. The transmission power for the 26/28 GHz band must be low to ensure a safe communication on human health. To increase the received power, it is mandatory to improve the gain of the antennas (in transmission and reception). The received power is given by the following Friis formula.

$$P_r = G_r G_t \left(\frac{\lambda}{4\pi d}\right)^2 \cdot P_t \quad (1)$$

where $G_r G_t$ antenna gains, $P_r P_t$ transmission and reception powers, λ the wave length and d the distance between transmitter and receiver.

The efficiency of RF-EH systems depends on the received power. By adopting the multi-beam forming technique, the RF_EH system stays operational for several orientation.

There are several techniques to increase the antenna gain, including increasing the size of the antenna [1] and using reflectors to concentrate the maximum electromagnetic power in a given direction which increases the gain [2]. The best technique remains the use of an array antenna which allows to considerably improve the gain of the antenna. The main challenge for this technique is the miniaturization. Designing an antenna array with high gain and optimal size facilitates it implementation.

Our work consists in the design and simulation of a multi-beam antenna array for the 26/28 GHz band of 5G network with high gain (12 dB and 15 dB, respectively) and an optimal size of 65 x 55 mm² designed on Teflon glass substrate with 0.67 mm of thickness. Table I presents a comparison between the proposed array antenna and other antennas proposed in references for the same order of gain and frequency band values.

TABLE I: COMPARISON BETWEEN THE PROPOSED ANTENNA AND OTHER PUBLISHED ANTENNAS

Ref.	Antenna size (mm^3)	F (GHz)	Gain (dBi)	Substrate
[3]	80x20x0.25	28-35	14-15	Hybrid
[4]	150x110x0.1	28	10	RO5880
[5]	150x70x0.2	28	11.16	Alron 430
[6]	150x75x0.75	28	14	RO4450B
Proposed Antenna	**65x55x0.67**	**26/28**	**12/15**	**Teflon glass**

For 26 GHz, the proposed antenna presents six beams located at θ (+,-) = 80°, 90° and 100° with a maximum directivity of 12.1 dBi for both. For 27.5 GHz, the antenna presents 14 beams located at θ(+,-) = 40°, 50°, 70°, 90°, 110°, 130° and 140°, with a maximum directivity of 15.1 dBi. The simulation and optimization have been carried out using CST (Computer Simulation Technology) microwave software.

In Section II, the antenna array structures are given with all dimensions and details. In Section III, simulations of S_{11} parameter and radiation pattern are performed and discussed. We conclude the paper in Section IV.

II. ANTENNA DESIGN

The proposed elementary antenna model consists of a horizontal dipole antenna with a length of 30 mm and six vertical elementary wires forming a sub-antenna array. This antenna has been designed on Teflon glass substrate with a

 ISBN: 978-1-61208-940-9

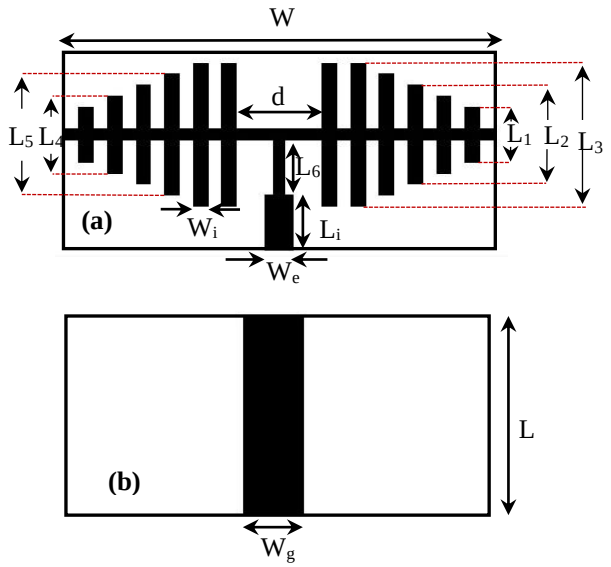


Figure 1. Proposed elementary antenna model, (a): the top side, (b): the bottom side

relative permittivity of 2.1 and a thickness of 0.67 mm. A microstrip line has been used to feed this antenna with an input impedance of 50 Ω. Figure 1 shows the proposed antenna model and its dimensions are given in Table II.

TABLE II: ALL ANTENNA DIMENSION IN (mm)

L_1	L_2	L_3	L_4	L_5	L_6	L_i	L	d	W_i	W_e	W_g	W
5	9	13	7	11	5	5	15	6	1	2	4	30

The first step of the proposed antenna array design is to stack two elementary antennas (of Figure 1) to form a two-antenna (2 x 1) array, as shown in Figure 2. Table III gives the (2 x 1) antenna array dimensions.

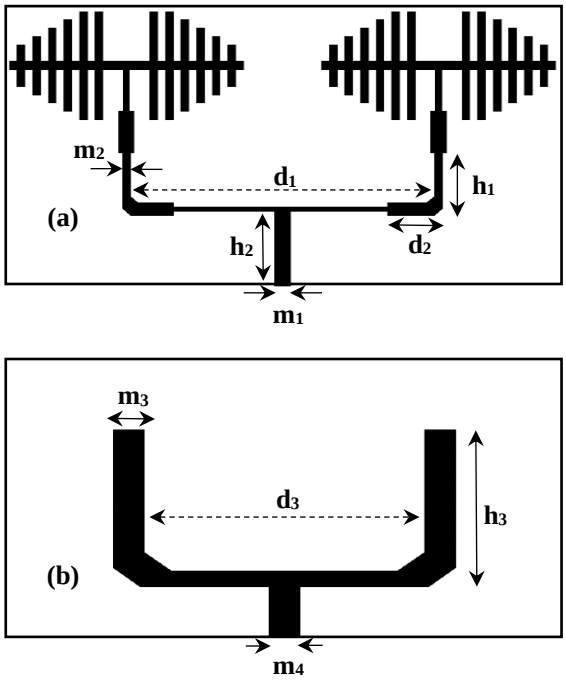


Figure 2. (2 x 1) antenna array model, (a): the top side, (b): the bottom side

TABLE III: (2 x 1) ANTENNA ARRAY DIMENSION IN (mm)

m1	m2	m3	m4	h1	h2	h3	d1	d2	d3
2	1	4	4	7.5	9	9	39	7	36

The second step is to design the (2 x 2) antenna array. In this part, two elementary antennas are added with the goal to increase the gain and directivity of the antenna. The challenge of this part is the impedance matching. Each antenna has to be matched to the equivalent impedance of the other three elementary antennas. For this purpose, a quarter-wave lines and an open stub are used, as shown in Figure 3. Table IV gives the complementary dimensions.

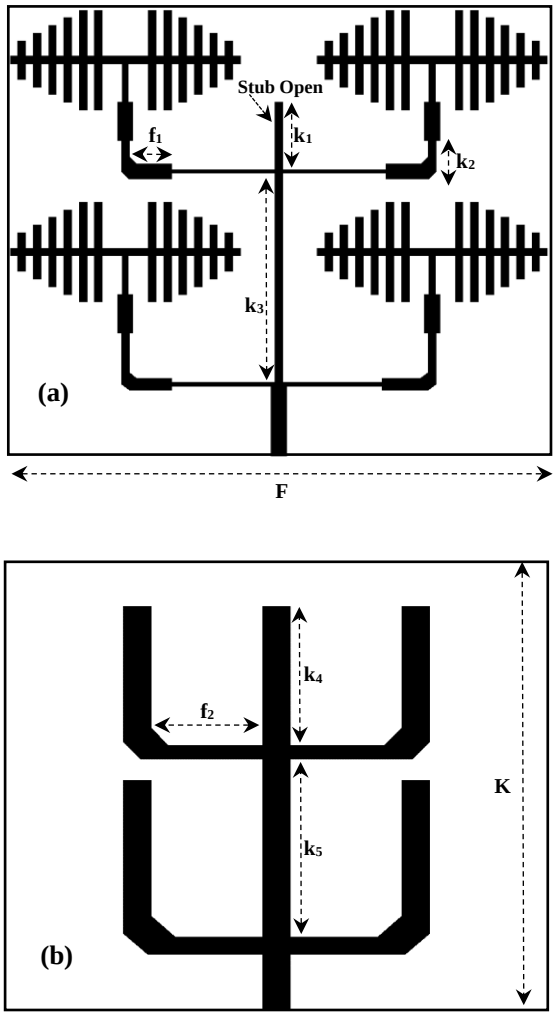


Figure 3. (2 x 2) antenna array model, (a): the top side, (b): the bottom side

TABLE IV: (2 x 2) ANTENNA ARRAY DIMENSION IN (mm)

k1	k2	k3	k4	k5	f1	f2	F	K
8.75	5	27.25	20	25.5	5.5	16	65	55

III. Simulation and Discussion

A. Elementary antenna

Figure 4 presents the simulated coefficient reflection of the elementary antenna (see Figure 1).

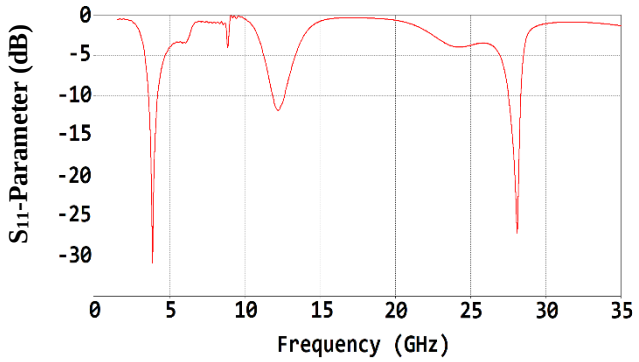


Figure 4. S_{11} parameter of elementary antenna

The elementary antenna is well adapted to the two 5G bands (3.5 GHz and 28 GHz) with an S_{11} parameter of -30 dB and -27 dB, respectively.

Figure 5 depicts the simulated 3D radiation pattern of the elementary antenna for 3.5 GHz and 28 GHz.

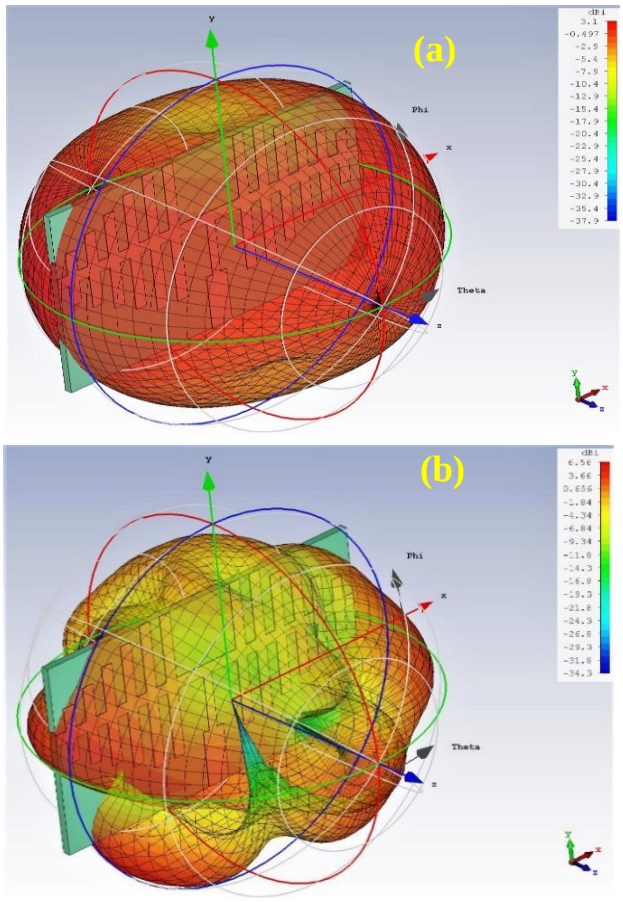


Figure 5. Simulated radiation pattern of elementary antenna. (a) : 3.5 GHz, (b) : 28 GHz.

The maximum directivity obtained by the proposed elementary antenna is 6.5 dBi at 28 GHz and 3 dBi at 3.5 GHz. The beam forming technology is adopted by 5G in the 24/26/28 GHz bands, etc. For this reason, we focus our study of the antenna array in the two 26/28 GHz 5G bands.

B. (2 x 1) proposed antenna array

Figure 6 shows the simulated coefficient reflection of the (2 x 1) antenna array (see Figure 2).

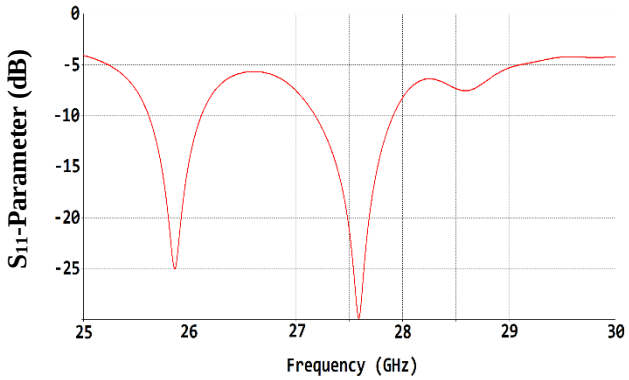


Figure 6. S_{11} parameter of the proposed (2 x 1) antenna array.

It can be seen that the proposed (2 x 1) antenna array is well adapted to both 26/28 GHz 5G bands, with a S_{11} of -25 dB and -30 dB respectively.

Figures 7 and 8 present the simulated radiation pattern of (2 x 1) antenna array (E and H) plans, respectively.

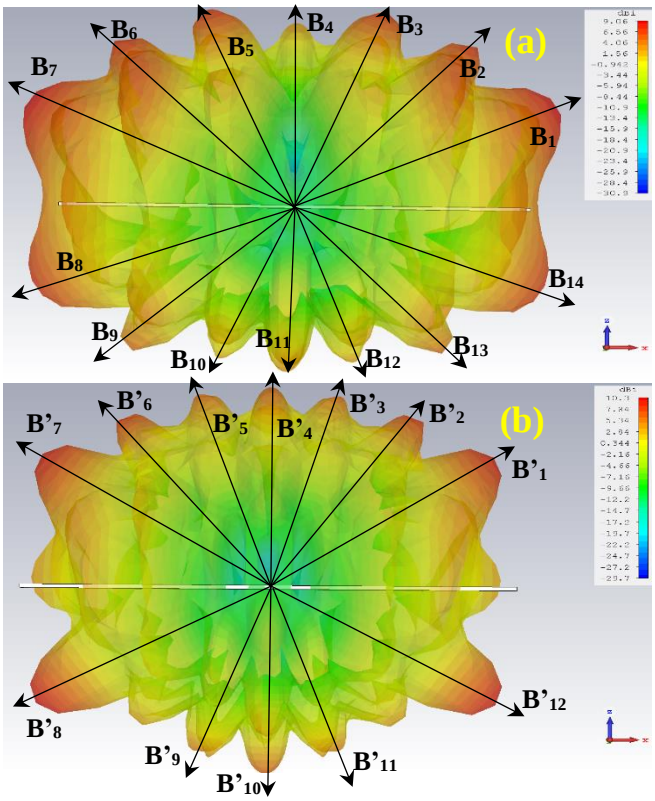


Figure 7. E_Plan simulated radiation pattern of (2 x 1) antenna array. (a) : 26 GHz, (b) : 28 GHz.

 ISBN: 978-1-61208-940-9

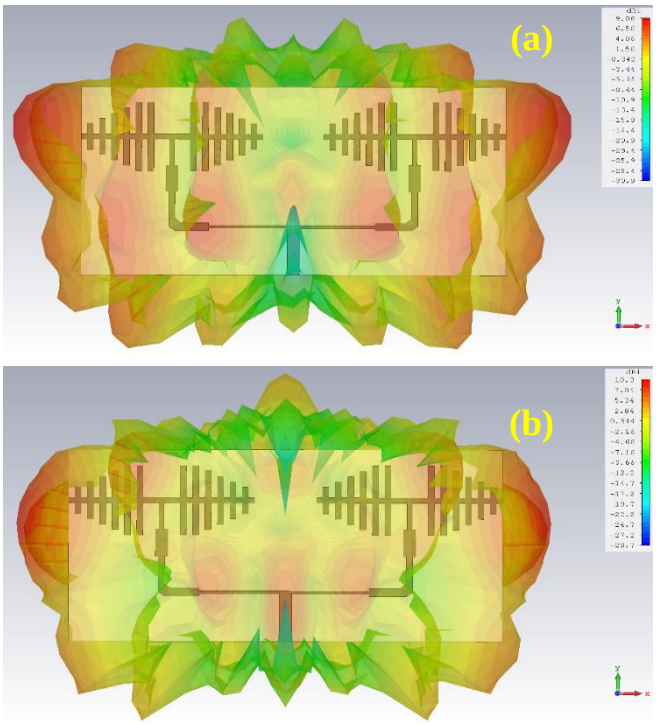


Figure 8. H_Plan simulated radiation pattern of (2 x 1) antenna array. (a) : 26 GHz, (b) : 28 GHz.

It can be noted that the array antenna (2 x 1) is a multi-beam with a high gain for both bands (26/28 GHz). For 26 GHz, the antenna has 14 lobes and 12 powerful lobes for 28 GHz. The maximum gain is of 9.6 dBi and 10.3 dBi, respectively for 26 GHz and 28 GHz 5G bands. Table V gives all directivity values for each beam.

TABLE V: DIRECTIVITY VALUES FOR EACH BEAM IN (dBi)

B1	**B2**	**B3**	**B4**	**B5**	**B6**	**B7**	**B8**	**B9**	**B10**
9	8.8	9	4	7.8	8.1	9	8.8	5.2	4.1
B11	**B12**	**B13**	**B14**	**B'1**	**B'2**	**B'3**	**B'4**	**B'5**	**B'6**
5	4.6	5.6	8.6	9.8	8.4	7.8	7.2	7.1	8.5
B'7	**B'8**	**B'9**	**B'10**	**B'11**	**B'12**				
9.7	10.2	4	8.1	4.8	10.1				

C. *(2 x 2) proposed antenna array*

Figure 9 depicts the simulated S_{11} parameter of the proposed (2 x 2) antenna array (see Figure 3).

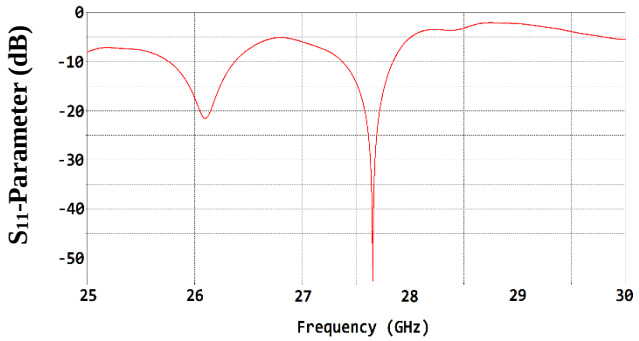


Figure 9. Simulated S_{11} parameter of (2 x 2) antenna array.

It can be noted that the proposed (2 x 2) antenna array is more adapted to 27.6 GHz band with a S_{11} of -55 dB, and adapted to 26 GHz band with a S_{11} of -23 dB.

Figures 10 and 11 depict the simulated radiation pattern of (2 x 2) antenna array (E and H) plans, respectively.

