JOURNAL OF TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY

3/2024

(Contents continued on back cover)

Editor-in-Chief

Adrian Kliks, Poznan University of Technology, Poland

Steering Editor

Paweł Pławiak, National Institute of Telecommunications, Poland

Editorial Advisory Board

Hovik Baghdasaryan, National Polytechnic University of Armenia, Armenia
Naveen Chilamkurti, LaTrobe University, Australia
Luis M. Correia, Instituto Superior Técnico, Universidade de Lisboa, Portugal
Pedro Crespo Bofill, Universidad de Navarra, Spain
Luca De Nardis, DIET Department, University of Rome La Sapienza, Italy
Nikolaos Dimitriou, NCSR "Demokritos" Athens, Greece
Ciprian Dobre, Politechnic University of Bucharest, Romania
Piotr Gawrysiak, Warsaw University of Technology, Poland
Filip Idzikowski, Poznan University of Technology, Poland
Andrzej Jajszczyk, AGH University of Science and Technology, Poland
Zbigniew Jaroszewicz, National Institute of Telecommunications, Poland
Albert Levi, Sabanci University, Turkey
Marian Marciniak, National Institute of Telecommunications, Poland
George Mastorakis, Technological Educational Institute of Crete, Greece
Constandinos Mavromoustakis, University of Nicosia, Cyprus
Takumi Miyoshi, Shibaura Institute of Technology, Japan
Klaus Mößner, Technische Universität Chemnitz, Germany
Imran Muhammad, King Saud University, Saudi Arabia
Mjumo Mzyece, University of the Witwatersrand, South Africa
Daniel Negru, University of Bordeaux, France
Jordi Perez-Romero, UPC, Spain
Michał Pióro, Warsaw University of Technology, Poland
Konstantinos Psannis, University of Macedonia, Greece
Salvatore Signorello, University of Lisboa, Portugal
Adam Wolisz, Technische Universität Berlin, Germany
Tadeusz A. Wysocki, Univeristy of Nebrasca, USA

Editorial Team

Content Editor: **Robert Magdziak**
Managing Editor: **Ewa Kapuściarek**

eISSN 1899-8852

© Copyright by National Institute of Telecommunications, Poland 2024

Testing Time Optimization Method for IEEE 802.15.4z Ultra-wideband Integrated Circuits

Grigor Tsaturyan[1] and Hovhannes Gomtsyan[1,2]

[1]*National Polytechnic University of Armenia, Yerevan, Armenia,*
[2]*European University of Armenia, Yerevan, Armenia*

https://doi.org/10.26636/jtit.2024.3.1623

Abstract — IEEE 802.15.4z-compliant ultra-wideband (UWB) devices are becoming ever more popular in contemporary radio engineering systems. Such systems are capable of precisely measuring distances (with their accuracy expressed in centimeters), are immune to interference, offer low latency and transmit data in an energy-efficient manner. Widespread adoption of UWB technology has triggered significant demand for testing integrated circuits these systems rely on, prompting the development of new testing methods to meet the ever increasing requirements in terms of testing speed and reliability. The same applies to sensitivity tests, in the course of which up to 2000 different packets may be received. The process of generating and analyzing such a large number of packets is time consuming. Furthermore, if multiple devices need to be tested simultaneously, the duration of the test will be multiplied accordingly. In such a context, the article investigates the lead time required to generate 2000 UWB packets using conventional methods and proposes a novel approach to significantly reduce packet generation time and improve testing efficiency.

Keywords — interframe spacing, packet generation, RF circuit testing, sensitivity measurements, test time reduction, UWB

1. Introduction

IEEE 802.15.4z UWB systems are short-range wireless communication solutions operating within the frequency range of 3.1 to 10.6 GHz. These systems rely on extremely low radiation power levels which are limited, under FCC regulations [1], to –41.3 dBm/MHz. Such a low spectral power allows UWB systems to coexist with narrowband type signal systems, minimizing potential interference. This advantage allows to take advantage of UWB devices' precise centimeter-level distance determination capabilities in environments in which Wi-Fi and other wireless solutions are used [2].

UWB devices determine distance through the measurement of time of flight (ToF). ToF detection can be achieved through such methods as single-sided two-way ranging or double-sided two-way-ranging, relying on packet exchange between two devices. The first device sends a packet containing a marker. By receiving the packet and detecting the position of the marker, the second device can determine the propagation time of the electromagnetic wave by measuring the time difference between the sent and received markers [3].

As UWB devices are required to be able to receive packets under different circumstances, their sensitivity needs to be measured. Receiver sensitivity is a parameter that describes the lowest power level at which the receiver can detect a UWB signal. In order to measure sensitivity of a receiver a test system is required to generate the necessary packets.

Therefore, in order to be able to reliably determine the sensitivity of UWB integrated circuits (IC), the test equipment should be able to generate 2000 different UWB data packets [4]. The device under test (DUT) receives these packets and assesses the sensitivity level based on the packet error rate (PER). If PER is lower than 1%, the test is considered passed [5]. Generating such a large number of packets will increase the time required for testing a single device.

Additionally, in scenarios involving multiple DUTs undergoing validation, verification or production test phases, the overall test duration multiplies accordingly. Thus, measures should be taken to reduce the time required to generate a large number of packets.

This paper proposes a novel method that greatly reduces the time required for generating a large number of packets by leveraging the National Instruments PXIe-5831 vector signal transceiver (VST) and utilizing script mode software capabilities.

The article is organized as follows. In Section 2, conventional methods currently used to generate the required waveforms are discussed. Thereafter, the algorithm of the proposed method aiming to reduce the waveform generation lead time is presented. In Section 3, PXIe-5831 VST is utilized to generate and receive waveforms using both the conventional approach and the proposed method, with the waveform generation lead time typical of each method being measured and compared. Conclusions are given in Section 4.