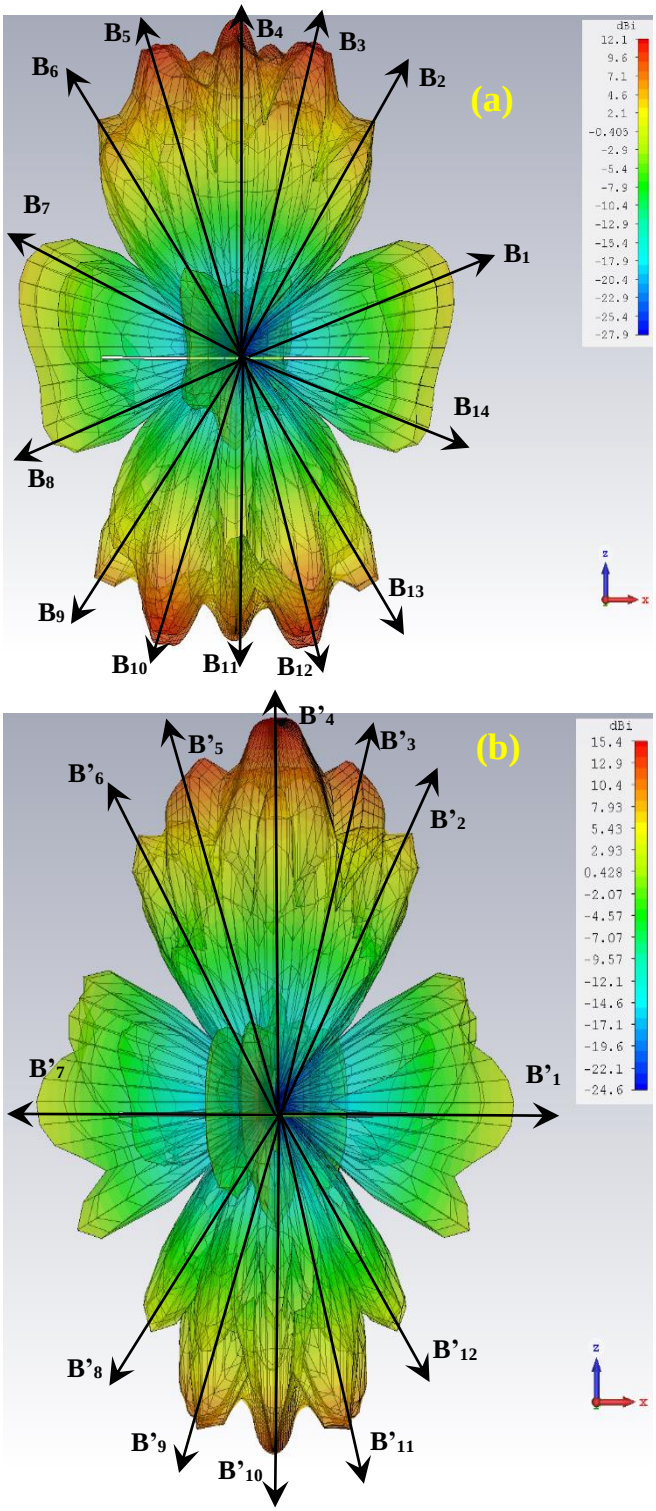


Figure 10. E_Plan simulated radiation pattern of (2 x 2) antenna array. (a) : 26 GHz, (b) : 27.6 GHz.

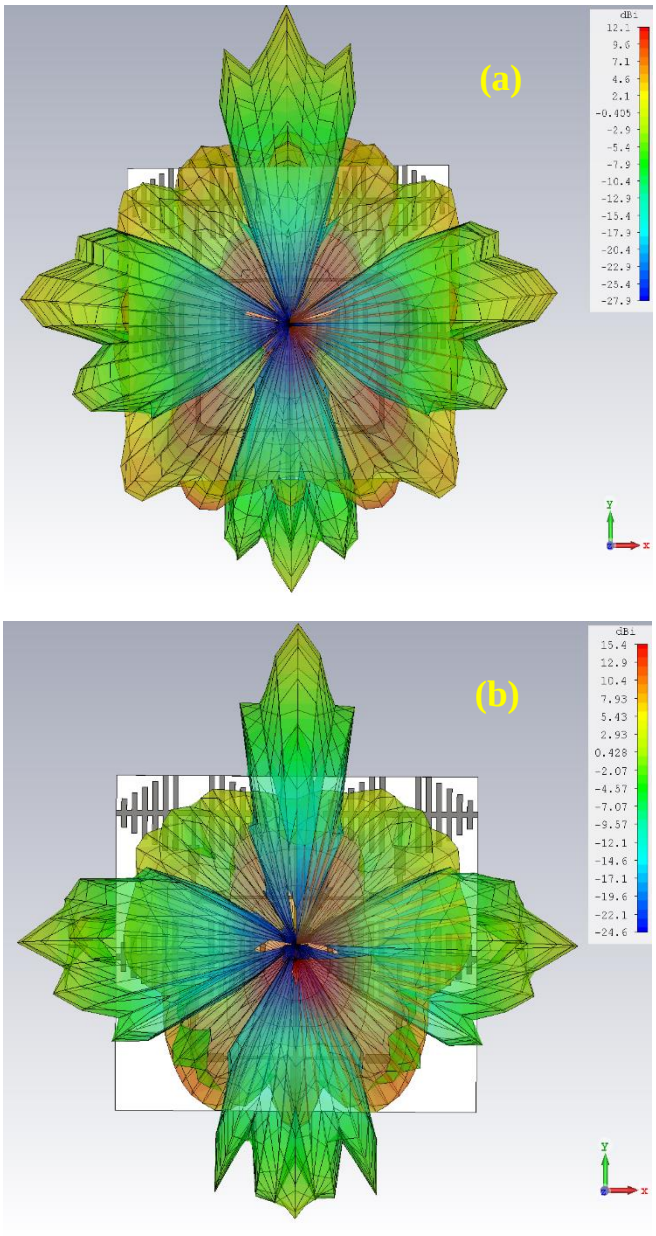


Figure 11. H_Plan simulated radiation pattern of (2 x 2) antenna array. (a) : 26 GHz, (b) : 27.6 GHz.

14 beams with high directivity are visualized by the radiation pattern of the antenna in E_Plan. The directivity values for each beam are given in Table VI.

TABLE VI: DIRECTIVITY VALUES FOR EACH BEAM IN (dBi)

B1	**B2**	**B3**	**B4**	**B5**	**B6**	**B7**	**B8**	**B9**	**B10**
5	7	10.1	12.1	10	7	5	4	7.1	11.1
B11	**B12**	**B13**	**B14**	**B'1**	**B'2**	**B'3**	**B'4**	**B'5**	**B'6**
9.2	11	7.2	4	3	7.3	12	15.4	12	7.3
B'7	**B'8**	**B'9**	**B'10**	**B'11**	**B'12**				
5	5.7	12.5	14.5	12.5	5.7				

For the 26 GHz and 27.6 GHz frequencies, 6 powerful beams are located, 3 in front and 3 in the back of the antenna. This allows the antenna to be operational at both sides. Figures 12 and 13 present the variation of directivity as a function of the direction angles for both frequencies in E and H plans, respectively.

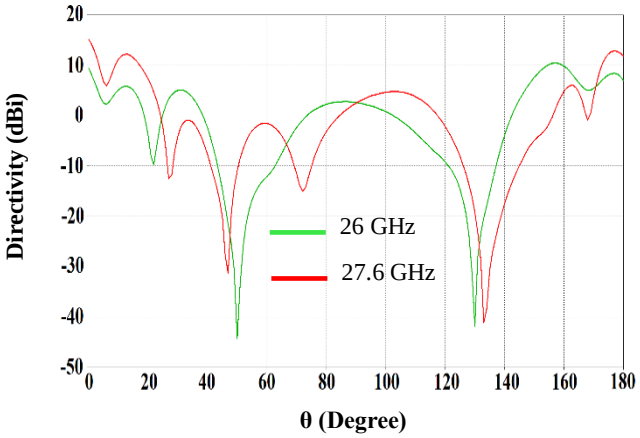


Figure 12. Directivity variation as a function of the direction direction angle in E_Plan

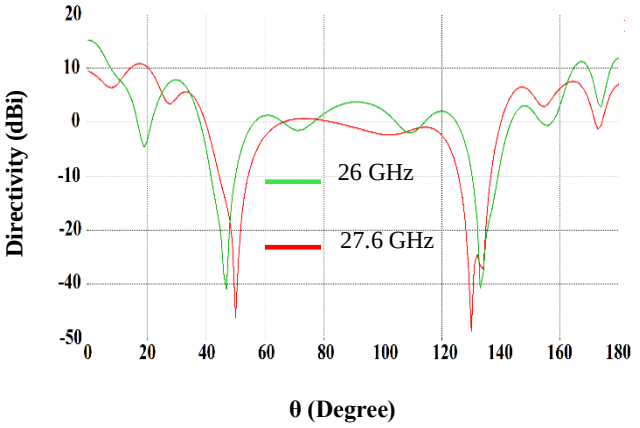


Figure 13. Directivity variation as a function of the direction angle in H_Plan

It can be noted that the maximum directivities are located at 160° for 26 GHz and 180° for 27.6 GHz. The maximum directivity values at the previous angles are 12 dBi and 15.4 dBi respectively.

IV CONCLUSION

In this paper, a high directivity multi-beam antenna for Radio Frequency Energy harvesting in 5G network is presented. This antenna has a size of 65 x 55 mm² designed on Teflon glass substrate with a relative permittivity of 2.1 and a thickness of 0.67mm. The maximum directivity obtained is 12.1 dBi for the 26 GHz frequency and 15.4 dBi for 27.5 GHz. The multibeam property makes this antenna capable of picking up waves at both frequencies (26 and 27.5) GHz in several directions with high directivity, which largely increases the received power.

 ISBN: 978-1-61208-940-9

REFERENCES

[1] A. J. Compston, J. D. Fluhler and H. G. Schantz, "A Fundamental Limit on Antenna Gain for Electrically Small Antennas," 2008 IEEE Sarnoff Symposium, 2008, pp. 1-5, doi: 10.1109/SARNOF.2008.4520113.

[2] J. Wu, C. Wang and Y. X. Guo, "A Compact Reflector Antenna Fed by a Composite S/Ka-Band Feed for 5G Wireless Communications," in IEEE Transactions on Antennas and Propagation, vol. 68, no. 12, pp. 7813-7821, Dec. 2020, doi: 10.1109/TAP.2020.3000858.

[3] M. Mirzaee and N. Tavassolian, "Low-Profile Wearable Wideband Antenna with High Gain Based on Franklin Array for Future 5G Wireless Body Area Networks," 2020 IEEE International Symposium on Antennas and Propagation and North American Radio Science Meeting, 2020, pp. 449-450, doi: 10.1109/IEEECONF35879.2020 9329727.

[4] M. Heino, C. Icheln, J. Haarla and K. Haneda, "PCB-Based Design of a Beamsteerable Array With High-Gain Antennas and a Rotman Lens at 28 GHz," in IEEE Antennas and Wireless Propagation Letters, vol. 19, no. 10, pp. 1754-1758, Oct. 2020, doi : 10.1109/LAWP.2020.3017 129.

[5] S. Kim and J. Choi, "1×8 Slotted Array Antenna with Fan-Beam Characteristics for 28 GHz 5G Mobile Applications , " 2020 International Symposium on Antennas and Propagation (ISAP), 2021, pp. 13-14, doi: 10.23919/ISAP47053.2021.9391404.

[6] L. Vähä-Savo et al., "Empirical Evaluation of a 28 GHz Antenna Array on a 5G Mobile Phone Using a Body Phantom, " in IEEE Transactions on Antennas and Propagation, doi: 10.1109/TAP.2021.3076535.

 ISBN: 978-1-61208-940-9

Secure PMIPv6-Based Mobility Solution for LoRaWAN

Hassan Jradi*†, Abed Ellatif Samhat*, Fabienne Nouvel†, Mohamad Mroue*, Jean-Christophe Prévotet†
*Lebanese University — CRSI, Hadath, Lebanon.
email: {samhat, mohamad.mroue}@ul.edu.lb
†INSA de Rennes — IETR, Rennes, France.
email: firstname.lastname@insa-rennes.fr

***Abstract*—The widespread use of Internet of Things (IoT) has stimulated the invention of new communication technologies like Long Range Wide Area Network (LoRaWAN) and Narrow Band-Internet of Things (NB-IoT) belonging to Low Power Wide Area Network (LPWAN) technologies. The wide use of these technologies brings new requirements like mobility management. Proxy Mobile IPv6 (PMIPv6) is a network-based mobility management protocol. However, integrating PMIPv6 in LPWAN rises a special challenge due to LPWAN constraints. In addition, PMIPv6 does not provide secure access to the operator domain. In this paper, we propose a new PMIPv6-based mobility solution for LoRaWAN boosted with an authentication scheme to access the operator domain and make this solution resist several types of attacks. In addition, we evaluate the performance of our scheme and we compare it with other works. Finally, we evaluate the security of the new scheme using Automated Validation of Internet Security Protocols and Applications (AVISPA) tool.**

Keywords – IoT; LPWAN; LoRaWAN; Mobility; Authentication.

I. INTRODUCTION

The Internet of Things (IoT) is the novel era of wireless communication carrying out several services ranging from data sensing to command execution [1]. New communication technologies designed to meet the needs of long communication range with low data rates and power consumption are invented and categorized under Low Power Wide Area Network (LPWAN) [2]. Long Range Wide Area Network (LoRaWAN) is one of the most prominent LPWAN technologies operating in the unlicensed Industrial, Scientific, and Medical radio band (ISM band) [3].

However, several applications like healthcare supervising and supply chain monitoring, require a secure mobility management protocol to ensure session continuity and secure access to the operator network [4]. Proxy Mobile IPv6 (PMIPv6) [5] is one of IPv6 protocol extensions designed to provide network-based mobility protocol. The mobility procedure is executed by an entity on behalf of the Mobile Node (MN) which minimizes the power consumed by it.

However, PMIPv6 does not deploy an authentication mechanism which is essential in case of an MN wishing to join the network. Several authentication schemes are proposed to be used in PMIPv6 as in [6][7]. Nonetheless, the used solution should be well adapted to work in LoRaWAN environment taking into consideration LPWAN constraints like payload length, data rate range, and number of messages per day.

Contribution. The main contribution of this paper is the proposal of a new mobility solution based on PMIPv6 protocol with an authentication scheme providing both intra-domain and inter-domain authentication for LoRaWAN.

Paper Organization. In Section II, we present some related work and we describe the problem. In Section III, we present the proposed mobility solution along with the authentication scheme. Section IV shows the results and the comparison with related work and Section V concludes the paper.

II. BACKGROUND AND RELATED WORK

Mobility is the movement of an MN leading to the release of the connection with the current Point of Attachment (PoA) and the establishment of the connection with a new PoA [8]. The deployed mobility management protocol has a major role in the connection release/establishment procedures from performance and security points of view.

Several types of mobility can be identified according to the handoff scenario. Handoff is the process of release of the old connection and establishment of the new connection. Thus, mobility can be homogeneous or heterogeneous if the previous and the new PoAs use the same or different link layer technologies, respectively. Mobility can be also intra-domain or inter-domain (also known as roaming) if the previous and the new PoAs belong to the same or different network operators, respectively. Consequently, several schemes are proposed to deal with the mobility challenge in LPWANs taking into consideration the security aspect.

The work done by Moosavi *et al.* [9] aims to provide a mobility management solution for IoT by splitting the network into two virtual layers. The intermediate processing layer consists of smart gateways that manage devices mobility, and the cloud layer consists of data analysis servers. This solution enables an end-to-end security solution between the MN and the end-user by providing authentication and data encryption, and at the same time provides session resumption after the handoff phase.

Another work done by Kang *et al.* [6] addresses the problem of lack of authentication in PMIPv6 protocol. This work focuses mainly on the PMIPv6 protocol without taking into consideration the IoT requirements. Thus, the proposed solution is a general solution and could not be adapted directly to IoT or LPWANs.

In Sharma *et al.* [7] work, the authors proposed a mobility management solution based on Fast Proxy Mobile IPv6 (FPMIPv6) to provide a proactive handoff approach, and based on Media Independent Handover (MIH) framework [10] which is a framework providing heterogeneous handoff using three

 ISBN: 978-1-61208-940-9

MIH services. Moreover, the authors propose an authentication scheme based on pre-shared keys to provide secure MIH communication between the MN and the network entities. This solution was intended for IoT and not for LPWAN having more constraints, thus it cannot be directly adapted into LPWANs.

In the work of Ayoub *et al.* [11], the authors proposed to use a modified version of Static Context Header Compression protocol (SCHC) [12] protocol named Dynamic Context Header Compression (DCHC) protocol along with the use of Mobile IPv6 (MIPv6) and a light version of MIH framework. This work was designed to operate in an LPWAN environment especially with LoRaWAN and Narrow Band-Internet of Things (NB-IoT).

Thus, as shown, several works try to deal with the mobility aspect of devices. However, these works either cannot be directly integrated into LPWAN, or do not provide a security mechanism for network access. In the next section, we present a new PMIPv6-based mobility solution boosted with an authentication mechanism taking into consideration LPWAN constraints.

III. PROPOSED SOLUTION

In this section, we present our mobility solution that provides both intra-domain and inter-domain mobility types for LoRaWAN, where the MN may move inside or outside its home operator network coverage.

A. Protocol Stack

The protocol stack used for the communication between the MN and the network is represented in Figure 1.

The upper layers consist of application and transport layers which are dependent on the deployment purpose of the network. These layers are used to send/receive the application data.

The network layer consists of IPv6 as a routing protocol and PMIPv6 as a network layer mobility management protocol. As we are dealing with LoRaWAN technology which is considered as a layer 2 or link-layer technology, we propose to modify the LoRaWAN protocol stack to be operable with the added IPv6 and PMIPv6 network layer functions. The integration of PMIPv6 requires the adoption of PMIPv6 network architecture, as discussed in the following subsection. However, the addition of this layer leads to additional overhead which should be examined carefully in LoRaWAN since the maximum payload length is 256 bytes. The advantage of using IPv6 is to achieve global mobility independently of the lower layer technology, since each technology can deploy its lower layer mobility protocol.

For that, we propose to use an adaptation layer to overcome the previous problem. In this layer, Static Context Header Compression protocol (SCHC) [12] is used to compress the IPv6 packet headers in order to fit suitable payload lengths for LPWANs. Upon a connection establishment, the sender and the receiver agree on a SCHC context. This context

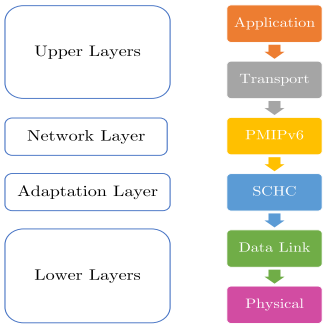


Figure 1. Mobile Node Protocol Stack.

contains several rules identified by RuleID. Each rule contains a list of entries. An entry contains a field identifier, a Compression/Decompression (C/D) action, a target value and a matching operator. An uplink packet header field is compared with the target value according to the matching operator, and if the comparison test succeeds, the C/D action is executed. In this way, the RuleID and the compression residues are sent instead of sending the entire header. At the receiver side, the reverse process is executed.

The lower layers consist of the data link and physical layers of the used LPWAN technology. In the case of LoRaWAN, the data link layer is LoRaWAN Media Access Control Layer and the physical layer is LoRa physical layer.

B. Network Architecture

The main entities in LoRaWAN are the Network Server (NS), the Join Server (JS) and the Gateways (GWs). The improved LoRaWAN architecture is called evolved LoRaWAN and is shown in Figure 2. We endeavored to integrate the two necessary PMIPv6 entities, which are the Media Access Gateway (MAG) and the Local Mobility Anchor (LMA), in the LoRaWAN architecture.

We propose to place the LMA functionalities within the NS since the latter is the anchor point of LoRaWAN architecture. Furthermore, a new entity called LoRa Mobile Access Gateway (LoRaMAG), is inserted between GWs and NS. Therefore, several GWs will be connected to one LoRaMAG and an MN should authenticate itself with the new LoRaMAG when it moves from a GW to another connected to a different LoRaMAG. LoRaMAG will play the role of the MAG of PMIPv6 architecture. It is responsible for the detection of MN movement, initiating the mobility signaling with the LMA, and data forwarding between MN and LMA through the dedicated tunnel.

The use of PMIPv6 adds more scalability where the NS functions are divided over several LoRaMAGs like the downlink GW selection. In addition, PMIPv6 is known to be to suitable for constrained devices since MIPv6 binding update messages are executed by MAG on the MN behalf.

Another entity used for the authentication between the MN and the LoRaMAG called the Authentication Server (AuS) is added also in each PMIPv6 domain (which is in this case the LoRaWA network). The detailed operation of the AuS function is shown in the next subsection.

 ISBN: 978-1-61208-940-9

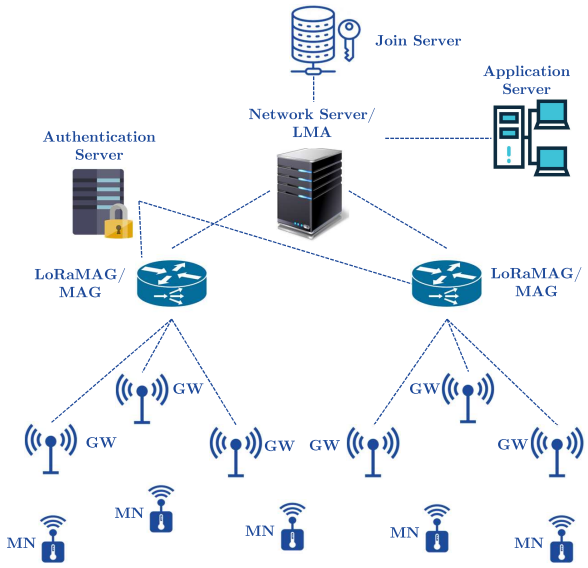


Figure 2. Evolved LoRaWAN Network Architecture.