2. Waveform Generation Techniques

Four different UWB packet configuration types have been defined in the IEEE 802.15.4z standard. The main characteristics of UWB packets include the following: synchronization (SYNC), start of frame delimiter (SFD), physical layer header (PHR), physical layer (PHY) payload, and scrambled

This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) License.
For more information, see https://creativecommons.org/licenses/by/4.0/

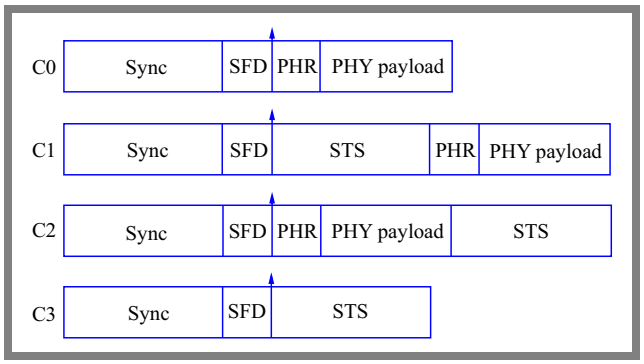

Fig. 1. UWB packet configurations.

timestamp sequence (STS). All of them are shown in Fig. 1. Configuration 0 (C0) does not include an STS inside the packet. In C1 and C2, only the STS location within the frame is changed, while C3 does not include PHR and PHY payload fields. In each configuration, the position of the ranging marker (RMARKER) [6] is indicated by an arrow.

2.1. Iterative Waveform Configuration and Generation Approach

To transmit such packets, the generator should support arbitrary or IQ waveform generation functionality. Typically, these instruments have the ability to save waveforms in their memory and then send them to the digital-to-analog converter (DAC) [7]. Exploiting this capability, one of the conventional methods used to generate a large number of UWB packets relies on creating a packet and downloading it into the built-in memory of a signal generator (SG) [8]. Then the SG generates the packet, repeating these steps in a loop until the required number of packets has been transmitted (Fig. 2).

The advantage of this method lies in minimal usage of SG resources, since an individual packet does not require a significant amount of memory. However, the method involves a time-consuming process of creating waveforms and writing them sequentially into the memory.

To cope with deterministic time intervals between the packets generated it is necessary to allow the receiver to process each frame before the arrival of the next frame. The waveform creation lead time for this method is considered to be idle time.

Nevertheless, because of its non-deterministic behavior, jitter exists in the interframe spacing between the generated packets. The other disadvantage of this method is that waveforms are not saved in the memory. Therefore, whenever it is necessary to generate the same waveforms to test different DUTs, the process of creating waveforms must be repeated, thereby increasing waveform generation lead time.

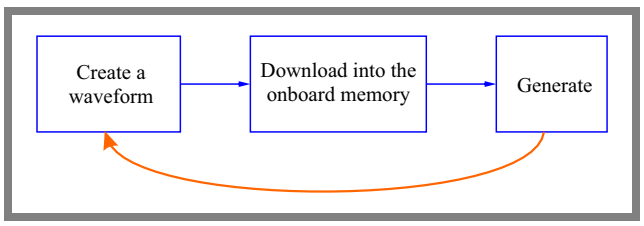

Fig. 2. Iterative waveform configuration and generation approach.

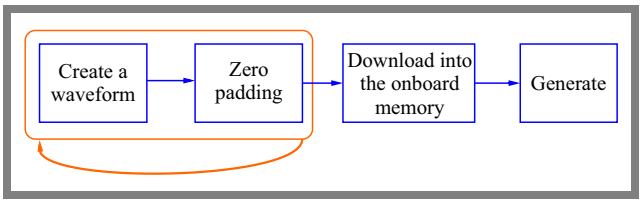

Fig. 3. Zero padding waveform generation approach.

2.2. Zero Padding Waveform Generation Approach

Another conventional waveform generation method addressing these disadvantages involves creating all the required waveforms by adding dummy samples (zeros) after each packet. These zeros do not contain any information and have no power, effectively representing the absence of a signal [9]. The number of zeros determines the interframe spacing time. Once multiple waveforms containing zeros are created, those can be downloaded into the onboard memory of SG, which then generates the entire waveform (Fig. 3).

This approach offers the possibility to define a deterministic interframe spacing time between the packets. However, it results in significant memory usage, as all packets and zeros are downloaded into the onboard memory at once. The equation for calculating the number of waveforms that can be downloaded into the onboard memory is:

$$N_W = \frac{M_{SG}}{M_{SP} + M_0}\,, \tag{1}$$

where N_W represents the number of waveforms, M_{SG} indicates the size of the onboard memory of SG, M_{SP} refers to the memory size of a single UWB packet, while M_0 represents the memory size of padded zeros.

In order to determine the $M_{SP} + M_0$ size, the number of samples N_s used to create the waveform should be considered:

$$N_s = t_w \cdot SR\,, \tag{2}$$

where t_w represents the length of the waveform and SR is the sample rate of SG. This article utilizes the PXIe-5831 with the maximum sample rate of 1.25 GS/s. t_w may vary depending on packet configuration. A UWB signal with the C2 packet format and a length of approximately 150 µs (Fig. 4) was created in the LabVIEW environment. The created packet, illustrated in Fig. 4, was used during the signal generation process.