C. Authentication Scheme

The proposed authentication scheme is used to authenticate the MN with the LMA in the PMIPv6 domain using AuS entity. In the following, we distinguish between two scenarios: intra-domain mobility and inter-domain mobility (or roaming).

In case of intra-domain mobility, the MN moves inside the coverage of a GW connected to a LoRaMAG belonging to its home domain, thus MN credentials are saved in the domain AuS where the MN is initially registered.

In case of inter-domain mobility, the MN moves towards the coverage of GWs of a visited domain having its visited LMA (vLMA), visited AuS (vAuS), and will be connected through visited LoRaMAG (vLoRaMAG). The MN is initially registered in its home domain in home AuS (hAuS) having MN credentials.

The proposed scheme consists of two phases: the registration phase and the authentication phase. We will present the authentication phase in case of roaming (device moving in the visited operator coverage).

In this scheme, we tried to integrate the PMIPv6 signaling with the authentication scheme signaling. This solution is compatible with class A LoRaWAN devices based on one transmission frame followed by two reception frames.

1) Registration Phase: The hAuS holds two secret keys X and Y which are only known by itself. Then, a MN_i having an identity ID_i and the hAuS holds two pre-shared keys:

- $X_i = H(H(X) \oplus ID_i)$.
- $Y_i = H(H(Y) \oplus ID_i)$.

We presume that we have secure links between {GWs and vLoRaMAG}, {vLoRaMAG and vAuS}, {vAuS and hAuS}. so that data confidentiality and integrity are ensured on these links. These links can be secured using Public Key Infrastructure (PKI) [13] or any authentication and key agreement scheme. We focus on the {MN and vLoRaMAG} link where LoRaWAN limitations are present.

2) Authentication Phase: In this phase, the MN tries to authenticate itself in the visited PMIPv6 domain with the vLoRaMAG through vAuS and hAuS, using the exchanges shown in Figure 3. Since the GWs only forward the messages, we do not represent them for more clarity. Moreover, this phase is divided into two sub-phases: home authentication sub-phase, and visited authentication sub-phase.

In home authentication sub-phase (red part in Figure 3), the MN sends its authentication request which passes to hAuS through vAuS. The hAuS checks the authentication request validity and derives two keys and shares them with vAuS. This sub-phase is executed **in case of roaming only and once per visited domain**.

In visited authentication sub-phase (green part in Figure 3), after the vAuS gets the visited keys from the hAuS, it uses them to authenticate the MN as long as it is in the visited domain without the need to send requests to hAuS. So after the first sub-phase, the second sub-phase can be repeated several times to authenticate the mobile when it moves between different vLoRaMAGs.

The message exchanges are detailed below. T_1 through T_4 are timestamp variables used to prevent replay attack.

1) MN_i computes $K_i = H(X_i) \oplus H(Y_i)$ and $MIC_1 = H(ID_i \| T_1 \| K_i)$ then sends a message with *AuthReq* tag consisting of $\{ID_i \| ID_{hAuS} \| T_1 \| MIC_1\}$.
2) vAuS checks the requested AuS by inspecting the second field of the request. In this case, the requested AuS is hAuS thus vAuS forwards this request to hAuS.
3) hAuS receives the request and gets the identity ID_i, then hAuS queries its database for the corresponding keys X_i and Y_i. Thereafter, hAuS computes $K_i = H(X_i) \oplus H(Y_i)$ and checks if $MIC_1 = H(ID_i \| T_1 \| K_i)$.
4) hAuS generates a random nonce N and computes two derived keys $vX_i = H(X_i \oplus N)$ and $vY_i = H(Y_i \oplus N)$. These two derived keys are intended to be sent to vAuS. Moreover, hAuS computes $MIC_2 = H(ID_i \| N \| T_2 \| K_i)$. hAuS sends a message with *RoamingAuthResp* tag consisting of $\{ID_i \| vX_i \| vY_i \| N \| T_2 \| MIC_2\}$. Note that this message is sent over a secure link.
5) vAuS receives the response and gets vX_i and vY_i, then saves them along with ID_i in its database. Thereafter, vAuS forwards the rest of the response to MN_i with its identity ID_{vAuS}. A mapping between ID_i and $DevAdd_i$ is saved in the vLMA/NS.
6) MN_i receives the response and checks if $MIC_2 = H(ID_i \| N \| T_2 \| K_i)$. MN_i gets N from the message then computes $vX_i = H(X_i \oplus N)$ and $vY_i = H(Y_i \oplus N)$ to be used for the authentication in the visited domain. At this step, home authentication sub-phase is finished and should not be executed again as long as MN_i is inside this domain.
7) After the reception of the $RoamingAuthResp$ by the vLoRaMAG in case of first authentication request (home authentication sub-phase), or in case of attach event detection by vLoRaMAG in second or upper MN_i at-

 ISBN: 978-1-61208-940-9

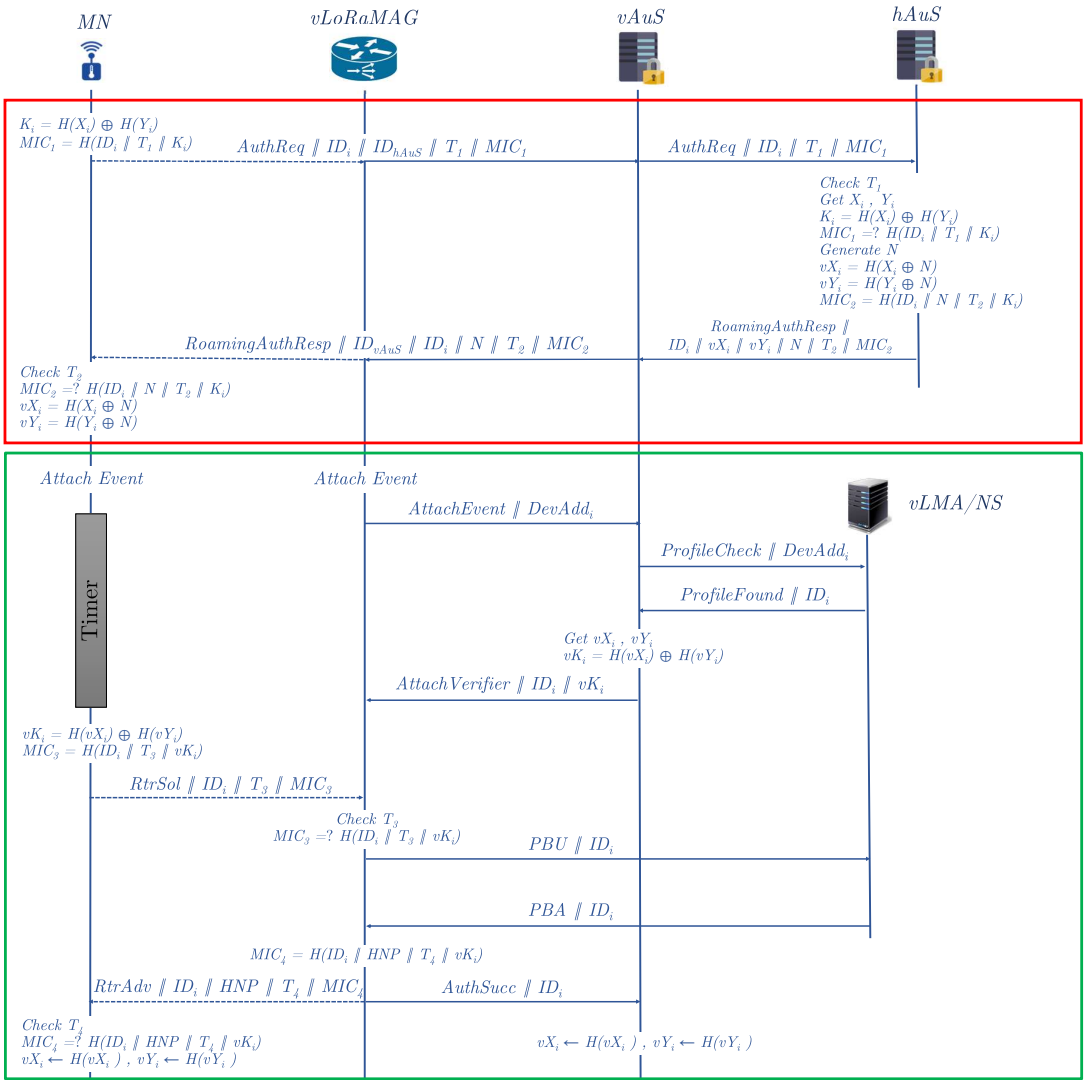


Figure 3. Message Exchange During Authentication Phase.

tachment (after a successful LoRaWAN Join Procedure); vLoRaMAG sends a message with *AttachEvent* tag consisting of $\{DevAdd_i\}$ to vAuS. This step is the first step of visited domain authentication sub-phase.

8) vAuS receives the attach event notification from vLoRaMAG then sends a message with *ProfileCheck* tag consisting of $\{DevAdd_i\}$ to vAuS/NS. vAuS/NS in its turn replies with a message with *ProfileFound* tag consisting of the corresponding $\{ID_i\}$.
9) vAuS queries its database based on ID_i to get vX_i and vY_i, then it computes $vK_i = H(vX_i) \oplus H(vY_i)$ and sends a message with *AttachVerifier* tag consisting of $\{ID_i \| vK_i\}$ to vAuS/NS. Note that this message is sent over a secure link.
10) After elapsing the timer launched by MN_i after the link layer attach, which is configured to be equivalent to the duration of the four previous exchanges, MN_i computes $vK_i = H(vX_i) \oplus H(vY_i)$ and $MIC_3 = H(ID_i \| T_3 \| vK_i)$. Then sends a message with *RtrSol* tag consisting of $\{ID_i \| T_3 \| MIC_3\}$ to vLoRaMAG in order to get a *RtrAdv* message to configure its network layer interface.
11) vLoRaMAG receives the *RtrSol* message and checks if $MIC_3 = H(ID_i \| T_3 \| vK_i)$. If so, vLoRaMAG sends a Proxy Binding Update (PBU) message along with ID_i to vLMA which is also the NS. Therefore vLMA performs the needed operations according to PMIPv6 protocol to register/update the Binding Cache Entry (BCE) of MN_i. Then it replies with Proxy Binding Acknowledgment (PBA) message along with ID_i to vLoRaMAG.
12) vLoRaMAG accepts the PBA message and computes $MIC_4 = H(ID_i \| HNP \| T_4 \| vK_i)$ and sends a message with *RtrAdv* tag consisting of $\{ID_i \| HNP \| T_4 \| MIC_4\}$ to MN_i. Home Network Prefix (HNP) is the network prefix corresponding to MN_i. vLoRaMAG sends another message with *AuthSucc* tag along with ID_i to vAuS to confirm the authentication success.

13) MN_i receives the *RtrAdv* message and checks if $MIC_4 = H(ID_i \| HNP \| T_4 \| vK_i)$ where it can now configure its network layer interface using HNP.
14) MN_i and vAuS update the two derived keys by performing the following operations $vX_i \leftarrow H(vX_i)$ and $vY_i \leftarrow H(vY_i)$ which will be used in the next authentication trial.

IV. RESULTS AND ANALYSIS

In this section, we present the performance evaluation and the security analysis of our solution. In addition, we compare the performance evaluation and the security features of our solution with related work.

A. Performance Evaluation

We evaluated the performance of the proposed authentication scheme by simulation using Network Simulator 3 (NS-3). The simulation scenario consists of the entities used in the authentication scheme. The link between MN and GW is a LoRaWAN radio link and is considered an unsecured link. The MN is trying to authenticate itself to the visited domain using the proposed scheme. The source code of implementation can be found in [14].

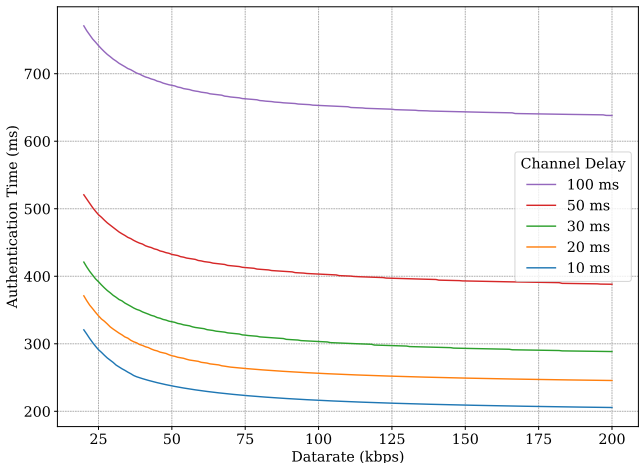


Figure 4. Authentication Time for Variable Data Rates and Channel Delays.

The evaluation metric is the time needed to perform the authentication scheme and the simulation parameters are the data rate (R_b) and the channel delay (τ) for the link between MN and GW. For that, we run simulations at LoRaWAN data rate range, i.e., R_b from 20 to 200 kbps and at $\tau \in \{10, 20, 30, 50, 100\}\ ms$. The results are shown in Figure 4.

In Figure 5, we show the overall handoff latency for related work presented in Section II. The results show that our solution provides competitive results with other solutions where we work on low data rates and we provide inter-domain authentication. Moosavi *et al.* [9] and Sharma *et al.* [7] use high data rates reaching 8 Mbps whereas in Ayoub *et al.* [11], and in this work, the data rates used are that used in LoRaWAN (between 20 to 200 kbps) forming a low latency mobility solution.

Moreover, the longest message payload on the LoRaWAN link is that tagged with $RoamingAuthResp$. The hash used in this scheme is SHA-256 thus hash length ($L_{Hash} = 32\ Bytes$). The lengths of identities, nonce and timestamp are respectively $L_{ID} = 4\ Bytes$, $L_{Nonce} = 8\ Bytes$ and $L_{Timestamp} = 10\ Bytes$. Thus $L_{Payload} = 2 \times L_{ID_i} + L_{Nonce} + L_{Timestamp} + L_{Hash} = 58\ Bytes < 256\ Bytes$ (Maximum LoRaWAN payload length). Thus, the authentication mechanism is suitable for LoRaWAN technology and more particularly for class A devices since it is based on reception after transmission.

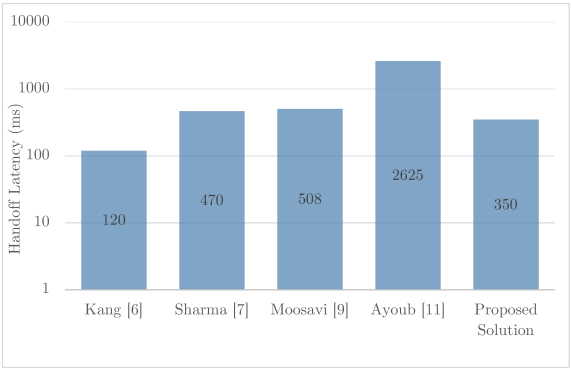


Figure 5. Performance Comparison of the Proposed Mobility Solutions.

B. Security Analysis

We assess the security of the proposed authentication scheme based on attacks related to device mobility as mentioned in our previous work [2].

- ◇ Device re-authentication: the proposed authentication scheme aims to provide secure access when the mobile node moves between different domains, thus it can be identified and authenticated in case of intra-domain and inter-domain mobility.
- ◇ Spoofing signaling message: the exchanged signaling messages between MN and network entities are integrity protected using MIC field through K_i or vK_i keys only known by the concerned entities. Thus, an attacker cannot modify the content of these messages without being detected.
- ◇ Address squatting and spoofing: an attacker cannot squat or spoof the device address since HNP is provided to MN during the authentication phase based on PMIPv6 specifications. And the network layer interface is configured after the authentication phase based on the provided HNP.
- ◇ Old address control: the MN IPv6 address is re-configured after the handoff phase based on the received IPv6 HNP. Thus, an attacker uses a device address without the completion of the authentication phase.
- ◇ Mutual authentication: the authentication between the mobile node and the hAuS is ensured using the hash key K_i, and between the MN and the vLoRaMAG using the hash key vK_i. These keys are confidential and cannot be derived by an attacker since it does have and cannot predict the key materials X_i, Y_i, vX_i and vY_i.

 ISBN: 978-1-61208-940-9

◇ Key freshness: the hash key vK_i is calculated during each authentication trial by a way it cannot be predicted in the next authentication trial based on the use of vX_i, vY_i. Even if vLoRaMAG having vK_i cannot predict it at next trial.
◇ Replay attack: this kind of attack is prevented by the use of timestamps $T1$ through $T4$.

In Table I, we compare the security features provided by related work presented in Section II.

TABLE I
COMPARISON OF SOLUTIONS ACCORDING TO SECURITY ISSUES

	Kang *et al.*[6]	Sharma *et al.*[7]	Moosavi *et al.*[9]	Ayoub *et al.*[11]	Proposed Solution
Device re-authentication	✓	✓	✓	✗	✓
Spoofing signaling message	✓	✓	✓	✗	✓
Address squatting and spoofing	✓	✓	✓	✗	✓
Old address control	✗	✓	✓	✗	✓
Mutual authentication	✓	✓	✓	✗	✓
Key freshness	✗	✓	✗	✗	✓
Replay attack	✗	✓	✓	✗	✓
Suitable for LPWAN	✗	✗	✗	✓	✓

✓ : Resistant ✗ : Vulnerable

C. Security Validation using AVISPA

We used Automated Validation of Internet Security Protocols and Applications (AVISPA) [15] as a validation tool for the security of the proposed authentication scheme. The implementation codes using HLPSL language can be found in [14]. Testing the implemented scheme using AVISPA shows that our solution is secure, as shown in Figure 6.

V. CONCLUSION

In this paper, we proposed an inter-domain mobility solution for LoRaWAN. We tried to solve the problem of domain access in PMIPv6 protocol by the use of the proposed authentication mechanism. Our solution is simulated using NS-3 and presents competitive results compared to other works in the literature. We conducted our scheme also through AVIPSA validation tool to prove its security.

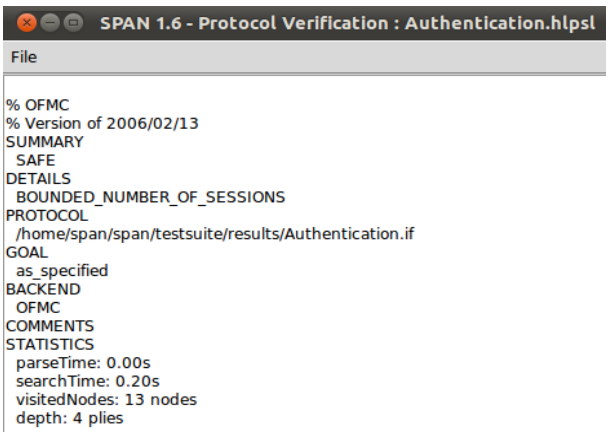


Figure 6. AVISPA Validation of the Proposed Authentication Scheme.

REFERENCES

[1] J. Lin *et al.*, "A survey on internet of things: Architecture, enabling technologies, security and privacy, and applications," *IEEE Internet of Things Journal*, vol. 4, pp. 1125–1142, 2017.
[2] H. Jradi, A. E. Samhat, F. Nouvel, M. Mroue, and J.-C. Prévotet, "Overview of the mobility related security challenges in lpwans," *Computer Networks*, p. 107 761, 2020.
[3] K. Mekki, E. Bajic, F. Chaxel, and F. Meyer, "A comparative study of lpwan technologies for large-scale iot deployment," *ICT express*, vol. 5, no. 1, pp. 1–7, 2019.
[4] W. Ayoub, A. E. Samhat, F. Nouvel, M. Mroue, and J.-C. Prévotet, "Internet of mobile things: Overview of lorawan, dash7, and nb-iot in lpwans standards and supported mobility," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 2, pp. 1561–1581, 2018.
[5] S. Gundavelli *et al.*, "Proxy mobile ipv6," *IETF RFC 5213*, Aug. 2008.
[6] D. Kang, J. Jung, D. Lee, H. Kim, and D. Won, "Security analysis and enhanced user authentication in proxy mobile ipv6 networks," *Plos one*, vol. 12, no. 7, pp. 1–20, 2017.
[7] V. Sharma *et al.*, "Mih-spfp: Mih-based secure cross-layer handover protocol for fast proxy mobile ipv6-iot networks," *Journal of Network and Computer Applications*, vol. 125, pp. 67–81, 2019.
[8] C. Perkins *et al.*, "Ip mobility support," *IETF RFC 2002*, Oct. 1996.
[9] S. R. Moosavi *et al.*, "End-to-end security scheme for mobility enabled healthcare internet of things," *Future Generation Computer Systems*, vol. 64, pp. 108–124, 2016.
[10] V. Gupta *et al.*, "Ieee802. 21 standard and metropolitan area networks: Media independent handover services," *Draft P802*, vol. 21, p. D00, 2009.
[11] W. Ayoub *et al.*, "Media independent solution for mobility management in heterogeneous lpwan technologies," *Computer Networks*, vol. 182, p. 107 423, 2020.
[12] A. Minaburo, L. Toutain, C. Gomez, D. Barthel, and J.-C. Zúñiga, "Schc: Generic framework for static context header compression and fragmentation," Technical Report RFC8724, Tech. Rep., 2020.
[13] C. Adams and S. Lloyd, *Understanding public-key infrastructure: concepts, standards, and deployment considerations*. Sams Publishing, 1999.
[14] H. Jradi, ns-3 and AVISPA Implementation source codes, retrieved: Mar, 2022. [Online]. Available: github . com / HassanJradi/secure-mobility-secure.git.
[15] L. Vigano, "Automated security protocol analysis with the avispa tool," *Electronic Notes in Theoretical Computer Science*, vol. 155, pp. 61–86, 2006.

 ISBN: 978-1-61208-940-9

Joint Power Control, Pilot Assignment, User Association and Flight Control for Massive MIMO Self-Organizing Drones using Reinforcement Learning

Gabriel Skidmore
Department of Computational Electrical and Computer Engineering
Miami University
Oxford, United States
Email: skidmogm@miamioh.edu

***Abstract*—Improving spectral efficiency is becoming increasingly important in mobile communications to keep up with the ever-increasing amount of data traffic coming from video streaming, Internet of Things, intelligent transportation systems, and augmented and virtual reality. In this work, a deep reinforcement learning algorithm (Deep Q-Learning) is implemented to maximize the sum spectral efficiency of ground users using Unmanned Aerial Vehicles (UAVs) as agents. The agents and environment are created by using OpenAI's Gym library to create a custom implementation of the agent, reward function, and environment. The problem is then relaxed by assigning users to UAVs that lead to the highest Single-Input Single-Output (SISO) Signal to Interference plus Noise Ratio (SINR) and allowing the UAVs to assign multiple pilot signals to ground users. Lastly, the implementation of the algorithm is compared to a convex relaxed version of the original reward function.**

Keywords- Wireless Drone Networking; Massive MIMO; Deep Q-Learning; Reinforcement Learning; Nonconvex Optimization.

I. INTRODUCTION

With recent technological innovations in telecommunications and the exponential increase in wireless data traffic from video streaming, Internet of Things (IoT), augmented reality, and surveillance, Unmanned Aerial Vehicles (UAVs) are being considered for use as Mobile Base stations (BSs) with massive Multiple-Input Multiple-Output (MIMO) networking capabilities [1], [2]. By enabling UAVs with Massive MIMO, different applications can be achieved, such as spectrum sharing through beamforming, unlicensed spectrum sharing with redeployable aerial drone base stations, secure wireless networking in congested environments through directional communication, and edge computing. Having UAVs act as massive MIMO mobile BSs is done to provide spectral efficient wireless networking for ground users. UAV BSs maximize spectral efficiency by changing location, controlling ground user transmit power, and assigning pilot sequence to users that maximizes the sum spectral efficiency of all users. It is worth noting that there are many factors that go into the operation of a drone, and they are impacted by many factors, such as weight, battery, energy consumption, and flight trajectory [4]. There have been many new approaches to extend the operation of a drone, such as battery swapping, using hybrid fuel cells and batteries, and solar cells for an extra power source [5]. These advancements are not the focus of this article which lets this article focus on the networking between ground users and the UAV BSs.

Massive MIMO adds the ability for wireless links to achieve higher throughput without the need of adding more BSs or increasing bandwidth. At the same time, UAVs have the advantage of being able to change location to follow demand. To maximize throughput, a massive MIMO UAV needs to find the best location and best pilot sequences to assign to users that minimizes interference between users. When multiple UAV BSs are used, each UAV BS has the added constraint of minimizing their connected users' interference with users connected to other BSs. It is reasonable to mount massive MIMO on an UAV BS because massive MIMO can reach smaller form factors if each ground node has distinct spatial channel characteristics [3]. For example, in a 2 GHz frequency band, for 100 dual-polarized antennas, the antenna array only requires 0.75 x 0.75 meters of space.

Controlling ground users' transmit power and pilot sequence allocation and the UAV BSs' movement is important in order to maximize ground users' sum spectral efficiency. However, this is a difficult problem because ground users can only be connected to one drone and can only be assigned one pilot sequence. This makes it a Mixed Integer Nonlinear Programming (MINLP) problem, which is generally NP hard. This means there are no known globally optimal solutions with polynomial computational complexity.

In recent literature, there are many solutions that do not use reinforcement learning. [7], [8] look to minimize user outage, [9] maximizes user spectral efficiency, [10], [11] maximize throughput, and [12] maximizes the Received Signal Strength Indicator (RSSI). There are also solutions that do use reinforcement learning. [13], [14] minimize user outage, [15], [16] maximize throughput, and [17] maximizes the RSSI. The main goal of these approaches is to maximize user experience. Some of these solutions, [7], [8] and [12]

 ISBN: 978-1-61208-940-9

focus on assisting a central terrestrial BS, [9] - [11], [13] - [15] focus on using UAVs, Long-Term Evolution (LTE) small cells, and access points as the BSs, and [17] only focuses on efficiently allocating spectrum for one cognitive radio network.

While many of these solutions are used in real systems, there are some drawbacks. First, LTE small cells, access points, and hotspots are not able to achieve the same level of spectral efficiency that massive MIMO can reach. Second, using terrestrial BSs is too expensive if there is only a short term spike in user demand. Finally, the last drawback is that solutions only meant to allocate spectrum to one radio network might not perform well if allocating resources to multiple radio networks at the same time.

To deal with these drawbacks, UAVs are used as mobile BSs mounted with massive MIMO antenna arrays. To maximize spectral efficiency, the locations of all the UAVs, the pilot signals assigned to the ground users and their transmit power are all maximized using Deep Q-Learning. Since this is a MINLP problem, the problem is relaxed by allowing the UAVs to assign multiple pilot sequences to ground users and users are assigned to UAVs that lead to the highest SISO SINR. These two steps remove the binary constraints in pilot assignment and user association. Results so far show that with two UAVs and two users, the UAVs are able to provide 95% of the optimum sum spectral efficiency.

The rest of the paper is organized as follows. Related work and background is reviewed in Section II. Problem formulation and implementation details are described in Section III. Sections IV goes over simulation results and includes a discussion of future work. Lastly, conclusions are drawn in Section V.

II. Related Work

A. Unmanned Aerial Vehicle Base Stations

An unmanned aerial vehicle network consists of an area with one or more UAVs mounted with BSs and multiple users or user equipment that need to connect to the network. The UAVs can change their location to provide better coverage, throughput, or spectral efficiency for users. This system could be paired with already existing ground BSs or be part of a public safety network after the occurrence of a natural disaster.

To see the advantages mobile BSs have over traditional static BSs, [10] checks the practical limits of UAVs in a cellular network and comes up with a mobility control algorithm to maximize the spectral efficiency at different UAV speeds. There are different scenarios where mobile UAV BS positioning is important. When using UAVs as mobile hotspots, [9] positions UAVs to minimize the distance between UAVs and users to boost packet throughput and the average packet throughput. UAV BSs can also be used for IoT networking where [12] used a drone mounted with a Raspberry Pi to create an IoT mesh to maximize the Received Signal Strength Indicator (RSSI). Drones can also be paired with traditional terrestrial BSs to further improve coverage. The authors in [8] use mobile BSs to provide downlink connectivity to ground users when their demand cannot be satisfied by the terrestrial station alone.

B. Massive MIMO

The main characteristic of massive MIMO is that it contains a massive number of antennas, generally 60 or more. It can take on several forms, but the main one used in this research is centralized massive MIMO where all the antennas are packed together in a single BS. To ensure that users receive their data stream with minimal interference, most forms of massive MIMO take advantage of spatial-division multiplexing. This form of multiplexing sends data streams at the same time and frequency [6]. What stops different users from receiving each other's signal is constructive and deconstructive interference with the antennas. The large number of antennas allows the BS to direct a signal at a specific user so only they will receive the signal. Massive MIMO BSs estimate channels using a pilot sequence that was sent by the user. One work that applies massive MIMO to mobile BSs is [11]. In this work, the authors applied a distributed algorithm price-based solution to UAVs to maximize the sum rate of all users and they did it by using convex relaxation to break the algorithm into three steps of access association, joint pilot assignment and power control and movement control. In this work, a massive MIMO array is attached to each drone and each drone controls a certain subset of users to control their pilot assignments and transmit powers with the goal of maximizing the sum of all the users' spectral efficiencies. [17] uses MIMO instead of massive MIMO and the only difference is the use of 2 antennas instead of 60 or more antennas.

C. Q-Learning

Q-Learning is an off-policy reinforcement algorithm that seeks to optimize the expected return based on Markov decision processes. The agent in Q-Learning starts by being in a certain state in the environment usually called the starting state and takes actions that will return a reward and transition the agent into a new state. In Q-Learning, an agent updates the q-value in its q-table where the rows are all the possible states, and the columns are all the possible actions. When an agent takes an action from an environment, it will update its q-value for the state it is in and the action that it took. All the values in the q-table are set to zero at the beginning and the table is updated using the following formula:

$$q_*^{new}(s,a) = (1 - \alpha)q(s,a) + \alpha\left(R_{t+1} + \gamma max\, q\left(s^{'}, a^{'}\right)\right) \quad (1)$$

where α is the learning rate and γ is the discount rate.

The way the agent finds the best policy is by exploration and exploitation. When the agent is exploring the

environment, it will select an action at random and when the agent is exploiting its environment, it will select the action with the highest q-value. At the beginning of training, the agent will explore its environment and will start to exploit it more and more as the episode number increases. This is an important step because the agent does not know anything about its environment in the beginning and once it starts to learn more, it will want to exploit it to find the best possible reward. One way this is achieved is by using the epsilon greedy strategy. This strategy sets a variable $\varepsilon = 1$, which decays exponentially and is updated using:

$$\alpha = \alpha_{end} + \left(\alpha_{start} - \alpha_{end}\right)\left(e^{-n_{step}\lambda}\right) \quad (2)$$

where α is the exploration rate, α_{end} is the ending exploration rate, α_{start} is the starting exploration rate, and λ is the exploration decay rate.

Reinforcement learning has been used in other UAV BS applications, such as determining the optimal positions for mobile BSs and [13] applies this to an emergency communication network. It may also be important to control the transmit power as well to minimize the interference between BSs, which is what [14] did by using a reinforcement learning algorithm to control transmit power and BS positioning to minimize user outage probabilities. Lastly, [17] uses reinforcement learning to analyze radio frequency channels to learn from past occupancy and conditions of the channels.

D. Deep Q-Learning

Unlike Q-Learning, which uses a q-table to keep track of its q-values, Deep Q-Learning uses a neural network where the input is the current state, and the output is the q-value for each action. The neural network can have any number of hidden layers and the main purpose of the neural network is to approximate the values of a q-table. For each action, state, reward and next state the agent experiences, this tuple is called experience replay. The experience replay includes the state of the environment, the action taken from that state, the reward given to the agent as a result of the previous state-action pair and the next state of the environment. Each experience replay is stored in an array called the replay memory up to some amount of experiences *N*. When the number of experiences gets larger than *N*, the first experience gets replaced with the most current experience. The replay memory gets sampled randomly to train the network, which breaks the correlation between successive samples to make the learning more efficient. When training the neural network, the loss is calculated by subtracting the q-value of the given state-action pair from the optimal q-value of the same state-action pair. The optimal q-value is calculated by passing the next state into the neural network to find the max q-value among all actions that can be taken in that state. The optimal q-values are not known at the beginning of training, so they are estimated using the neural network and get updated when the weights of the neural network get updated. To avoid instability, the optimal q-values are updated and calculated using a separate network called the target network. The target network is a copy of the original policy network, but only gets updated *x* timesteps.

Deep reinforcement learning has been used in [15] to find the best way to allocate resources in a distributed environment when the channel state information is not known. [16] uses this tool to also control transmit powers to mitigate interference, which helps maximize throughput.

E. Discussion of Related Work

Only one previous work [11] applies massive MIMO to UAV BSs, but they use a pricing algorithm to get within 90% the global optimum. None of the previous works have applied reinforcement learning or deep reinforcement learning to control the user association, pilot assignment and UAV movements to maximize the sum rate of the users. While [17] did use reinforcement learning, it does not take into account multiple BSs with multiple users. Therefore, this paper proposes a solution using Deep Q-learning, where each UAV is an agent that will try to maximize its reward based on the sum spectral efficiencies of all users. This approach is similar to [11] in that it is a distributed system and each UAV does not need to collect the full statistical Channel State Information (CSI), the locations of the other UAVs, the noise power and other network parameters. Also, since this is a distributed system, it does not encounter failure from a single point like centralized systems. Also, distributed systems scale better and have lower latencies than centralized systems.

III. Implementation Details

The deep reinforcement learning algorithm was written and simulated in python. Python was used so PyTorch and Gym can be used. Pytorch is used to create the target and policy neural networks in the Deep Q-Learning model and Gym is used to create the reinforcement learning environment. The Gym library is a toolkit for developing reinforcement learning algorithms and has many good environments to choose from when building the massive MIMO UAV agents and environment. A Gym environment class has four main functions that are needed to manage the agents in the environment. The *init* method is used to initialize all variables and constants, the *step* method executes one time step in the environment, the *reset* method resets the environment to the initial state, and the *render* method prints the current state of the environment to the screen.

A. Massive MIMO UAV BSs

All UAV BSs can move freely in a 500 $m^2 \times 500$ m^2 area, but their heights are kept at a constant 100 m above the ground. The number of UAV BSs is set to {1, 2} and the

 ISBN: 978-1-61208-940-9

number of users is set to {1, 2, 3, 4}. The users do not move through the environment and are randomly placed in the grid at the start of the simulation. The number of antennas in each UAV BS is set to {10, 20, 30, 40, 50, 100}. The range of transmit powers the UAV BSs can set for the ground users is in between [5, 500] mW and the total number of power levels is set to {3, 4, 5}. The path factor is set to 2, the average noise power is set to 10^{-8}, the length of each pilot sequence is set to 10 symbols and the total number of available pilot sequences is set to {2, 3, 4, 5, 6}.

B. Deep Reinforcement Learning Implementation

Since this problem is a MINLP problem, to find the best sum spectral efficiency, the pilot assignment is relaxed to allow a UAV to assign multiple pilot signals to a single user and the problem was broken down into two subproblems. In the first subproblem, the users are connected to the UAV that has the best SISO SNR. The next subproblem is the deep Q-Learning algorithm, which controls the UAVs' movements and the power allocation for each pilot sequence. For the implementation details of the reinforcement learning algorithm:

Agent: UAV BSs

State: Includes the position on the agents, the users the agents are connected to, the number of pilot sequences, and the number of power levels. To limit the total number of states, the agents operate in a square grid, and the total number of transmit powers are divided into *n* power levels. The level of different transmit powers is divided evenly between the max and min user transmit powers.

Action: Includes moving up, down, left, right, increasing user transmit power assigned to a pilot sequence, and decreasing user transmit power assigned to a pilot sequence.

Reward: Based on the sum of the spectral efficiency for the users connected to UAVs. The spectral efficiency is calculated by dividing the capacity from (3) by *B* Hz.

$$C_g = B\log_2(1+\gamma_g) \tag{3}$$

The capacity is then calculated with (3), which includes channel-estimation error, the type of linear spatial multiplexing/demultiplexing, power control, and noncoherent inter-cell interference (4) [11].

Lastly, to get the reward from the spectral efficiency equation, it is multiplied by the power level the UAV chooses for a pilot sequence divided by the max power level. This was done to encourage the agents to choose only one pilot sequence.

$$R_g = C_g\left(\frac{p_{gw}}{p_{max}}\right) \tag{5}$$

The neural network in this article has one hidden layer with 500 nodes. The input to the neural network is one-hot encoded, which makes the input have as many nodes as the number of states, and the number of nodes at the output is determined by the total possible actions that the agent can take. After each action, the agent stores the action-reward pair in memory. After the agent takes a certain number of actions, it updates the target network based on a random batch of action-reward pairs from memory using the Adam optimizer. The purpose of the target network is to help reduce instability when both the training q-values and the optimum q-values are both being updated throughout the simulation. The loss is calculated using mean square error between the q-value calculated in the training network and the q-value calculated in the target network. The agent repeats the above process until the training is over.

C. Relaxed Centralized Solution

The simulations were compared to a relaxed centralized solution to see how close they were to the max possible reward. This problem is solved by maximizing (6), which finds the x,y, and z positions of the UAVs. This returns the largest sum spectral efficiencies across all users. This relaxed centralized solution is not the main focus of the paper because the UAVs do not know the locations of the users. Also, a centralized system is more prone to failure if a single component fails while a decentralized system is more resilient to failure.

$$Maximize: \sum_{g\in G} C_g(x,y,z) \tag{6}$$

In (6), all interferences are ignored and all users are assumed to be using separate pilot sequences. Also, $C_g(x,y,z) = Blog_2(1+\gamma_g(x,y,z))$, and since $\gamma_g(x,y,z) >> 1$, $C_g(x,y,z)$ can be approximated.

$$\begin{aligned} &\approx Blog_2(\gamma_g(x,y,z)) \\ &\leq Blog_2(M\tau\rho_g p_0 \zeta_{gg}^2 H_{gg}^2(x,y,z)) \\ &= Blog_2\left(\frac{M\tau\rho_g p_0\zeta_{gg}^2}{d_{gg}^x(x,y,z)}\right) \\ &= Blog_2(M\tau\rho_g p_0\zeta_{gg}^2) - xBlog_2(d_{gg}(x,y,z)) \end{aligned} \tag{7}$$

When looking at (7), the first term is constant since none of the variables are a function of x, y, or z. Another way to write this maximization problem is to write it as a minimization of the second term.

$$Minimize: \sum_{g\in G} -log_2(d_{gg}(x,y,z)) \tag{8}$$

The exact solution to this problem is approximated by calculating a 20000 x 20000 grid of all the possible x and y positions of the UAV. This finds a solution very close to the

$$\gamma_{gw}(\tilde{\mathbf{p}}) = \frac{(M-|\mathcal{G}_{a(g)}|)\tau\rho_g\beta_{gg}^2 p_{gw}}{(1+\tau\mathcal{E})(1+\sum_{g'\in\mathcal{G}}\sum_{w'\in\mathcal{W}}\mu_{g'g}(\boldsymbol{\mu})p_{g'w'}) + (M-|\mathcal{G}_{a(g)}|)\sum_{g'\in\mathcal{I}_g\setminus g}\sum_{w'\in\mathcal{W}}\rho_{g'}\beta_{g'g}^2 p_{g'w'}} \tag{4}$$

 ISBN: 978-1-61208-940-9

optimum solution because there is only 25 mm separation for each square in the grid.