1 ms interframe spacing is used to allow the receiving device to process each frame before the arrival of the next. With a 150 µs packet, the total length for one waveform would be 1.15 ms. As for numerator of Eq. (1), the typical value for M_{SG} is 2 GB. It may be applied in this scenario as well, as PXIe-5831 has an onboard memory of 2 GB [10]. These values are applied to Eq. (1):

$$N_W = \frac{2\text{ GB}}{1.15\text{ ms} \cdot 1.25\text{ GS/s} \cdot 4} = \frac{2\text{ GB}}{0.00575\text{ GB}} = 347.8\,, \tag{3}$$

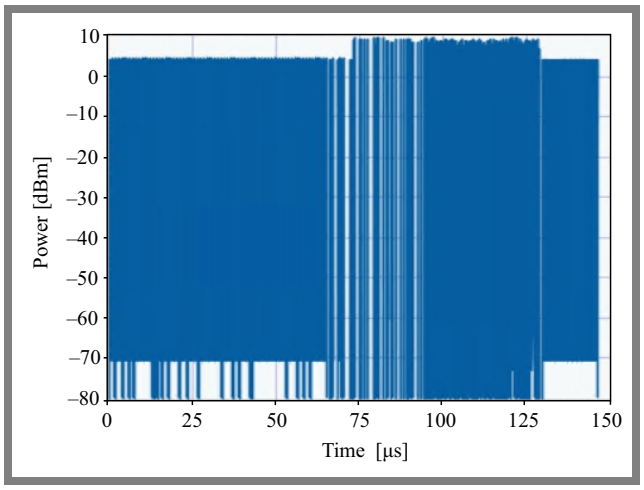

Fig. 4. Generated UWB packet with a length of approximately 150 µs.

where the coefficient of 4 used in the denominator represents the amount of memory allocated to each sample, as the I and Q DACs of the PXIe-5831 have the resolution of 16-bits, i.e. 2 bytes per DAC.

Equation (3) shows that the maximum number of waveforms that can be downloaded into the memory at once is limited to 347 only. This limitation indicates that whenever a greater number of waveforms is re required, this method cannot be utilized.

2.3. Memory-Efficient Interframe Spacing Waveform Generation Approach

The proposed method aims to address the limitations of the previous two approaches. It is similar to the approach shown in Fig. 3. However, instead of adding zeroes to define interframe spacing, a specified number of samples that do not consume memory and effectively represent the absence of a signal is generated.

An algorithm is proposed to generate such multiple waveforms downloaded into the onboard memory in a desired sequence. Between these waveforms, the number of samples required to properly represent the interframe spacing length is defined. An internal, highly deterministic counter algorithm is implemented in LabVIEW to precisely define the length of interframe spacing without memory allocation. The concept relies on integrating a delay function into the SG and counting every generated sample of the instrument. PXIe-5831 offers a maximum sample rate of 1.25 GS/s. Therefore, the minimum possible delay sample will be 0.8 ns, ensuring sufficient interframe spacing accuracy. To represent a 1 ms interframe spacing, the next packet should be delayed by 1 250 000 samples.

Figure 5 illustrates a single waveform, including the packet shown in Fig. 4, and 1 250 000 additional samples to represent 1 ms interframe spacing without consuming the onboard memory.

In this approach, zeros are not used to represent interframe spacing. Therefore, the number of waveforms that can be downloaded into the onboard memory is:

$$N_W = \frac{M_{SG}}{M_{SP}} = \frac{2\ \text{GB}}{0.15\ \text{ms} \cdot 1.25\ \text{GS/s} \cdot 4} = \frac{2\ \text{GB}}{0.00075\ \text{GB}} = 2666.6\ , \tag{4}$$

Equation (4) shows that up to 2666 waveforms can be stored in the instrument's onboard memory at once, effectively surpassing earlier limitation. This represents a significant improvement compared to the zero padding approach, where only 347 waveforms could be accommodated.

3. Results and Analysis

The PXIe-5831 vector signal transceiver is used to generate 2000 packets and to assess the two methods described in Section 2. Initially, the iterative waveform configuration and generation approach is tested, followed by an evaluation of the proposed memory-efficient interframe spacing waveform generation method. A comparative analysis of these methods is performed.

Figure 6 illustrates the test system used, with the PXIe-5831 instrument mounted inside the chassis together with the PXIe-8881 controller. PXIe-5831's generation port is connected to the analyzer port to evaluate the generated waveforms [11].

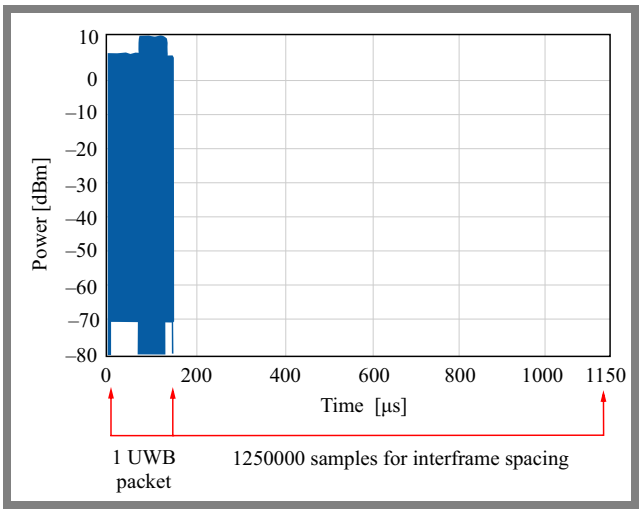

Fig. 5. Generated UWB packet and interframe spacing samples.

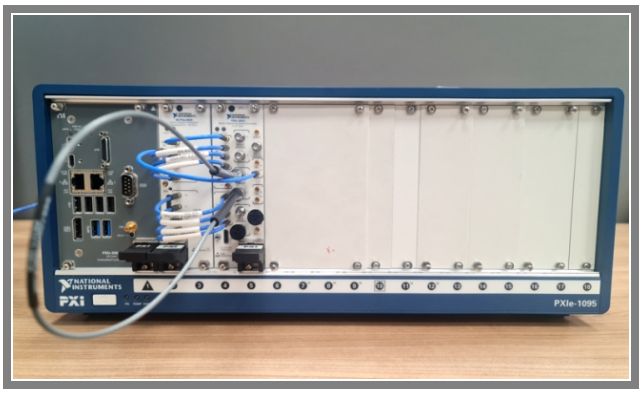

Fig. 6. Test system overview.

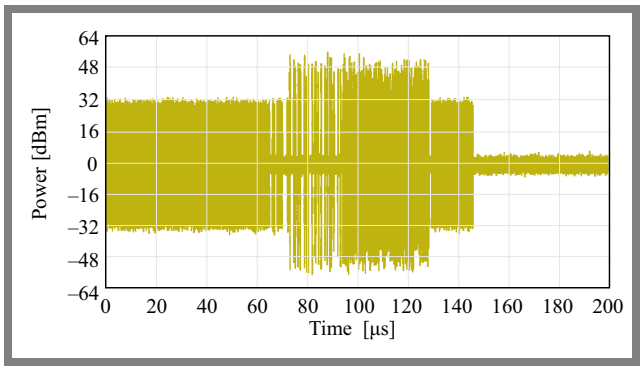

Fig. 7. First packet captured with the use of the iterative waveform configuration and generation approach.

2000 waveforms are used for both methods. The only parameter changed for each waveform is the payload information, which does not modify the waveform length, ensuring validity of the analysis. The algorithm for waveform generation is implemented in LabVIEW using the following UWB PHY parameters [12]:

- mean PRF = 62.4 MHz,
- PHR rate = 0.85 Mb/s,
- data rate = 6.81 Mb/s,
- packet configuration = SP2,
- code index = 9,
- preamble symbol repetitions = 64,
- symbols in SFD = 8.

UWB channel 9 with a carrier frequency of 7987.2 MHz [13] is used for measurements. The signal power level is configured to –10 dBm. The analyzer portion of the PXIe-5831 is configured to capture waveforms based on IQ power level, with a trigger level of –20 dBm.

3.1. Iterative Waveform Configuration – Performance Evaluation

To determine the time required for generating 2000 waveforms with this approach, two timestamps are captured: the first at the start of the initial packet configuration process, and the other after generation of the 2000th packet has been completed. The first captured packet with a length of approximately 150 µs is shown in Fig. 7.

Although interframe spacing was set to 1 ms, the subsequent waveform cannot be observed within this period due to the configuration process. This takes approximately 480 ms and its duration is unstable, adding uncertainty to the interframe spacing length. This duration is measured by recording timestamps at the start and completion of the waveform configuration process. Figure 8 shows, in a histogram, the Gaussian distribution of configuration times measured for each generated waveform, with a standard deviation of $\sigma = 0.01$ and a mean value of 0.48 s.

To assess the generation lead time required on multiple occasions, which is essential for multi-DUT testing, the mathematical representation of generating 2000 waveforms will be:

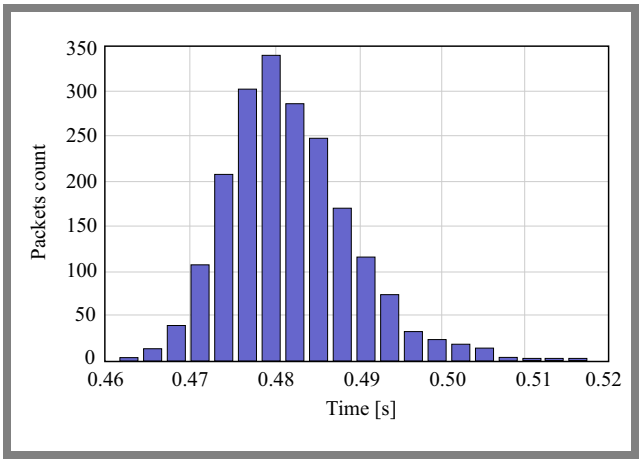

Fig. 8. Gaussian distribution of 2000 configuration times measured.

$$T = N \cdot 2000 \cdot \left[t_{cfg}(\sigma_{cfg} = 0.01) + t_{gen}(\sigma_{gen} \approx 0) \right] , \quad (5)$$

where N represents the number of DUTs to be tested, t_{cfg} is the waveform configuration time, t_{gen} is the packet generation time, and T is the total generation time for N DUTs.

The standard deviation for the total generation time σ_T depends on the standard deviation of the configuration time σ_{cfg}, as the packet generation time is deterministic ($\sigma_{gen} \approx 0$). Assuming that 100 DUTs need to be tested, the total time required for the waveform generation phase is:

$$T = 100 \cdot 2000 \cdot (480 \text{ ms} + 1.15 \text{ ms}) = 96230 \text{ s} . \quad (6)$$

As $\sigma_T = \sigma_{cfg}$ from Eq. (5), 96230 s ≈ 26.73 h and will vary by ±0.27 h.

While this method successfully generates 2000 waveforms, it suffers from a drawback consisting in its inability to accurately define interframe spacing. Additionally, regenerating the same 2000 waveforms requires that they be configured again, as the waveforms are not stored in the instrument's onboard memory. Consequently, the process of generating these 2000 waveforms 100 times would take over 26 hours. It is obvious that the zero padding waveform generation approach, where only 347 waveforms can fit into the onboard memory, will suffer from the same drawbacks.