IV. Simulation Results

For the experiment, the data that was recorded was the cumulative reward and it is the total reward the agent received during each episode. This benchmark is used because it shows the agent gradually choosing better actions that return a higher reward more often in a single episode. The simulation parameters are stored in Table 2 and Table 1 describes the meaning of all the parameters in Table 2. In Figure 1, there are 2 UAVs, 2 users, and 2 pilot sequences to choose from. In the beginning of the graph, there are four different colors stacked on top of each other. The blue line on the bottom represents one or both of the UAVs assigning the first pilot sequence to the first user. The orange line above the blue line represents one or both of the UAVs assigning the second pilot sequence to the first user. The green line above the orange line represents one or both of the UAVs assigning the first pilot sequence to the second user. Lastly, the red line above the green line represents one or both of the UAVs assigning the second pilot sequence to the second user. Since this is the cumulative reward, for each episode there are 2000 iterations where the UAV can change the pilot sequence assigned to a user. In the beginning of the simulation, the UAVs do not know which pilot sequence to assign to the users, so it picks randomly. As the simulation progresses, the UAVs learn that if a user is assigned a pilot sequence by one or the other UAV, they or the other UAV should assign the other pilot sequence to the other user. This can be seen near the end of the simulation where there is only a green line above an orange line, or a red line above a blue line (Figure 1). This means that either the first user was assigned the first pilot sequence and the second user was assigned the second pilot sequence or the first user was assigned the second pilot sequence and the second user was assigned the first pilot sequence. There is also an average in the graph which is represented by the black line on top. The average takes into account the past 50 episodes and can be seen increasing until it levels out at about the 20000th episode.

The optimum cumulative reward is found by finding the optimum location for the UAVs using the relaxed centralized solution and then multiplying it by the number of iterations for one episode in the simulation. The optimum cumulative reward is shown in the graph by a yellow horizontal line on the top of the graph and it is also labeled for clarity. The simulation in Figure 1 can be seen to come very close to the optimum cumulative reward found using the relaxed centralized solution. In this simulation, the drone network average was able to come within 95% of the cumulative relaxed centralized solution.

A. Discussion

The UAVs were able to find the optimum solution even though the reward function was not fully convex. The function that the reinforcement learning algorithm is optimizing over greatly determines how quickly a solution can be found. Further investigation will look into adding more UAVs, more users, and simulations where there are more users than pilot sequences. Further investigation will also look into different ways the policy network can be updated to improve convergence when more agents and users are added. This will be done to test results with more realistic scenarios where there are more users and not enough pilot sequences for each user.

V. Conclusion

In this article, deep reinforcement learning was implemented and evaluated to optimize the sum spectral efficiency of ground users. To verify this, a simulation was built in Python using OpenAI's Gym library to create a custom agent, reward function, and environment. Details are included on how the deep reinforcement learning algorithm is set up and the convex relation techniques used to help the aerial drone find the maximum spectral efficiency. Lastly, the results are compared to the log of the euclidean distance to see how the simulations compare to the near optimum sum spectral efficiency.

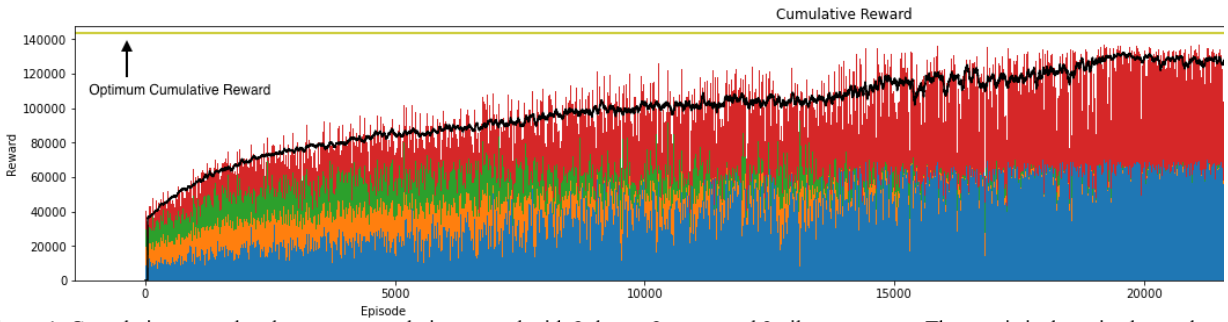


Figure 1: Cumulative reward and average cumulative reward with 2 drone, 2 users, and 2 pilot sequences. The x-axis is the episode number and the y-axis in the cumulative reward.

 ISBN: 978-1-61208-940-9

References

[1] R. Iyer and E. Ozer, "Visual IoT: Architectural Challenges and Opportunities; Toward a Self-Learning and Energy-Neutral IoT," in IEEE Micro, vol. 36, no. 6, pp. 45-49, Nov.-Dec. 2016, doi: 10.1109/MM.2016.96.

[2] V. J. Hodge, S. O'Keefe, M. Weeks and A. Moulds, "Wireless Sensor Networks for Condition Monitoring in the Railway Industry: A Survey," in IEEE Transactions on Intelligent Transportation Systems, vol. 16, no. 3, pp. 1088-1106, June 2015, doi: 10.1109/TITS.2014.2366512.

[3] E. Björnson, E. G. Larsson and T. L. Marzetta, "Massive MIMO: ten myths and one critical question," in IEEE Communications Magazine, vol. 54, no. 2, pp. 114-123, February 2016, doi: 10.1109/MCOM.2016.7402270.

[4] M. Mozaffari, W. Saad, M. Bennis, Y. -H. Nam and M. Debbah, "A Tutorial on UAVs for Wireless Networks: Applications, Challenges, and Open Problems," in IEEE Communications Surveys & Tutorials, vol. 21, no. 3, pp. 2334-2360, thirdquarter 2019, doi: 10.1109/COMST.2019.2902862.

[5] M. N. Boukoberine, Z. Zhou and M. Benbouzid, "Power Supply Architectures for Drones - A Review," IECON 2019 - 45th Annual Conference of the IEEE Industrial Electronics Society, 2019, pp. 5826-5831, doi: 10.1109/IECON.2019.8927702.

[6] T. L. Marzetta, "Massive MIMO: An Introduction," in Bell Labs Technical Journal, vol. 20, pp. 11-22, 2015, doi: 10.15325/BLTJ.2015.2407793.

[7] A. Vallamreddy and P. Indumathi, "Outage performance analysis of MIMO cognitive radio network users in fading environment," 2014 2nd International Conference on Devices, Circuits and Systems (ICDCS), 2014, pp. 1-4, doi: 10.1109/ICDCSyst.2014.6926207.

[8] W. Shi et al., "Multiple Drone-Cell Deployment Analyses and Optimization in Drone Assisted Radio Access Networks," in IEEE Access, vol. 6, pp. 12518-12529, 2018, doi: 10.1109/ACCESS.2018.2803788.

[9] A. Fotouhi, M. Ding and M. Hassan, "Flying Drone Base Stations for Macro Hotspots," in IEEE Access, vol. 6, pp. 19530-19539, 2018, doi: 10.1109/ACCESS.2018.2817799.

[10] A. Fotouhi, "Towards intelligent flying base stations in future wireless network," 2017 IEEE 18th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2017, pp. 1-3, doi: 10.1109/WoWMoM.2017.7974302.

[11] Z. Guan, N. Cen, T. Melodia and S. M. Pudlewski, "Distributed Joint Power, Association and Flight Control for Massive-MIMO Self-Organizing Flying Drones," in IEEE/ACM Transactions on Networking, vol. 28, no. 4, pp. 1491-1505, Aug. 2020, doi: 10.1109/TNET.2020.2985972.

[12] M. F. Ahmed, E. M. Naveed, P. Nar and S. K. Jindal, "Design of an Autonomous drone for IoT deployment analysis," 2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN), 2019, pp. 1-4, doi: 10.1109/ViTECoN.2019.8899609.

[13] P.V. Klaine, J.P.B Nadas, R.D. Souza, M.A. Imran, “Distributed Drone Base Station Positioning for Emergency Cellular Networks Using Reinforcement Learning”. *Cogn Comput* 10, 790–804, 2018. pp. 790-804, doi: 10.1007/s12559-018-9559-8

[14] R. de Paula Parisotto, P. V. Klaine, J. P. B. Nadas, R. D. Souza, G. Brante and M. A. Imran, "Drone Base Station Positioning and Power Allocation using Reinforcement Learning," 2019 16th International Symposium on Wireless Communication Systems (ISWCS), 2019, pp. 213-217, doi: 10.1109/ISWCS.2019.8877247

[15] J. Jang, H. J. Yang and S. Kim, "Learning-Based Distributed Resource Allocation in Asynchronous Multicell Networks," 2018 International Conference on Information and Communication Technology Convergence (ICTC), 2018, pp. 910-913, doi: 10.1109/ICTC.2018.8539654.

[16] G. Zhao, Y. Li, C. Xu, Z. Han, Y. Xing and S. Yu, "Joint Power Control and Channel Allocation for Interference Mitigation Based on Reinforcement Learning," in IEEE Access, vol. 7, pp. 177254-177265, 2019, doi: 10.1109/ACCESS.2019.2937438.

[17] L. Bondan, M. A. Marotta, L. R. Faganello, J. Rochol and L. Z. Granville, "ChiMaS: A spectrum sensing-based channels classification system for cognitive radio networks," 2016 IEEE Wireless Communications and Networking Conference, 2016, pp. 1-7, doi: 10.1109/WCNC.2016.7564911.

TABLE I.

Variables	***Meaning***
Batch Size	Number of samples used to train the neural network
Gamma	Discount rate
Starting Epsilon	Staring value of epsilon decay
Ending Epsilon	Ending value of epsilon decay
Epsilon Decay	Value to decrease epsilon by
Target Update	Updates target network every *N* Episodes
Memory Size	Max number of experiences in replay memory
Learning Rate	Learning rate of neural network
Number of Episodes	Max number of episodes in simulation
Max Steps Per Episode	Max number of steps before episode ends

TABLE II.

Values of Variables used in Simulation	
Variables	***Values***
Batch Size	10
Gamma	0.99
Starting Epsilon	0.9
Ending Epsilon	0.001
Epsilon Decay	$4*10^{-8}$
Target Update	60
Memory Size	1000
Learning Rate	0.0001
Number of Episodes	50000
Max Steps Per Episode	2000

 ISBN: 978-1-61208-940-9

Optimization of the Virtual Network Function Reconfiguration Plan in 5G Network Slicing

Hanane Biallach
Orange Innovation - Heudiasyc
Châtillon, France
hanane.biallach@hds.utc.fr

Mustapha Bouhtou
Orange Innovation
Châtillon, France
mustapha.bouhtou@orange.com

Dritan Nace
Heudiasyc
Compiègne, France
dritan.nace@hds.utc.fr

***Abstract*—It is widely acknowledged that the forthcoming 5G architecture will be essentially based on network slicing, which enables to provide a flexible approach to realize the 5G vision. Thanks to the emerging Network Function Virtualization (NFV) concept, the network functions are decoupled from dedicated hardware devices and realized in the form of software. One of the main technical challenges is the reconfiguration of Virtualized Network Functions (VNFs). In this work, we have modeled the problem through integer linear programming, which is compared to the topological sorting algorithm. Experimental results show the benefits of our model and demonstrate its ability to achieve reconfiguration plans with minimum migration duration and service interruption.**

***Index Terms*—5G Network slicing, VNF reconfiguration, VNF migration, Integer linear programming.**

I. Introduction

5G comes with new needs that may vary considerably depending on the use case. This may involve very low latency, very high throughput or a massive amount of connections, all with a high Quality of Service (QoS) requirement. One of the visions of 5G is network slicing [1] [2], which is a concept for running multiple logical customized networks on a shared common infrastructure complying with agreed Service Level Agreements (SLAs) for different vertical industry customers and requested functionalities. In the 3rd Generation Partnership Project (3GPP) [3], three standardized slice types are currently defined: Massive Machine Type Communications (mMTC), Enhanced Mobile Broadband (eMBB) and Ultra-reliable and Low Latency Communications (uRLLC). The goal of the first slice is to respond to the exponential increase of connected objects. It allows as many objects as possible to connect to the network. The second slice is put forward for data-intensive applications and requires high data rates of several giga bits per seconds with moderate latency. The last slice is intended for applications requiring extremely high reactivity and high message transmission guarantee.

The network slicing is essentially based on Network Function Virtualization (NFV), which decouples network functions from proprietary hardware platforms and implements them into software to make the provision of these functions more efficient and flexible. Typically, a slice service is represented by a Service Function Chain (SFC) that is a set of Virtualized Network Functions (VNFs) that are executed according to a given order (see Fig. 1). A VNF may be made up of one or more VNF Components (VNFC), which implement a subset of the VNF's functionality. The VNF (hence VNFC) can be implemented in either VM (software application behaving as a separate computer system, exhibited by VMWare [4], Xen, KVM, etc.) or container (lightweight, standalone, executable package of software such as Docker [5] and Kubernetes [6]).

Fig. 1. Example of SFC for internet connection. **UE**: *User Equipment*, **RAN**: *Radio Access Network*, **DU**: *Distributed Unit*, **CU**: *Centralized Unit*, **UPF**: *User Plane Function*, **DN**: *Data Network*

The NFV concept allows dynamic changes that can occur in the network due to its flexibility and agility. Optimally reconfiguring Virtualized Network Functions (VNFs) remains important, since the dynamic slicing changes can impact the performance of one or more slices, which can lead to broken SLA requirements, and therefore penalties. In real life, slice demands arrive dynamically and the network state changes continuously. Due to the stochastic nature of users behavior, the traffic variation cannot be perfectly handled in advance. Consequently, all the placement decisions that might be optimal, at a certain point in time, may become sub-optimal due to the new demands of VNFs deployments. This may end up with an inefficient resource usage, hence the need for network reconfigurations time to time in response to changing network conditions is an indispensable component to maintain high QoS and profit.

Today, most of existing work on network slicing is mainly focused on the flow routing together with Service Function Chain (SFC) deployment, and less from VNF reconfiguration point of view. In [7], an Integer Linear Programming (ILP) model was formulated to solve the problem of mapping and reconfiguring the SFC for dynamic situations, with the objective to minimize the service provider's operational overhead. Yet, they did not consider the migration cost. In [8], the authors consider the problem of both rerouting traffic flows and improving the mapping of network functions onto nodes in the presence of dynamic traffic, with the objective of bringing the network back to a close to optimal operating state, in terms of resource usage. However, they do not take into consider-

 ISBN: 978-1-61208-940-9

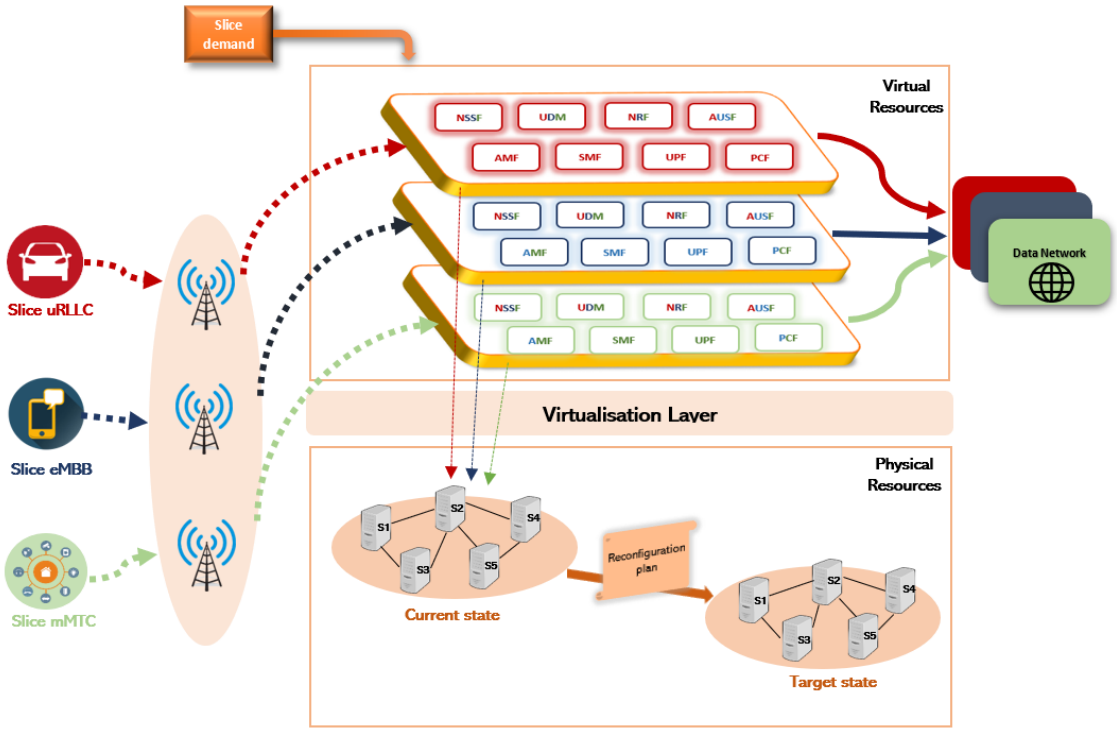


Fig. 2. Presentation of VNFs reconfiguration problem. **NSSF**: *Network Slice Selection Function*, **UDM**: *Unified Data Management*, **NRF**: *Network Repository Function*, **AUSF**: *Authentication Server Function*, **AMF**: *Access and Mobility Management Function*, **SMF**: *Session Management Function*, **PCF**: *Policy Control Function*

ation the service interruption. Eramo et al. [9] [10] proposed algorithms to handle the VNF placement, SFC routing, and VNF migration in response to the change of user workload. A migration policy was proposed to establish when and where migrations of VNF have to be accomplished so as to minimize a total cost characterized by the sum of the energy cost and the reconfiguration cost occurring when the VNFs are moved from the initial location. Nevertheless, the VNF interruption has not been taken into account.

Our contributions in this paper are as follows:

- We propose an ILP model for the VNF reconfiguration problem in the context of network slicing in 5G networks. The model takes into account two types of migrations (hot and cold migrations). Our algorithm generates a reconfiguration plan to pass rapidly and efficiently from an initial state where the placed VNFs are not optimally allocated to a new state, computed beforehand, respecting the resource capacity with a minimal interruption, migration and SLA costs.
- We provide abundant numerical data illustrating the findings of our work.

The organization of this paper is as follows: Section II presents the system model. Section III introduces the problem statement and formulation. The experiments and evaluation results are presented in Section IV. Some concluding remarks and perspectives are presented in Section V.

II. SYSTEM MODEL

A. Problem definition

The problematic we are interested in is the reconfiguration of 5G network slices. To define it simply, a reconfiguration is a reallocation of the NFV to adapt the utilization of network resources to the occurred changes. Fig. 2 shows an example of VNFs reconfiguration problem. Three service slices are presented: self driving car (slice uRLLC), live streaming (slice eMBB) and smart home (slice mMTC). Each slice is a set of virtualized network functions (VNFs), and each VNF is deployed in one Virtual Machine (VM) with different capacities (CPU, RAM). The slices (virtual resources) are deployed in the VNF infrastructure (physical resources) presented by five servers in ”current state”, which represents the initial state of our problem. At time t, a new demand for slice deployment is presented. In this case, the current VNF placements become sub-optimal and inefficient. The VNFs placement should be reconfigured by migrating VNFs to another optimal state (target state). In our problematic, the current and target states are known beforehand. Our objective is to reconfigure all the realized migrations to attain the target state (new placement of VNFs) and generate a reconfiguration plan that allows to pass rapidly and optimally from the current state to the target state while respecting resource capacities, minimizing the service interruption and migration duration.

The VNF migrations are performed by two types of migrations. In a **hot (live) migration**, the running VNFs are moved

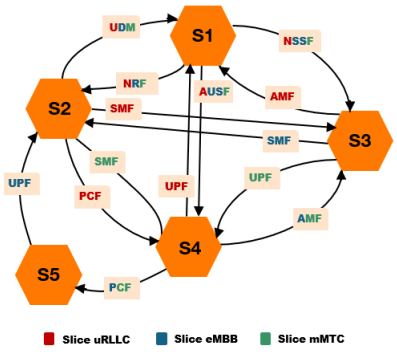


Fig. 3. The graph representation of VNFs migrations

between the source and target servers without disconnecting the service or application. The hot migration ensures minimal downtime for the VNF. In a **cold (non-live) migration**, the VNFs are moved between servers by powering off the VNF on the source server, moving it to the target server then powering it back up on the target server. The cold migration ensures an important downtime that should be minimised.

Our objectives are to minimize the total migration duration, so as the VNF migration be performed as quickly as possible and to minimize the service interruption, which is important especially for some use cases such as smart grids, intelligent transport systems and remote surgery. These services require an ultra-high network reliability of more than 99.999% and very low latency (of 1 millisecond) for packet transmission. Minimizing the service interruption ensures the slice availability and the avoidance of SLA violations. Moreover, the VNF interruption degrades the service not only for one slice, but for many other slices too as they can be shared between several of them.

B. Problem modeling

The VNF reconfiguration problem is a NP-Hard optimisation problem. [11] demonstrates the NP-hardness of the reconfiguration problem in the context of distributed systems. In this paper, we model the network as a connected directed graph $G = (V, E)$. The nodes $v \in V$ stand for the servers and the links $(v, v') \in E$ connecting the nodes represent the VNF migration from node v to node v' (see Fig. 3). The VNF can be shared between different slices (ex in Fig. 2: *NSSF*, *UDM*, *NRF* and *AUSF* are shared VNFs between all slices) or dedicated to one slice (ex in Fig. 2: *SMF*). In this paper, we assume that each VNF is implemented in one VM and there is a low communication delay between them. Thus, the flow routing is not taken into consideration.

Our objective is to propose an optimisation algorithm that takes as input the current and target states and generates as output the reconfiguration plan while minimizing the total migration duration and the service interruption. There is a property that already has been demonstrated in [11], and consists of finding the reconfiguration plan in polynomial time without service interruption using Topological Sorting (TS) algorithm, in the case where the network topology is an acyclic graph.

The topological sorting for Directed Acyclic Graph (DAG) [12] [13] is a linear ordering of nodes such that for every directed arc (v, v'), node v comes before v' in the ordering. The TS algorithm is not possible if the graph is not a DAG (for the case of cyclic graph). For this reason, we propose an exact model that can be applied to different types of graphs (acyclic and cyclic) and where the solution can be optimised in terms of service interruption and total migration duration.

III. Problem Statement and Formulation

Linear programming constitutes the basis of the solution method developed in this work. In this section, we present the VNF reconfiguration problem statement and its formulation as an Integer Linear Programming (ILP).

A. VNF reconfiguration problem: Problem statement

The VNF reconfiguration problem is presented as follows with notation given in Table I for easy reference:

- ***Given***: The placement of VNFs in the current and target states.
- ***Find***: in which stage k the VNF_i should be migrated, which type of migration to use (cold or live migration) while respecting the resource constraints.
  The total number of stages N required for all VNFs to be migrated can be equal to N_v the number of VNFs in worst cases, where each VNF migrates separately in one stage. Or less than that, in case where we have parallel migrations.
- ***Subject to***: the VNF occupied CPU capacity cap_i^{cpu}, the VNF occupied RAM capacity cap_i^{ram}, the VNF interruption duration δ_i and the VNF migration duration T.
- ***Objective***: minimizing the VNF migration and interruption duration.

B. Problem formulation

To formulate the integer linear programming model, we introduce the decision variables, the constraints to be satisfied, and the objective function.

1) Decision variables: We have the following decision variables to model VNF migrations between servers (Knapsacks):

$$x_{ik} = \begin{cases} 1, & \text{if the } VNF_i \text{ is migrated in stage } k \text{ ;} \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

$$y_{ik} = \begin{cases} 1, & \text{if the } VNF_i \text{ is interrupted in stage } k \text{ ;} \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

2) Problem constraints:

- *Integrity constraint for migration*

Equation (3) insures that VNF_i can only be migrated once to the destination server.

TABLE I
TABLE OF NOTATIONS

Notation	Description
N	number of stages
N_v	number of VNFs
N_s	number of servers
k	order / stage of the reconfiguration
x_{ik}	a binary variable indicating that VNF_i is migrated in order k
y_{ik}	a binary variable indicating the stage where VNF_i is interrupted in source host
$O_{(s)}$	set of VNFs originating from server s
$D_{(s)}$	set of VNFs targeting server s
C_s^k	represents the residual CPU capacity of server s in stage k
R_s^k	represents the residual RAM capacity of server s in stage k
cap_i^{cpu}	represents the occupied CPU capacity of VNF_i
cap_i^{ram}	represents the occupied RAM capacity of VNF_i
δ_i	represents the interruption duration of VNF_i
T	represents the migration duration of a given VNF
β_i	represents the cost of service interruption, which is the SLA availability of each VNF_i
α	represents the migration cost of all VNFs

$$\sum_{k=1}^{N_v} x_{ik} = 1 \ ; \ \forall i \in \{1, .., N\} \tag{3}$$

- *Integrity constraint for interruption*

Equation (4) shows that VNF_i can only be interrupted once in the source server.

$$\sum_{k=1}^{N_v} y_{ik} = 1 \ ; \ \forall i \in \{1, .., N\} \tag{4}$$

- *Capacity constraint*

Equations (5) and (6) ensure that resource capacities of each server (for CPU and RAM, respectively) in each stage k, are not exceeded. VNFs that are interrupted free the resources of their origin server, while VNFs that are placed consume the resources of their destination server.

$$\forall s \in \{1, .., N_s\} \ ; \ \forall k \in \{1, .., N\} \ ;$$
$$C_s^k - \sum_{i \epsilon D(s)} x_{ik} cap_i^{cpu} + \sum_{i \epsilon O(s)} y_{ik} cap_i^{cpu} = C_s^{k+1} \tag{5}$$

$$\forall s \in \{1, .., N_s\} \ ; \ \forall k \in \{1, .., N\} \ ;$$
$$R_s^k - \sum_{i \epsilon D(s)} x_{ik} cap_i^{ram} + \sum_{i \epsilon O(s)} y_{ik} cap_i^{ram} = R_s^{k+1} \tag{6}$$

$$C_s^k \geq 0 \ ; \ \forall s \in \{1, .., N_s\} \ ; \ \forall k \in \{1, .., N\} \tag{7}$$

$$R_s^k \geq 0 \ ; \ \forall s \in \{1, .., N_s\} \ ; \ \forall k \in \{1, .., N\} \tag{8}$$

- *Interruption duration constraint*

Equation (9a) ensures that the VNF is migrated with cold migration during the whole process. It considers that interruption and migration could be performed in one same stage. Otherwise, equation (9b), which refers also to cold migration, gives more flexibility. The interruption and migration could be performed in one or more stages. In cold migration, the VNF is interrupted first in the source host then it is migrated to the destination host. The VNF interruption should be before VNF migration.

Equation (9c) ensures that the VNF is migrated with live migration during the whole process. In this case, the migration of VNF_i occurs at one stage before interruption. Here, we consider that VNF is interrupted in the source host after totally being migrated to the destination host.

Equation (9d) encompasses the cold and live migration.

$$\sum_{k=1}^{N_v} k y_{ik} = \sum_{k=1}^{N_v} k x_{ik} \ ; \ \forall i \in \{1, .., N\} \tag{9a}$$

$$\sum_{k=1}^{N_v} k y_{ik} \leq \sum_{k=1}^{N_v} k x_{ik} \ ; \ \forall i \in \{1, .., N\} \tag{9b}$$

$$\sum_{k=1}^{N_v} k y_{ik} = \sum_{k=1}^{N_v} k x_{ik} + 1 \ ; \ \forall i \in \{1, .., N\} \tag{9c}$$

$$\sum_{k=1}^{N_v} k y_{ik} \leq \sum_{k=1}^{N_v} k x_{ik} + 1 \ ; \ \forall i \in \{1, .., N\} \tag{9d}$$

The interruption time is considered as the number of stages between the VNF interruption and VNF migration. In live migration, the interruption time is negligible, therefore, we consider δ_i = 0. In cold migration, the VNF interruption is performed at least in one stage δ_i = 1. Equation (10) refers to the formulation of VNF interruption time.

$$\delta_i = (\sum_{k=1}^{N_v} k x_{ik} + 1) - (\sum_{k=1}^{N} k y_{ik}) \tag{10}$$

- *Migration duration constraint*

Equation (11) finds the maximum migration duration that should be minimized.

 ISBN: 978-1-61208-940-9

$$\sum_{k=1}^{N_v} kx_{ik} \leq T \; ; \; \forall i \in \{1, .., N\} \tag{11}$$

3) Objective function:

$$min(\sum_{i=1}^{N_v} \beta_i \delta_i + \alpha T) \tag{12}$$

The objective function consists of minimizing the VNFs interruption time, that represents the number of stages during which the VNF_i is interrupted, and minimizing the VNFs migration duration, that represents the number of stages during which the VNF_i is migrated. The weight β_i associated with δ_i represents the SLA availability for each VNF belonging to a given slice. The service availability of the slice is divided into three ranges: high availability (β_i = 100%), average availability ($\beta_i \geq$ 99%), and low availability ($\beta_i <$ 99%).

IV. Experimental results

A. Simulation Setup

1) Topology Dataset: The ILP model is solved using CPLEX Optimisation studio V12.8 integrated in python. Experiments were conducted on a machine with Core i7-6600U CPU and 16 Go of RAM. We use randomly generated topologies to evaluate our model for both acyclic and cyclic graphs. The graphs are randomly generated with different sizes (small and medium graphs) using the NetworkX, which is a well-known pyhton lib, and we are inspired from the code proposed by [14] [15]. The nodes of the graph represent the servers and the links represent the VNF migration. The node and link capacities are generated randomly, (1∼50) for CPU capacity and (10∼90) for RAM capacity.

2) VNF and slice Datasets: The type number of VNFs is randomly generated in range (20∼150). The slices are randomly generated by choosing the set of connected VNFs, taking into consideration the shared and dedicated VNFs. Each slice contains at least 5 VNFs. The datasets are presented in Table II and Table III.

TABLE II
DATASETS OF ACYCLIC GRAPHS

Instances	Servers	VNFs	Slices
DC-acy1	10	25	6
DC-acy2	20	35	11
DC-acy3	40	60	12
DC-acy4	50	120	24
DC-acy5	80	150	35

TABLE III
DATASETS OF CYCLIC GRAPHS

Instances	Servers	VNFs	Slices
DC-cy1	10	30	8
DC-cy2	20	45	12
DC-cy3	40	70	15
DC-cy4	50	120	25
DC-cy5	80	146	32

B. Evaluation Metrics

To show the performance of our model, we use the following evaluation metrics:

- **Scalability**: To evaluate the scalability of our model, we adopt two metrics. These metrics are the model execution time in seconds and the estimated gap to optimal in % after one hour.
- **Migration duration**: We evaluate the total migration duration for the entire process, as well as the number and the percentage of migrated VNFs per each step of the migration.
- **Interruption duration**: We evaluate the ratio of interrupted VNFs to the total VNFs as well as the interruption duration for each slice demand.
- **Migration and interruption costs**: We evaluate the interruption and migration cost by giving each VNF the corresponding SLA availability β_i and varying the weight α.

C. Simulation Result and Analysis

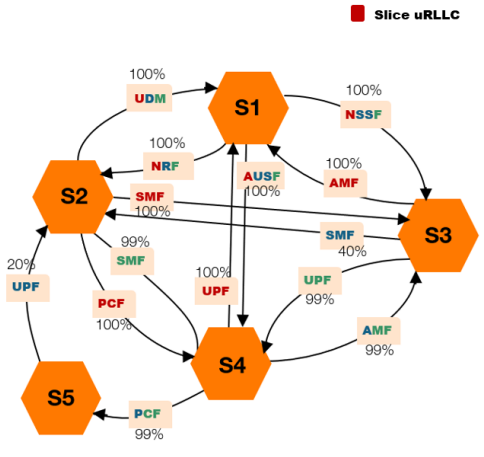
1) Evaluation according to the nature of slices: We evaluate the example presented in Fig. 2 with V = 5 and E = 14. As we mentioned earlier, each slice has its SLA availability that should be respected. In Fig. 4(a), we present the considered values of each service availability: high availability for slice uRLLC (β_i = 100%), average availability for slice mMTC (β_i = 99%) and low availability for slice eMBB ($\beta_i <$ 99%).

Fig. 5(a) shows the total migration and interruption duration for all slices according to different variations of α, where the number of α is varied between (1∼200). We can see that the total migration duration decreases and the total interruption duration increases with the rise of the α. From $\alpha = 100$, we can observe clearly that the total migration and interruption duration stagnates respectively in steps 3 and 7. This is because the ILP model finds the optimal solution that minimizes both migration and interruption duration.

To evaluate the interruption duration for each slice, we set the α to 100. Fig. 5(b) presents migration duration of VNFs for each slice. We can see that in the slice uRLLC there is no interruption as it demands a high availability, then for mMTC there is one VNF interrupted for duration of 1 step, while the eMBB has more interrupted VNFs. To have more details about the interrupted VNFs, Fig. 4(b) shows the reconfiguration plan of VNFs migrations. We can see that UPF and SMF dedicated to eMBB are interrupted for 3 steps. This is because of the low SLA availability of 20% and 40% respectively. Then, the AMF shared between eMBB and mMTC is interrupted for 1 step, which explains the importance of the service availability. The ILP model takes into consideration the availability of each slice while minimizing the interruption duration.

2) Evaluation according to the nature of datasets: In this section, we evaluate the datasets presented in Table II and Table III for acyclic and cyclic graphs, respectively. In this experiments, we set β_i and α to 1 to focus more one the nature of graphs and their impact on the results.

 ISBN: 978-1-61208-940-9

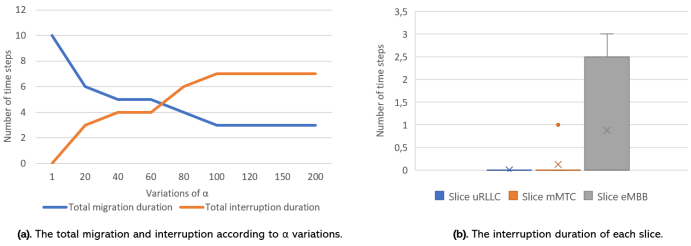


(a). The VNFs migration with the SLA availability according to each service slice.

Start Step	Actions	Is interrupted ?	Interruption duration
1	Migrate NSSF to S3	NO	--
1	Migrate AUSF to S4	NO	--
1	Migrate SMF to S3	NO	--
1	Migrate UPF to S1	NO	--
1	Migrate PCF to S5	NO	--
2	Migrate NRF to S2	NO	--
2	Migrate UPF to S4	NO	--
2	Migrate UDM to S1	NO	--
2	Migrate PCF to S4	NO	--
3	Migrate AMF to S1	NO	--
3	Migrate SMF to S2	YES	3
3	Migrate AMF to S3	YES	1
3	Migrate UPF to S2	YES	3

(b). The reconfiguration plan for all slices.

Fig. 4. The reconfiguration plan of all VNFs migrations taking into consideration the SLA availability



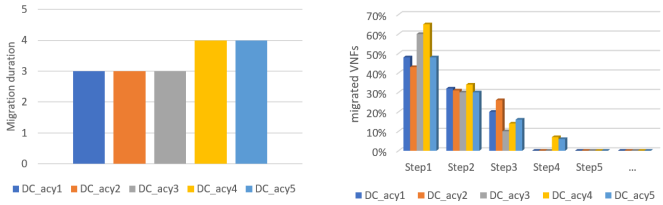
(a). The total migration and interruption according to α variations.

(b). The interruption duration of each slice.

Fig. 5. The evaluation results of migration and interruption cost

TABLE IV
COMPARISON BETWEEN THE ILP MODEL AND TS ALGORITHM FOR ACYCLIC GRAPH

Instances	**ILP**: Execution time (s)	**ILP**: Best objective	**TS**: Execution time (s)	**TS**: Best objective
DC-acy1	0.33	3	0.000532	25
DC-acy2	1.06	3	0.000324	35
DC-acy3	2.08	3	0.000949	60
DC-acy4	17.76	4	0.001346	80
DC-acy5	36.19	4	0.001257	150



(a). Total migration duration.

(b). Percentage of migrated VNFs per steps of migrations.

Fig. 6. The evaluation results of migration duration for acyclic graph

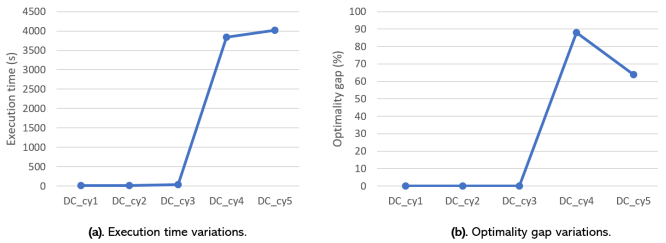
- **Acyclic graph**

For acyclic graphs, our ILP model solves the VNF reconfiguration problem without interruption and with 0% of optimality gaps. As we mentioned in Section II-B, in the case of an acyclic graph, we can find the reconfiguration plan in polynomial time without interruption using the TS algorithm. In Table IV, we compare our ILP model with the TS algorithm. We can see that the TS finds a reconfiguration plan in milliseconds comparing to our ILP model. However, the best objective of our ILP model is more interesting than the TS algorithm. This is because the ILP finds the optimal solution that minimizes the migration duration while the TS finds a feasible solution without taking into consideration the migration duration. This means that the ILP gives a solution where the VNFs are migrated from the early steps and in parallel as long as possible (see Fig. 6(b)) and with minimum migration duration (see Fig. 6(a)). Fig. 6 shows that the ILP migrates all VNFs in the first four steps for all instances.

- **Cyclic graph**

Figures 7(a) and 7(b), respectively, show the evolution of the execution time and the gap to optimal estimated by CPLEX at the end of the execution time according to the different instances. These two metrics significantly increase for DC-cy4 and DC-cy5 instances due to the np-hardness of the problem. Optimality gaps are often at 0% except in case of DC-cy4 and DC-cy5 instances which need more time to find an optimal solution. The ILP converged to optimality for medium graphs (120 to 146 VNFs) about over an hour of simulation. It provides an interesting solution in terms of migration duration and VNF interruption. Fig. 8(a) shows that the VNFs can migrate over 7 steps for DC-cy4 and over 6 steps for DC-cy5. The interrupted VNFs are less than 20% and 10%, respectively, for DC-cy4 and DC-cy5 (see Fig. 8(a)). In Fig. 8(b), we can see that most VNFs are migrated without interruption (hot migration where $\delta_i = 0$).

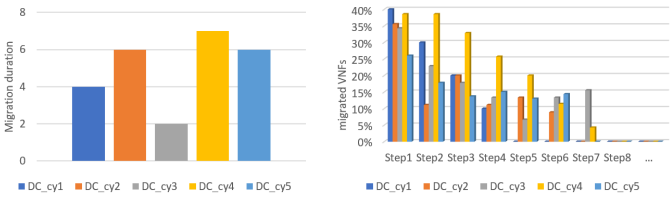
Like acyclic graphs, the ILP succeeds for cyclic graphs



(a). Execution time variations.

(b). Optimality gap variations.

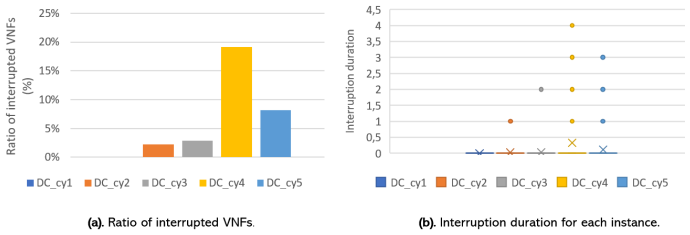
Fig. 7. Scalability evaluation results for cyclic graph



(a). Total migration duration.

(b). Percentage of migrated VNFs per steps of migrations.

Fig. 8. The evaluation results of migration duration for cyclic graph



(a). Ratio of interrupted VNFs.

(b). Interruption duration for each instance.

Fig. 9. The evaluation results of interruption duration for cyclic graph

to migrate efficiently and quickly the VNFs from the first steps with minimum interruptions. However, in acyclic graphs the migrations is performed without interruptions and slightly faster when compared to the case of cyclic graphs. This leads to conclude that the ILP model complexity depends strongly on the presence of cycles.

V. Conclusion and Future Works

In this paper, we have proposed an ILP-based solution for the problem of slice reconfiguration in the context of 5G networks. The ILP finds a reconfiguration plan, consisting of a series of migrations that will relocate the VNFs from their current servers to those computed beforehand, while minimizing the migration and interruption duration. We evaluate the proposed model according to the service importance taking into consideration the SLA availability metric, and according to the nature of datasets (whether it is an acyclic or a cyclic graph). The simulations reveal some strengths of our model in terms of slice service availability. In addition, evaluation results show that the ILP model yields good solutions in terms of minimizing the total migration and VNF interruption duration. As a future work, we plan to propose a heuristic based on topological sorting algorithm in order to improve the convergence time and allow dealing with larger instances.