3.2. Memory-Efficient Interframe Spacing Waveform – Performance Evaluation

This method focuses on two key phases of the process: configuration and generation. Since all the waveforms can be saved in SG's onboard memory, configuration is required only once. Thereafter, the waveforms may be generated as needed. Consequently, in multi-DUT testing scenarios, the 2000 waveforms need to be configured once only. To determine the time required for configuring 2000 waveforms (UWB packets with interframe spacing samples), timestamps are recorded at the start of the process of configuring the first packet and at the end of the process of configuring the last (2000th) packet. The difference between these timestamps represents the time required for the configuration phase. To calculate the gen-

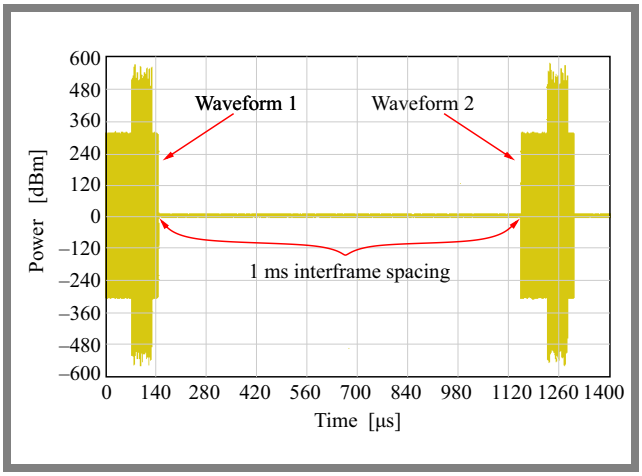

Fig. 9. Time between two subsequent waveforms captured with the use of the memory-efficient interframe spacing waveform generation approach.

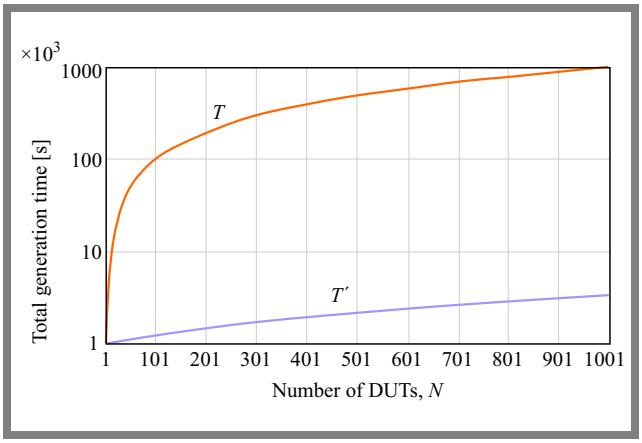

Fig. 10. Relationship between T and T'.

eration time, two timestamps are captured at the end of the configuration phase and after completing the generation of the last (2000th) packet.

The first and the second waveforms captured are illustrated in Fig. 9, showing that the second waveform appears within the next 1 ms, which represents the defined interframe spacing length. Unlike other methods, this approach does not involve a configuration step between the generated waveforms. Therefore, by adding delay samples between the packets, it becomes possible to define the interframe spacing with nanosecond accuracy.

As the waveform configuration phase takes place only once, Eq. (5) should be modified as follows:

$$T' = 2000 \cdot \left[t_{cfg(\sigma_{cfg}=0.01)} + N \cdot t_{gen(\sigma_{gen}\approx 0)}\right] . \quad (7)$$

Assuming that the same number of 100 DUTs needs to be tested, the total time required for the waveform generation is:

$$T' = 2000 \cdot \left[480\ \text{ms} + 100 \cdot 1.15\ \text{ms}\right] = 1190\ \text{s} . \quad (8)$$

As $\sigma_T = \sigma_{cfg}$ from Eq. (7), 1190 s ≈ 0.33 h and will vary by ±0.003 h.

From Eqs. (6) and (8) it is clear that $T' \ll T$, showing that the proposed method optimizes the time required for generation large numbers of packets. Furthermore, the difference between T and T' increases when the number of tested DUTs goes up, as illustrated in Fig. 10.

4. Conclusion

An investigation of test time optimization techniques for IEEE 802.15.4z UWB devices was performed, focusing on measurements requiring a large number of packet generations. Conventional UWB packet generation methods were analyzed, exhibiting such drawbacks as time-consuming waveform configuration processes, difficulties with accurate definition of interframe spacing, and significant memory requirements.

To address these challenges, a novel memory-efficient generation technique was proposed. This method utilizes a specified number of delay samples between packets to define interframe spacing, eliminating the need for zero padding. This has been achieved by integrating a delay function into the signal generator, thus allowing to define the interframe spacing with an accuracy of 0.8 ns.

By leveraging National Instruments PXIe-5831 VST and LabVIEW software environment, an algorithm has been implemented showcasing significant improvements in waveform generation time and memory utilization. Calculations demonstrated that it is feasible to store up to 2666 UWB packets in the generator's onboard memory, which is a significant improvement compared to the conventional approaches, where only 347 packets could be stored. Also, the waveform generation lead time has been drastically reduced to 0.33 hours, compared to 26.73 hours when testing 100 DUTs. Experimental results confirmed that the proposed method offers the flexibility of configuring 2000 waveforms only once, making it highly suitable for multi-DUT testing scenarios.

References

[1] G. Tsaturyan, S. Antonyan, and L. Movsisyan, "Testing of Integrated Circuit Operating with Ultra-wideband Technology", *Scientific Proceedings of Vanadzor State University*, 2022.

[2] M. Stocker *et al.*, "On the Performance of IEEE 802.15.4z-Compliant Ultra-wideband Devices", *Workshop on Benchmarking Cyber-Physical Systems and Internet of Things (CPS-IoTBench)*, Milan, Italy, 2022 (https://doi.org/10.1109/CPS-IoTBench56135.2022.00012).