References

[1] X. Foukas, G. Patounas, A. Elmokashfi, and M. K. Marina, "Network slicing in 5g: Survey and challenges," *IEEE Communications Magazine*, vol. 55, no. 5, pp. 94–100, 2017.

[2] I. Afolabi, T. Taleb, K. Samdanis, A. Ksentini, and H. Flinck, "Network slicing and softwarization: A survey on principles, enabling technologies, and solutions," *IEEE Communications Surveys Tutorials*, vol. 20, no. 3, pp. 2429–2453, 2018.

[3] "5g; procedures for the 5g system (3gpp ts 23.502 version 16.5.0 release 16)," (2020-07). [Online]. Available: https://www.etsi.org/deliver/etsi_ts/123500_123599/123502/16.05.00_60/ts_123502v160500p.pdf

[4] Vmware. [Online]. Available: https://www.vmware.com/

[5] Docker. [Online]. Available: https://www.docker.com/

[6] Kubernetes. [Online]. Available: https://kubernetes.io/fr/

[7] J. Liu, W. Lu, F. Zhou, P. Lu, and Z. Zhu, "On dynamic service function chain deployment and readjustment," *IEEE Transactions on Network and Service Management*, vol. 14, no. 3, pp. 543–553, 2017.

[8] A. Gausseran, F. Giroire, B. Jaumard, and J. Moulierac, "Be scalable and rescue my slices during reconfiguration," in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, 2020, pp. 1–6.

[9] V. Eramo, M. Ammar, and F. G. Lavacca, "Migration energy aware reconfigurations of virtual network function instances in nfv architectures," *IEEE Access*, vol. 5, pp. 4927–4938, 2017.

[10] V. Eramo, E. Miucci, M. Ammar, and F. G. Lavacca, "An approach for service function chain routing and virtual function network instance migration in network function virtualization architectures," *IEEE/ACM Transactions on Networking*, vol. 25, no. 4, pp. 2008–2025, 2017.

[11] R. Sirdey, J. Carlier, H. Kerivin, and D. Nace, "On a resource-constrained scheduling problem with application to distributed systems reconfiguration," *European Journal of Operational Research*, vol. 183, pp. 546–563, 2007. [Online]. Available: https://hal.inria.fr/inria-00311377

[12] D. E. Knuth and J. L. Szwarcfiter, "A structured program to generate all topological sorting arrangements," *Information Processing Letters*, vol. 2, no. 6, pp. 153–157, 1974.

[13] D. Ajwani, A. Cosgaya-Lozano, and N. Zeh, "A topological sorting algorithm for large graphs," *ACM J. Exp. Algorithmics*, vol. 17, sep 2012. [Online]. Available: https://doi.org/10.1145/2133803.2330083

[14] J. Sun, D. Ajwani, P. K. Nicholson, A. Sala, and S. Parthasarathy, "Breaking cycles in noisy hierarchies," in *Proceedings of the 2017 ACM on Web Science Conference*, ser. WebSci '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 151–160. [Online]. Available: https://doi.org/10.1145/3091478.3091495

[15] Breaking cycles in noisy hierarchies. [Online]. Available: https://github.com/zhenv5/breaking_cycles_in-noisy_hierarchies

 ISBN: 978-1-61208-940-9

Using Distributed Ledger Technology for Command and Control and Decentralized Operations

David Last, Michael Atighetchi, Partha Pal, Edward Lu
Raytheon BBN Technologies
Cambridge, Massachusetts, USA
{david.last, michael.atighetchi, partha.pal, edward.lu}@raytheon.com

Ryan Toner
Air Force Research Laboratory
Rome, New York, USA
ryan.toner@us.af.mil

***Abstract* — The US military is developing the warfighting philosophy of Multi-Domain Command and Control (MDC2), which integrates land, sea, air, space, and cyberspace into a unified operation environment. MDC2 depends on the consistent sharing of operational plans and intelligence reports, which will be contested by adversary advances in communications-denying technology. Thus, the MDC2 system of the future must enable to development and dissemination of plans within the context of intermittent communications. We are developing a proof-of-concept MDC2 prototype to explore the requirements and constraints of this space. This system will be built on top of a distributed database; after evaluating the available options, we believe that Distributed Ledger Technology (DLT) is a strong candidate to meet the particular requirements of the MDC2 use case. Here, we investigate several DLT options and compare their capabilities to the MDC2 requirements, analyzing the design tradespace. We also examine several DLT alternatives and identify why they do not meet these requirements. We develop two initial prototype MDC2 systems, a baseline system based on an SQL-type relational database and one based on DLT. We run experiments to compare the performance of the two prototypes, and we discuss how these results relate to their suitability for MDC2. Finally, we outline the future path for this research in order to complete a full-functionality prototype MDC2 system.**

Keywords—Distributed Ledger Technology; blockchain; Command and Control; Multi-Domain Command and Control.

I. INTRODUCTION

With a view towards the battlefield of the future, Department of Defense (DoD) policy over the past half-decade has been moving towards the Multi-Domain Command and Control (MDC2) concept (also called Distributed Maritime Operations, Multi-Domain Operations, and All-Domain C2) [1] [2] [3]. MDC2 integrates the warfighting domains of land, sea, air, space, and cyberspace into a unified planning process under a single Joint Forces Commander. At the same time, adversaries are advancing their battlefield capabilities for jamming and otherwise hindering communications. Thus, the battlefield of the future will not resemble communications-permissive battlefields the DoD has enjoyed the last few decades. Based on these two trends, there is an increased need for front line units to share military plans and intelligence, but also the potential for significant barriers to doing so.

In order to address this situation, the DoD needs to develop a next-generation MDC2 system that allows commanders to share plans and orders across an entire theater of operations, but also empowers frontline units to collaboratively update plans in response to changing battlefield conditions. This MDC2 system must maintain a consistent view of the data among all parties (when in communication) and reconcile conflicts in the data (when re-connecting after a period of denied communication).

This MDC2 system should be built on top of a distributed database that can operate through intermittent communication and also reconcile divergent database copies that result from evolving data during network disconnection. Upon analysis of the system requirements, we believe that Distributed Ledger Technologies (DLT) are a promising option for this MDC2 system. The research presented in this paper describes an investigative study to evaluate the suitability of DLT for such a system.

This paper is organized as follows. In Section II, we present the MDC2 use case that motivates our research. Section III explores our reasons for selecting DLT as a solution for this problem, and Section IV details the different DLT implementations we considered. There are other solutions to this problem besides DLT; Section V presents some industry-standard alternatives and discusses their advantages and disadvantages compared to DLT for our use case. Section VI discusses our experiments with our DLT-based prototype, and Section VII recommends avenues for future research (both experimentation and development).

II. MULTI-DOMAIN COMMAND AND CONTROL USE CASE

Consider the following scenario: an Air Force base in a war zone is under threat of imminent attack. In an effort to protect his forces, the Air Commander divides his air units into small groups called Dispersed Units (DU); these DUs are organized into a hierarchy of Parent Dispersed Units (PDU). These DUs are then deployed to forward operating bases to geographically separate them (Figure 1).

Prior to dispersing the DUs, the Air Commander and his planning staff plan out the air war for the next 2 weeks. They assign different DUs to destroy or recon different targets, and they distribute these plans to the DUs. During the mission, the forward-deployed DUs are disconnected from the mission planners and other DUs due to geography, adversary jamming, cyber attack, etc. Additionally, the DUs discover that the battlefield is changing from its initial state as seen by the mission planners. Targets are in different locations than originally thought, new threats are discovered, etc. These changes invalidate the original mission plans, and necessitate updates/modifications to the plans in order to achieve mission success and deal with these new threats. Normally, the DUs would request plan updates from the mission planners, but they are now out of communication. Therefore, front line units in the DUs must be delegated authority to re-plan and re-task local units, and they must record these updates to the plan. When the DUs re-establish communications with the home base, these changes must be communicated to the Air Commander and mission planners. The original plans and the updated plans must be reconciled into a consistent, updated view of the plans so that all parties will be on the same page.

There are several critical requirements for a database to store these plans and enable decentralized operations. First, the database must provide ***consistent data*** between different network participants (to the extent possible). Second, when plans do diverge due to disconnected communications, different versions of the plans must be ***deconflicted*** once communications are re-established. Third, this database must be

 ISBN: 978-1-61208-940-9

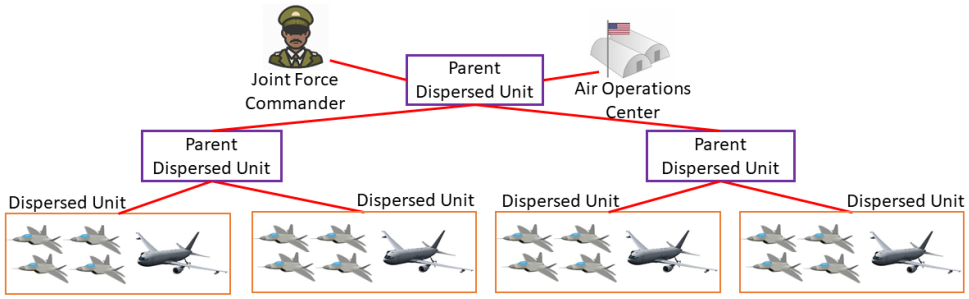


Figure 1. Organizational Hierarchy of Dispersed Units.

leaderless; there can be no single point of failure where disrupted communications means that network participants are unable to update or disseminate the plans. Fourth, the database must provide ***immutable***, ***auditable provenance***; any party examining the database should be able to examine the entire evolution history of the plans to see who made what changes, when, and why. This provenance history must be tamper-proof so that adversaries and bad actors cannot modify the history after the fact. Fifth, the database must also provide ***non-repudiation*** guarantees so that parties writing to the database can be held accountable for their updates.

Although we have presented a Department of Defense use case here, such a system would have wide applications. Consider a wilderness search and rescue scenario. Drones and human volunteers are collaborating to search a National Park for a lost hiker. This is a remote, austere environment where the searchers (human and drone) use relatively low-powered radios for communication, and communications will often be interrupted due to distance and geography. The search is led by a single Search and Rescue Coordinator, and on-the-ground searchers are divided into groups (DUs) and assigned to different search areas. The different groups need to collaboratively re-plan and re-assign roles during the search in response to changing intelligence and environmental conditions, and they need to communicate their search results back to the Coordinator after they complete their search pattern.

This type of command and control system has many other applications. It could be used to coordinate Border Patrol units or to manage drug enforcement operations. This system would be useful for collaborative planning between groups of autonomous drones working towards a common goal, such as mapping a remote area. It could even be used to manage logistics systems like the United States Postal Service (USPS), UPS, FedEx, or DHL.

III. WHY DLT?

The Command and Control system described here requires a distributed database that allows multiple parties to modify the same data. It must allow updates to the data in disconnected network partitions, and it must automatically reconcile data conflicts taking into account mission context in order to determine which version of the data is authoritative. In this research, we explore Distributed Ledger Technology (DLT), e.g., blockchain, as a potential solution for this system. DLT/blockchain was designed as a leaderless, distributed database that can tolerate network disconnection, and it has properties that are attractive for distributed Command and Control systems.

A. How DLT\blockchain works

In a DLT/blockchain network, all network participants maintain a local copy of the distributed database (ledger). All data operations on this database (create/update/delete) are encoded as Transactions and submitted to a pool of Proposed Transactions. Certain network participants (called "miners" in Bitcoin networks) select a set of Proposed Transactions, check them for "correctness" (according to a set of rules based on the characteristics of the system), bundle them into a Block, append that Block to the end of a chain of Blocks (hence blockchain), and advertise the new Block to other network participants. Other network participants verify the correctness of the Transactions and then accept the new Block. The Block contains a cryptographic hash of its own contents and the contents of the previous Block in the chain, which makes the Block ***immutable*** (any tampering with the Block will invalidate the hash, making the tampering instantly detectable).

In some cases, two different miners will create and advertise two different next Blocks. This results in different network participants having different versions of the blockchain; this is called a blockchain ***fork***. Different blockchain implementations have methods for avoiding forks, and they also have methods for determining which fork will be accepted as the authoritative blockchain; the other fork will be discarded.

A key aspect of DLT/blockchain is the consensus algorithm. If it is easy for miners to generate and propose new Blocks, then it will be easy for bad actors to manipulate the system. The consensus algorithm makes it difficult to generate a valid Block, but easy to check the validity of the Block. The consensus algorithm is also used to limit the blockchain mining speed, which reduces the frequency of blockchain forks and makes it more difficult for bad actors to manipulate. There are different types of consensus algorithms. Proof of Work requires a large amount of compute power to solve a difficult, but easily verifiable, math problem to generate a new Block; this algorithm is used in the Bitcoin network [4]. Proof of Stake requires network participants to "stake" a certain amount of owned cryptocurrency in order to become validators (same role as miners) who are randomly selected to create the next Block; this algorithm is used in the Ethereum network [5]. Proof of Authority is similar to Proof of Stake except that validators are chosen to create the next Block with probability of being chosen being proportional to the validator's reputation (based on its past behavior in the network) [6].

B. Design tradeoffs

Distributed Ledger Technology has a number of advantages related to the requirements of the distributed Command and Control use case. First, DLT/blockchain guarantees eventual consistency of the data (when all nodes have the same copy of the blockchain). Second, blockchain is designed for leaderless management so that there is no single point of

 ISBN: 978-1-61208-940-9

failure that can bring down the network. Although the command hierarchy in Figure 1 seems to indicate the Joint Forces Commander as a single leader for the C2 system, once authorities to modify plans are conditionally delegated to frontline units, this situation more closely resembles a leaderless network. Third, blockchain provides the required immutable provenance auditing and non-repudiation. Finally, DLT also enables smart contracts, which can be used for the delegation of authority described previously.

However, there are also drawbacks to using blockchain for distributed databases. Because the system is leaderless, it relies on the consensus algorithm to reach a consistent data state (rather than a single leader dictating the data state). The consensus algorithm is necessarily slower and more complex than a single master database. This has significant implications for the latency and throughput of Write transactions for a blockchain-based distributed database.

Based on these design tradeoffs for blockchain-based distributed databases, we evaluate that blockchain is best suited for data that evolves slowly and can tolerate latency. Therefore, high-volume data like sensor information or video streams are not well-suited for blockchain. The Command and Control data for collaborative mission planning are a good fit for blockchain-based systems; however, C2 applications with real-time requirements may not be suited for blockchain databases.

This paper outlines an investigatory effort that explores the feasibility of using blockchain for Command and Control collaborative planning.

IV. DLT FRAMEWORKS

As part of this research, we investigated a number of different open-source DLT implementations to serve as the basis for a distributed, collaborative Command and Control system. The final Command and Control system has the following requirements, so the basis DLT implementation should support these:

- Flexible roles/leaderless operation for network nodes (no single point of failure)
- Network partitions must support continued operation to some degree
- Configurable access control/authorities management
- Support for blockchain forking and reconciliation
- Not resource intensive (operation on mobile platforms with constrained communications)
- Permissioned network – all participants are known and authorized

With these requirements in mind, we evaluated to applicability of several different DLT implementations.

Option 1: **Hyperledger Fabric**

Hyperledger Fabric is an open-source project developed under the Linux Foundation. In contrast to many blockchain implementations (like the Bitcoin network), Fabric is permissioned, rather than permissionless. All network participants are known, and only certain network participants are authorized to add Transactions to the blockchain. Network participants are authorized with X.509 security certificates issued by a certificate authority. Network participants are divided into *organizations*, which share a single distributed ledger. Organizations can be grouped into a *consortium*, which allows participants from different organizations to access the distributed ledger of the other organizations. Hyperledger Fabric encodes access control policies into chaincode, which is used to govern database read/write operations. Because Hyperledger Fabric uses X.509 certificates and defined network validators, it does not require a resource-intensive consensus algorithm like Proof of Work [7].

Option 2: **R3 Corda**

R3 Corda is a distributed ledger that was developed for the financial services sector [8]. It uses the Unspent Transaction Output (UTXO) model (similar to Bitcoin) for managing data assets. R3 Corda does not use Proof of Work as a consensus algorithm; rather, it defines the concept of a Notary, where a single Notary must control all data assets consumed by a Transaction. If a single Notary does not control all the data assets, then control most be transferred before the proposed Transaction can be executed. The need for Notaries poses a significant problem for Command and Control in disconnected environments. A disconnected Command and Control system will require multiple ownership of data assets since we do not know a priori which node will need to modify which assets.

Option 3: **Algorand**

Algorand is a blockchain-based digital currency like Bitcoin that was created in 2017 by MIT professor Silvio Micali [9] to address some of the shortcomings of Bitcoin. Algorand uses Proof of Stake as a consensus algorithm. The Algorand network is permissionless, so any party can join as a network node. Additionally, Algorand supports only one class of user; all nodes in the system have the same authority level (although weighted by their stake in the system). This disallows the designation of "trusted" parties with varying levels of authority, which is necessary for the Command and Control delegation of authority. Because of these reasons, Algorand is not suitable for a Command and Control system.

There are many more blockchain implementations in this technology space, but most of them share basic characteristics with these three systems. Based on our analysis, we selected Hyperledger Fabric as the basis for our Command and Control system. Hyperledger Fabric supports some of the requirements for the Command and Control system (leaderless operation, configurable authorities, operation of network partitions, permissioned network), but other capabilities will need to be built around it as part of our prototype (blockchain forking and reconciliation). One key point in favor of Hyperledger Fabric is that **since it does not use a resource-intensive consensus algorithm like Proof of Work**, it does not suffer from well-known energy consumption concerns about cryptocurrency implementations like Bitcoin.

V. NON-DLT ALTERNATIVES

Blockchain is a relatively new approach to distributed databases. There are a number of more traditional distributed database solutions that should be considered as direct competitors to a blockchain-based system; any blockchain system should be evaluated against these other solutions.

SQL (Structured Query Language)–type databases are relational databases organized into tables with columns and rows. In a standard configuration, a single SQL-type database serves Read and Write requests from multiple Clients. SQL databases like MySQL, PostgreSQL, and Microsoft SQL Server do not natively support a distributed database configuration, but they can be configured into a single master/multiple replica configuration to support distributed Read operations but not distributed Write operations [10].

Git is a popular open-source Distributed Version Control System (VCS) that was developed in 2005 to manage soft-

 ISBN: 978-1-61208-940-9

ware development projects with multiple contributors. Git users can download local copies of a master database, make updates to the data, and then push the updates to be merged with the master database. Any local updates that do not conflict with the master database are merged automatically, but local updates that conflict with updates to the master database are flagged to the human user for manual merging. Git records a full history of who made what changes to the database, at what time [11]. Git works as a distributed database, but it is not leaderless; it relies on one database node serving as the authoritative master node.

The **Interplanetary File System (IPFS)** was designed as a Peer-to-peer file sharing system that works as a distributed database. Users of the database create files locally and IPFS divides the files into chunks, generates a cryptographically hashed Content ID (CID) for each chunk, and advertises the CIDs to other users in the network. If other users wish to download the file, IPFS queries the network for the location of the data associated with the relevant CIDs and downloads the chunks. That user then becomes a secondary provider for those CIDs until they are deleted. When new versions of a file are added to IPFS, they are stored using new CIDs; old versions of the file cannot be tampered with or erased (unless all providers of the CID delete their local copy). Within IPFS each file exists as an independent entity; there is no concept of conflicting versions of the same data and no merge/reconcile functionality [12].

VI. Experimentation

A. Prototypes

The goal of this research and the initial experiments it encompasses is to evaluate the feasibility of using Distributed Ledger Technology as a distributed database for a Command and Control system, as opposed to other distributed database solutions. To fulfill this goal, we built two prototypes for experimentation: one representing the state of the practice (built on top of PostgreSQL, a relational database), and one representing the state of the possible (built on top of Hyperledger Fabric). These two prototype networks each contain three nodes that function as a distributed database (Figure 2). PostgreSQL is not natively a distributed database, so that prototype is set up with a single master and two replicated copies. In this configuration, clients can only write to the master node, and these write operations are propagated to the replicated copies.