[3] P. Sedlacek, P. Masek, and M. Slanina, "An Overview of the IEEE 802.15.4z Standard and its Comparison to the Existing UWB Standards", *29th International Conference Radioelektronika*, Pardubice, Czech Republic, 2019 (https://doi.org/10.1109/RADIOELEK.2019.8733537).

[4] R. Yang and R.S. Sherratt, "Multiband OFDM Modulation and Demodulation for Ultra Wideband Communications", in: *Novel Applications of the UWB Technologies*, 2011 (https://doi.org/10.5772/16700).

[5] Rohde & Schwarz, "Generation of IEEE 802.15.4 Signals", Application Note, 2016 (https://www.rohde-schwarz.com/us/appli

cations/generation-of-ieee-802-15-4-signals-application-note_56280-95360.html).

[6] IEEE, "P802.15.4z/D06, Mar2020 - IEEE Draft Standard for Low-Rate Wireless Networks Amendment: Enhanced High Rate Pulse (HRP) and Low Rate Pulse (LRP) Ultra Wide-band (UWB) Physical Layers (PHYs) and Associated Ranging Techniques", 2020 (ISBN: 9781504465335).

[7] Keysight, "Keysight Fundamentals of Arbitrary Waveform Generation", 2015 (https://www.keysight.com/us/en/assets/9018-03815/reference-guides/9018-03815.pdf).

[8] National Instruments, "NI-FGEN User Manual", 2024 (https://www.ni.com/docs/en-US/bundle/ni-fgen/page/user-manual-welcome.html).

[9] M. Łuczyński, A. Dobrucki, and S. Brachmański, "Active Tone Elimination Algorithm Using FFT with Interpolation and Zero-padding", *Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)*, Poznan, Poland, 2020 (https://doi.org/10.23919/SPA50552.2020.9241255).

[10] National Instruments, "PXIe-5820 Specifications", 2024 (https://www.ni.com/docs/en-US/bundle/pxie-5820-specs/page/specs.html).

[11] National Instruments, "PXIe-5831 Specifications", 2024 (https://www.ni.com/docs/en-US/bundle/pxie-5831-specs/page/specs.html).

[12] IEEE, "802.15.4-2020 - IEEE Standard for Low-rate Wireless Networks", 2020 (https://doi.org/10.1109/IEEESTD.2020.9144691).

[13] R. Juran *et al.*, "Hands-on Experience with UWB: Angle of Arrival Accuracy Evaluation in Channel 9", *2022 14th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, Valencia, Spain, 2022 (https://doi.org/10.1109/ICUMT57764.2022.9943504).

Grigor Tsaturyan, M.Sc.
Institute of Information and Telecommunication Technologies and Electronics
https://orcid.org/0009-0000-4465-7176
E-mail: grigor.tsaturyan99@gmail.com
National Polytechnic University of Armenia, Yerevan, Armenia
https://polytech.am/en/home/

Hovhannes Gomtsyan, Ph.D.
Institute of Information and Telecommunication Technologies and Electronics
https://orcid.org/0009-0003-2479-7923
E-mail: hovhannes.gomcyan@polytechnic.am
National Polytechnic University of Armenia, Yerevan, Armenia
https://polytech.am/en/home/
European University of Armenia, Yerevan, Armenia
https://www.eua.am

Performance Analysis of Antenna-relay Selection in CNOMA Systems under Practical Impairments

Menaa Saber and Khelil Abdellatif

Echahid Hamma Lakhdar University, El-Oued, Algeria

https://doi.org/10.26636/jtit.2024.3.1589

Abstract — Selection strategies prove to be a valuable approach in mitigating complexity associated with antenna selection (AS) and relay selection (RS), optimizing signal transmission through a streamlined number of antennas/relays, and enhancing overall system performance. This paper offers a comprehensive analysis, deriving closed-form expressions for the outage probability (OP) and throughput in proposed scenarios that leverage the best relay selection (BRS) and transmit antenna selection (TAS) protocol for cooperative non-orthogonal multiple access (CNOMA), along with partial relay selection (PRS) and TAS protocol for CNOMA. The study extends to Rayleigh fading channels, considering practical impairments such as successful interference cancellation (SIC) error, channel estimation error (CEE), and feedback delay error. In comparing the proposed system to conventional CNOMA, our findings highlight the substantial impact of SIC, CEE, and feedback delay imperfections on the performance of both proposed scenarios. Notably, the application of BRS-based TAS protocol outperforms PRS-based TAS in terms of OP and throughput. The close alignment between analytical, asymptotic, and simulation results attests to the credibility of conducted analysis.

Keywords — channel estimation error, CNOMA, outage probability, relay selection, transmit antenna selection

1. Introduction

Non-orthogonal multiple access (NOMA) positions itself as a transformative technology with the potential to shape the landscape of B5G and 6G networks due to its capacity to significantly boost network capacity by optimizing the utilization of scarce spectrum resources [1], [2].

In power-domain NOMA, multiple users are served on the same time-frequency resources [3]. For that, downlink information is efficiently broadcasted using superposition coding (SC) at the base station (BS). Subsequently, successive interference cancellation (SIC) is employed at the users to eliminate multiuser interferences [4].

Recently, there has been a significant gain in interest regarding the utilization of NOMA in cooperative communication for 5G and beyond deployment scenarios [5]. This increased attention is driven by the inherent advantages of cooperative relaying strategies, which include expanded coverage, improved reliability, and effective mitigation of challenges arising from multipath propagation [6].