It is important to note that these two prototypes do not provide the same functionality. Because the PostgreSQL prototype is not a true distributed database, it has no need for a consensus algorithm, because only one node (the master node) is the arbiter of the correct data state; this also represents a single point of failure. Because there is no need for consensus in the PostgreSQL database, the message exchange between nodes will necessarily be much more complex in the Hyperledger Fabric prototype than the PostgreSQL prototype. Therefore, we fully expect that the PostgreSQL prototype will outperform the Hyperledger Fabric prototype in terms of throughput and latency when processing Write operations. The main advantage of Hyperledger Fabric over a more traditional database is that it does not contain a single point of failure, and that a partition of the network can continue operation even when disconnected from the rest of the network. The PostgreSQL prototype does not support either of these capabilities. The following experiments demonstrate how much of a performance downgrade Hyperledger

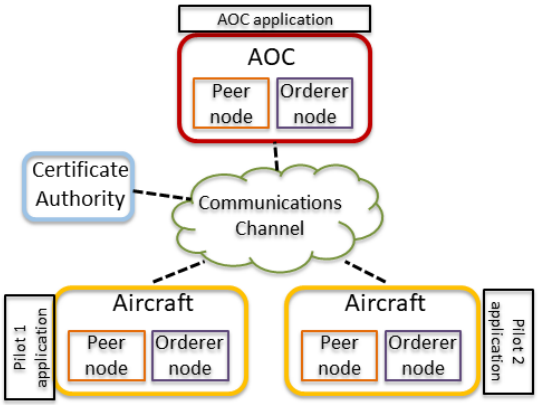


Figure 2. Hyperledger-based MDC2 Initial Prototype.

Fabric suffers as opposed to a more traditional approach in order to analyze the tradeoffs between basic performance metrics and the special functionality that blockchain provides. These experiments also provide indications as to which use cases are best suited for a blockchain-based approach.

B. Experimentation

We ran several experiments to compare the performance of the Hyperledger Fabric- and PostgreSQL-based prototypes. The different nodes of the network were run in Docker containers on an Ubuntu Linux VM. We used the Pumba tool [13] to simulate bandwidth degradation and disconnection between different network nodes. For these experiments, we use Write Transactions that write representations of Link 16 J2.2 messages to the distributed database (Link 16 J2.2 messages are Air Force self-position reports for military aircraft) [14].

C. Experiment 0 – Hyperledger Fabric parameters

Our first experiment was an initial exercise of the Hyperledger Fabric prototype to explore its capabilities and experiment with major configuration parameters to identify the optimal configuration for our use case. We experimented with two independent variables: Batch Timeout and Traffic Density.

One of the major configuration parameters for Hyperledger Fabric is Batch Timeout. This value, expressed in fractions of a second, instructs the Hyperledger Fabric nodes that once they receive a Write Transaction, how long they should wait for additional Transactions before bundling all available Transactions into a new Block to be added to the blockchain. If this parameter is set to a low (short) value, then new Transactions will be bundled into Blocks almost as soon as they are received. This may improve the Transaction throughput, but an increased number of Blocks being processed by the consensus algorithm can increase Transaction latency. On the other hand, if this parameter is set to a high (long) value, it can increase Transaction throughput (because there are fewer Blocks, there is less network overhead per Transaction), but because there are fewer Blocks, it can also (counterintuitively) decrease the average Transaction latency. In this experiment, we vary the value of Batch Timeout to find the optimal setting for our use case.

For the Batch Timeout experiment, we use the 3-node network configuration shown in Figure 2. Two Clients commit Link 16 J2.2 Write Transactions at a frequency of 500 milliseconds for each Client. These Clients commit Transactions

 ISBN: 978-1-61208-940-9

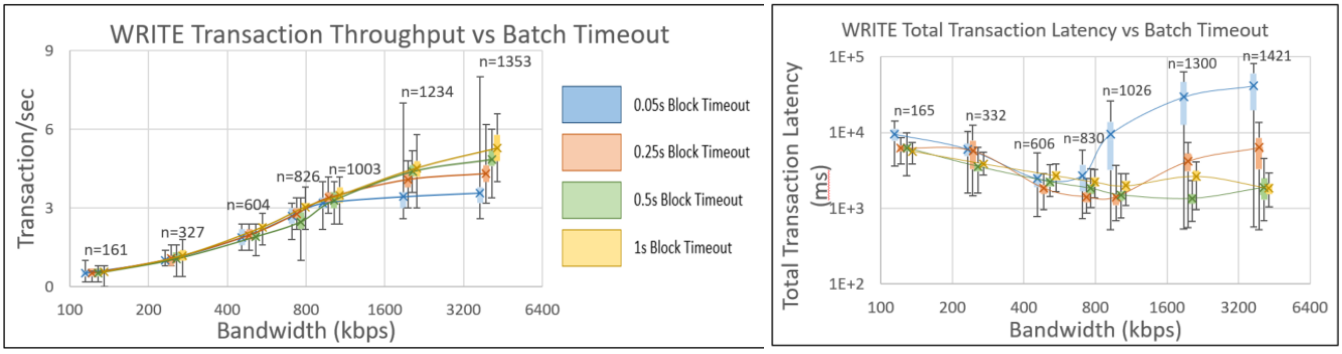


Figure 3. Experiment 0 - Hyperledger Fabric prototype Write Transaction throughput and latency vs. Batch Timeout.

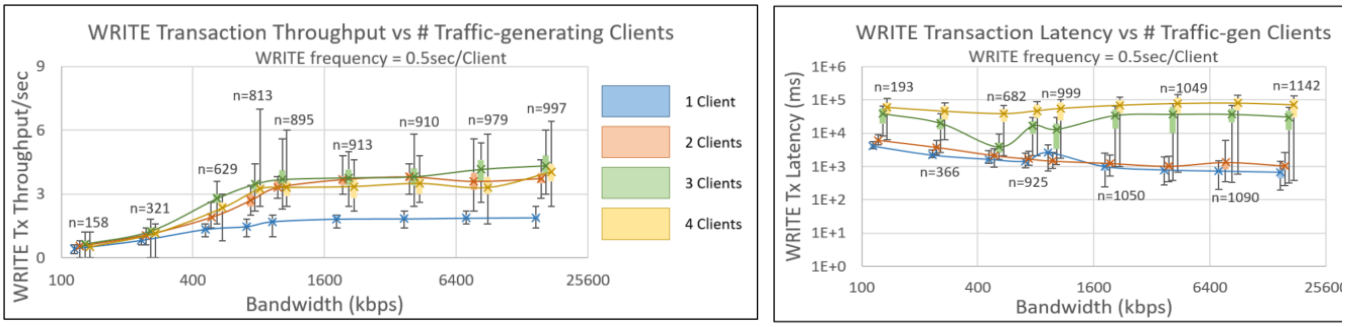


Figure 4. Experiment 0 - Hyperledger Fabric prototype Write Transaction throughput and latency vs. Traffic Density.

for a period of 300 seconds. We vary the Batch Timeout parameter between 0.05s, 0.25s, 0.5s, and 1s, and we also vary the bandwidth of the Hyperledger Fabric network to simulate degraded communications. We measure the Transaction throughput and latency to determine the optimal Batch Timeout for Hyperledger Fabric. The experimental results are shown in Figure 3; we determine that the optimal Batch Timeout parameter is 0.5 seconds.

We also run an experiment to determine the maximum traffic density the Hyperledger Fabric network can handle before it begins to affect performance. We use 1-4 Clients, which each submit Link 16 J2.2 Write Transactions to the network at a frequency of 500 milliseconds, and we run the experiment for 300 seconds. We use a Batch Timeout vale of 0.5 seconds based on the previous experiment. We measure the Transaction throughput and latency, and the results are shown in Figure 4. We also ran these experiments for 5 and 6 Clients, but those results are similar to the results for 4 Clients. Based on the experimental results, we determine that the amount of traffic generated by 2 or 3 Clients represents the best tradeoff between throughput and latency, depending on the situation.

D. *Experiment 1*

Following the Hyperledger Fabric parameter tuning in Experiment 0, we run the first experiment that compares the Hyperledger Fabric and PostgreSQL prototypes head-to-head. In this experiment, 2 Clients write Link 16 J2.2 messages to the distributed database at a frequency of 300 milliseconds per client; the database Transactions are approved (according to the prototype's approval mechanism) and then propagated to all nodes in the network. The experiment lasts for a period of 300 seconds. We vary the bandwidth available to the networks to simulate degraded communications, and we measure the throughput and latency of the networks.

Based on the difference in complexity of the Transaction approval mechanism between the two prototypes, we expect the PostgreSQL prototype to outperform Hyperledger Fabric in both throughput and latency; however, we wish to see if the difference between these metrics is sufficiently low to justify the benefits of Hyperledger Fabric in our Command and Control use case. We show the experimental results in Figure 5.

E. *Experiment 3*

We also run an initial experiment to compare the performance of the two prototype systems in a disconnected communications scenario. In this experiment, we measure how long it takes to merge new database Transactions into a database node that has not yet received them. At the beginning of this experiment, a set of Clients write 500 Link 16 J2.2 Transactions to the database; these Transactions are propagated to all nodes in the network. Then, one of the database nodes is partitioned from the rest of the network (in PostgreSQL, this is one of the replicated nodes). The Clients write an additional 500 J2.2 Transactions to the main network; the partitioned node does not receive these Transactions. We then reconnect the partitioned node to the network, and the network automatically pushes the new Transactions to the reconnected node. We measure how long it takes for the reconnected node to be brought fully into sync with the rest of the distributed database nodes. The results are shown in Figure 6.

F. *Future Experiments*

In the future, we plan to run additional experiments to further evaluate the performance of the Hyperledger Fabric prototype against the PostgreSQL prototype.

Experiment 2 – Hardware and Network Requirements: This experiment will use the same procedure as Experiment 1. We will measure the disk storage required at each node, the processing power for 1 Write Transaction, and the network overhead for submitting 1 Write Transaction and propagating it to all distributed nodes.

Experiment 4 – Dynamic Data Merging: This experiment will use the same procedure as Experiment 3, except that the

 ISBN: 978-1-61208-940-9

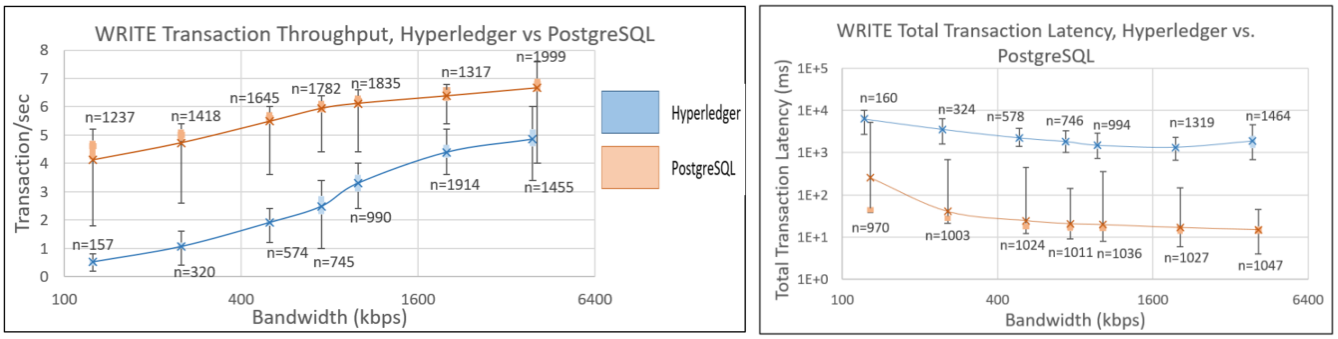


Figure 5. Experiment 1 - Hyperledger Fabric vs. PostgreSQL Prototypes Write Transaction throughput and latency.

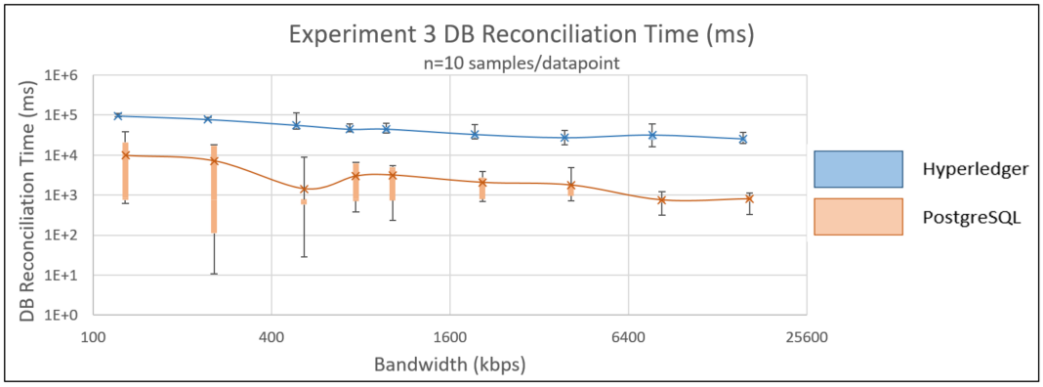


Figure 6. Experiment 3 - Hyperledger Fabric vs. PostgreSQL prototypes Write Transaction throughput and latency.

Clients will continue to write new Transactions to the database during the Merge period. We will measure how long it takes for the partitioned node to come back into synchronization with the rest of the nodes, and what is the effect of the increased network traffic due to the new Write Transactions.

G. Discussion

The experiments discussed here show that in initial performance tests, a blockchain-based Command and Control system performs worse in terms of latency and throughput than a simpler, non-collaborative SQL-type relational database. The key distinction between the two is the consensus algorithm; it enables leaderless operation and disconnected communication tolerance, but it introduces significant complexity in validating and committing Write Transactions. Therefore, the key questions for evaluating the suitability of blockchain are:

- Based on its performance constraints, what type of data is blockchain best suited for?
- Do the benefits of blockchain (security, leaderless operation to tolerate degraded communications) outweigh its performance penalties for our specific use case?

The Experiment 0 results indicate that a blockchain-based system is best-suited for low Write volume data with a moderate tolerance for Write latency. Therefore, blockchain is most applicable to high value data that does not change frequently (like military campaign plans), but not high-volume, low latency data (like real-time control signals, sensor data, or video streams). Experiments 1 and 3 bolster this determination. Over the course of a multi-day military engagement, plans will probably be added to the database at less than 3-5 Transactions per second, and the Command and Control system can tolerate a 1-10 second latency on the dissemination of these plans to frontline units (Figure 5). Additionally, isolated frontline units re-establishing communication with the main group can tolerate 20-100 seconds to download Command and Control updates (Figure 6).

VII. Recommended Future Work

The research discussed in this paper was performed as a study to answer the question, “Is it feasible to use blockchain as a Command and Control distributed database?” Since this research has answered this question in the affirmative, the next step in this research is to build a full prototype. This prototype will incorporate several innovations beyond the initial study.

The first innovation will be blockchain branching and merging. Most current blockchain implementations address blockchain forks by requiring a majority of network nodes to write to the blockchain, thus explicitly preventing forks (Hyperledger Fabric uses this approach), or they resolve forks by determining one fork branch to be authoritative and discarding the other branches (Bitcoin’s blockchain implementation uses this approach). Neither of these approaches are sufficient for the Command and Control scenario: requiring a majority of nodes to write new Transactions prevents minority partitions from writing new data, and discarding a blockchain fork (which represents the collaborative planning of a minority partition) invalidates previous planning and decision-making, throwing the entire Command and Control system into chaos. Therefore, the full prototype needs new functionality to allow blockchain forks in minority partitions, as well as merging these blockchain forks within an understanding of the context of the larger mission.

The second innovation for the full prototype will be the implementation of a conditional authority calculus. In a full communications environment, planning decisions should be made by the highest-ranking authority and disseminated to lower-ranking units. If communications are disconnected, these lower-ranking units must be authorized to make these planning decisions. However, if there are no constraints on which units can make which planning decisions, this can lead to an explosion of blockchain branches that will be very com-

 ISBN: 978-1-61208-940-9

plex to maintain and merge. Our conditional authority calculus will use a dynamic ruleset that is evaluated in the context of the mission environment to determine which parties are allowed to make which planning decisions at a specific point in time. This will constrain the complexity of the planning process and the resultant blockchain merges.

As we develop this full prototype, we will also pursue opportunities to deploy it during Department of Defense field exercises in order to evaluate its performance in operational scenarios and begin building acceptance within the user community.

VIII. Conclusion

In recent years, the DoD has been moving towards the Multi-Domain Command and Control philosophy as the most effective way to integrate warfighting domains. However, adversary advances in communications-denying technologies jeopardize the ubiquitous communications needed to realize MDC2. Therefore, the DoD needs an advanced MDC2 system that enables collaborative planning and information sharing in the presence of constrained, intermittent communications. Based on our investigation, we believe that Distributed Ledger Technology is a strong candidate for such a system that supports the communication requirements of the MDC2 scenario. In this paper, we investigate different DLT implementations and evaluate them against the MDC2 scenarios; we identify Hyperledger Fabric as meeting the key requirements for MDC2. We built two different MDC2 prototypes: one based on standard distributed database technology, and one based on Hyperledger Fabric. We ran a number of experiments to evaluate the performance of the two systems, and to evaluate whether Hyperledger's performance is sufficient for an MDC2 system. Our experimental results are encouraging, so we chart a path forward to build a production-grade DLT-based MDC2 system that can operate in modern, communications-denied environments.

Statements/Disclaimers

Distribution Statement "A" (Approved for Public Release, Distribution Unlimited). Case # AFRL 2022-0076. This effort is sponsored by the Air Force Research Laboratory (AFRL).

The views expressed are those of the authors and do not reflect the official guidance or position of the United States Government, the Department of Defense or of the United States Air Force.

Statement from DoD: The appearance of external hyperlinks does not constitute endorsement by the United States Department of Defense (DoD) of the linked websites, or the information, products, or services contained therein. The DoD does not exercise any editorial, security, or other control over the information you may find at these locations.

References

[1] US Air Force, "Air Force future operating concept: A view of the Air Force in 2035." Washington, DC: Government Printing Office, 2015. Accessed: Apr. 20, 2022. [Online]. Available: https://www.af.mil/Portals/1/images/airpower/AFFOC.pdf

[2] J. M. Richardson, "A design for maintaining maritime superiority, Version 2.0." Naval College War Review, Dec. 17, 2018. Accessed: Apr. 20, 2022. [Online]. Available: https://media.defense.gov/2020/May/18/2002301999/-1/-1/1/DESIGN_2.0.PDF

[3] US Army Training and Doctrine Command, "The U.S Army in multi-domain operations 2028." Training and Doctrine Command, Ft. Eustis, VA, Dec. 06, 2018. Accessed: Apr. 20, 2022. [Online]. Available: https://adminpubs.tradoc.army.mil/pamphlets/TP525-3-1.pdf

[4] "Proof of work," *bitcoin.it*. https://en.bitcoin.it/wiki/Proof_of_work (accessed Apr. 20, 2022).

[5] "Proof of stake (POS)," *ethereum.org*, Dec. 09, 2021. https://ethereum.org/en/developers/docs/consensus-mechanisms/pos/ (accessed Apr. 20, 2022).

[6] "Proof of authority explained," *binance.com*, Dec. 09, 2020. https://academy.binance.com/en/articles/proof-of-authority-explained (accessed Apr. 20, 2022).

[7] E. Androulaki *et al.*, "Hyperledger Fabric: A distributed operating system for permissioned blockchains," 2018.

[8] R. G. Brown, J. Carlyle, I. Grigg, and M. Hearn, "Corda: an introduction," *R3 CEV August*, vol. 1, p. 15, Aug. 2016.

[9] E. Auditore, "Algorand: origins," *Algorand: Origins*. https://community.algorand.org/blog/algorand-origins/ (accessed Apr. 20, 2022).

[10] M. Kamaruzzaman, "Top 10 databases to use in 2021," *towardsdatascience.com*, Jan. 20, 2021. https://towardsdatascience.com/top-10-databases-to-use-in-2021-d7e6a85402ba (accessed Apr. 20, 2022).

[11] D. Spinellis, "Git," *IEEE Softw.*, vol. 29, no. 3, pp. 100–101, Jun. 2012.

[12] "Interplanetary File System (IPFS)," *IPFS*. https://ipfs.io/ (accessed Apr. 20, 2022).

[13] A. Ledenev, "Pumba: chaos testing tool for Docker (Github)," *Pumba: chaos testing tool for Docker (Github)*. https://github.com/alexei-led/pumba (accessed Apr. 20, 2022).

[14] Air Land Sea Application Center, "Introduction to Tactical Digital Information Link J and quick reference guide (TADIL J)." Jun. 2000. Accessed: Apr. 20, 2022. [Online]. Available: https://apps.dtic.mil/dtic/tr/fulltext/u2/a404334.pdf

 ISBN: 978-1-61208-940-9