In relay communication, two widely recognized protocols are employed: the decode-and-forward (DF), in which the relay decodes and re-encodes the information signal before forwarding, and the amplify-and-forward (AF), where the relay amplifies the received signal from the source and transmits it directly to the destination [7].

To enhance the performance of the NOMA system, some studies have explored the combined utilization of NOMA and cooperative communication. In [8], the authors examined a downlink NOMA system with cooperative full-duplex (FD) relaying, employing the near user as a relay for the far user. They derived closed-form expressions for the outage probability (OP) and ergodic sum rate (ESR), considering both fixed power allocations and optimal power allocations aimed at minimizing the OP, as well, the OP performance in a downlink CNOMA with an AF relay [9].

The authors in [10], developed novel closed-form expressions for the bit error rates (BER) of full-duplex cooperative NOMA (FD-CNOMA) systems in the presence of imperfect SIC and residual self-interference (RSI). The authors of [11] have investigated a hybrid relaying scheme, that switches between full and half duplex (FD/HD) to improve the performance of the DF CNOMA system for two users and multi-users [12].

Additionally, the investigation in [13] focuses on evaluating the OP and ESR of an FD/HD coordinated direct and DF-based relay scheme in a NOMA system. This analysis specifically considers scenarios with imperfections in CSI and SIC.

To enhance the coverage and performance for the far user, a proposed scheme in [14] involves utilizing a multi-hop DF relay-aided NOMA (MH-DF-R-NOMA). This scheme includes deriving closed-form expressions for BER and OP in the presence of imperfect SIC and CSI.

While, CNOMA offers benefits like increased spectral efficiency and broader coverage, it also presents challenges such as complexity, interference management, varying channel conditions, limited coverage, power allocation, synchronization issues, and inaccuracies in channel estimation. Addressing these drawbacks is imperative to fully harness the potential of CNOMA in the context of 5G and 6G networks.

To address the drawbacks of CNOMA, relay selection (RS) presents numerous advantages over conventional CNOMA

This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) License.
For more information, see https://creativecommons.org/licenses/by/4.0/

systems. These include heightened reliability, expanded coverage, enhanced throughput, flexible network deployment options, and improved energy efficiency. The advantages position relay selection as a valuable technique for enhancing the performance and efficacy of 5G and 6G systems.

The impact of RS strategy on the performance of cooperative NOMA is examined to reach the minimal OP across all possible RS schemes and meet the quality of service (QoS) criteria for both users [15]. The authors in [16], delves into the performance analysis of OP and the sum rate of NOMA schemes within AF relay systems considering partial relay selection (PRS). Furthermore, regarding the best relay selection (BRS) considered in [17], the authors analyzed the OP of the optimal RS schemes for cooperative downlink NOMA, considering both fixed and adaptive power allocations (PAs) at the relays, respectively. The OP performance of a dual-hop multi-relay NOMA system using a DF scheme over Nakagami-m fading channels is investigated in [18].

In [19], the authors analyzed the secrecy outage performance for a NOMA network assisted by multiple relays operating over Nakagami-m fading channels, utilizing relay selection. The authors in [20], investigate the analysis of BER performance for both AF and DF relay selection in cooperative NOMA systems over Rayleigh fading channels. Indeed, the AF-assisted max-min relay selection method exhibits superior performance for the far user compared to the DF relay. The implementation of BRS in the downlink scenario of NOMA-based cognitive relay networks (NCRNs) is investigated under the assumption of Rayleigh fading channels in [21].

Beyond the contributions of cooperative communications to enhancing NOMA system performance, researchers suggest the implementation of multiple-input multiple-output (MIMO) techniques to ensure these improvements. Considering that MIMO techniques entail higher complexity and power consumption, the strategy of antenna selection (AS) is deemed a robust option to ensure the desired performance [22].

According to AS techniques, three main protocols are widely known, namely: transmit antenna selection (TAS) [23], receive antenna selection (RAS) [24], and joint transmit and receive antenna selection (JTRAS) [25]. The authors in [26] investigated three NOMA downlink models, including a single-input-single-output (SISO) scenario featuring a single antenna, a multi-input-single-output (MISO) scenario, and a MIMO scenario. The TAS protocol was employed in all scenarios. The authors conducted a comparison of the OP and system throughput, considering all users over Rayleigh fading channels.

In [27], the researchers introduced AS for an FD-CNOMA system to maximize the end-to-end signal-to-interference-plus-noise ratio (SINR) for both the near and far users. This involved employing TAS/RAS protocols at the relay. The performance was characterized in terms of OP and ESR. A comprehensive analysis of the OP performance for two hybrid AS schemes, namely TAS and maximal ratio combining (MRC) at the first hop, as well as JTRAS at the second hop, in a MIMO-NOMA based downlink AF relaying network. The analysis considers the impact of channel estimation error (CEE) and feedback delay over Nakagami-m fading channels [28].

In addition, new combined antenna selection strategies, such as max-max-max and max-min-max, have been developed in [29]. Nevertheless, the max-max-max and max-min-max schemes have the common goal of enhancing the stronger and weaker user, respectively. Furthermore, the impact of hardware impairments (HWI) on the uplink SIMO CNOMA (SIMO-CNOMA) using BRS under SIC and CSI imperfections is examined in terms of OP and system throughput in [30].

Therefore, employing the TAS protocol as a MIMO technique in conventional CNOMA systems yields numerous advantages. These encompass heightened diversity, augmented spectral efficiency, enhanced throughput, reduced interference, and enhanced deployment flexibility.

Drawing from the aforementioned literature, CNOMA studies have predominantly focused on single-relay scenarios, often overlooking factors like imperfect SIC, CEE, and feedback delay. Moreover, investigations rarely delve into the joint application of RS and TAS within CNOMA frameworks. Consequently, the integration of RS and TAS protocols into CNOMA systems while addressing practical challenges such as SIC, CEE, and feedback delay remains largely unexplored. This paper fills this gap by evaluating and analyzing the performance of CNOMA systems using schemes that combine best relay selection with TAS (BRS-TAS CNOMA) and partial relay selection with TAS (PRS-TAS CNOMA), with a focus on OP and throughput performance. Therefore, the primary contributions of this study are outlined as follows:

- We investigate a realistic downlink CNOMA scheme influenced by SIC, CEE, and feedback delay. The study includes different scenarios with multiple relays and antennas, employing the RS technique and the integration of the TAS protocol to improve performance. The transmitting base station (BS) and relays are equipped with multiple antennas, while single antennas are employed at the reception.
- Exact integral expressions for the OP and throughput of all considered scenarios are derived, considering SIC, CEE, and feedback delay imperfections. These expressions are validated through both asymptotic analysis using the

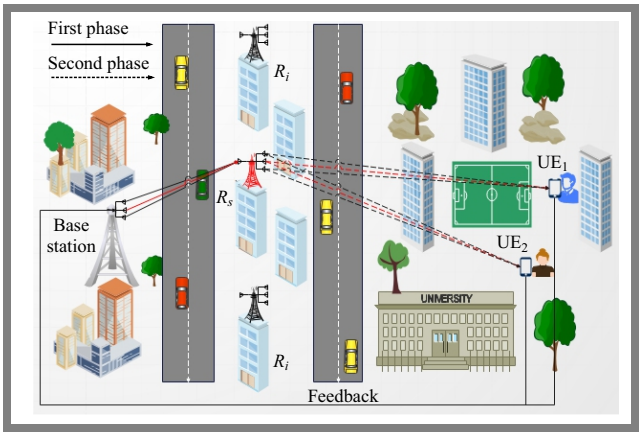

Fig. 1. The concept of system model.

McLaurin series expansion and simulations carried out with computer simulation.

- The impact of key system parameters on the OP and throughput performance is evaluated based on the obtained analytical expressions. These parameters include the number of relays and antennas, the distance between the relay and the base station, the power allocation factor, and practical impairments.
- A performance comparison between the analyzed schemes and their HD relay-aided NOMA system counterpart is provided.
- Finally, numerical results show that the systems under study in the high SNR region are limited by practical constraints, as indicated by the presence of an error floor (EF). Increasing the number of antennas and relays helps reduce these effects.

The rest of the paper is organized as follows: In Section 2, we delve into the MISO-NOMA system utilizing the TAS protocol, exploring both PRS-TAS CNOMA and BRS-TAS CNOMA scenarios while considering imperfections like SIC error, CEE, and feedback delay, taking conventional CNOMA as a benchmark. Section 3 provides a detailed derivation of closed-form expressions for OP in the investigated system. Numerical and simulation results are then presented in Section 4. The paper concludes with Section 5, summarizing the key findings and insights.

2. System Model

Here, we examine the downlink CNOMA system comprising a base station, N relays, and two users: a far user UE_1 and a near user UE_2. As shown in Fig. 1, the BS and relays are equipped with N_t transmit antennas, while each user and relay use a single receiving antenna, resulting in a two-hop communication setup using the TAS scheme. So, the TAS scheme is applied in the two hops, ensuring that N_t at both the BS and relay offers the highest SNR, respectively.

To represent this system under the influence of imperfect SIC and imperfect CSI including CEE and feedback delay, first some notations and definitions are introduced.

$h_{j.SR_i}$ and $h_{j.R_iD}$ are the $N_t \times 1$ fading channel coefficients vector of both hops between the BS–R and R–UE_j, respectively, modeled as a complex Gaussian random variable:

$$h_{j.SR_i} \sim CN\left(0, \sigma^2_{j.I}\right)$$

and

$$h_{j.R_iD} \sim CN\left(0, \sigma^2_{j.II}\right) \ .$$

The variances are defined as:

$$\sigma^2_{j.SR_i} = E\left\{|h_{j.SR_i}|^2\right\}$$

and

$$\sigma^2_{j.R_iD} = E\left\{|h_{j.R_iD}|^2\right\} \ ,$$

with $j = 1, 2$.

$\hat{h}_{j.SR_i}$ and $\hat{h}_{j.R_iD}$ are the $N_t \times 1$ estimated channel coefficients vector of the two-hops can be expressed by:

$$\hat{h}_{j.SR_i} = h_{j.SR_i} - \kappa_{j.I}$$

and

$$\hat{h}_{j.R_iD} = h_{j.R_iD} - \kappa_{j.II} \, ,$$

where $\kappa_{j.I}$ and $\kappa_{j.II}$ are the CEE independent of $\hat{h}_{j.SR_i}$ and $\hat{h}_{j.R_iD}$ and modelled as:

$$\kappa_{j.I} \sim CN\left(0, \sigma^2_{\kappa_{j.I}}\right) \ ,$$

$$\kappa_{j.II} \sim CN\left(0, \sigma^2_{\kappa_{j.II}}\right) \ .$$

The variances are defined as:

$$\sigma^2_{\kappa_{j.I}} = \sigma^2_{j.SR_i} - \widehat{\sigma}^2_{j.SR_i} \ , \ \sigma^2_{\kappa_{j.II}} = \sigma^2_{j.R_iD} - \widehat{\sigma}^2_{j.R_iD},$$

where:

$\widehat{\sigma}^2_{j.SR_i} = E\left\{\left|\widehat{h}_{j.SR_i}\right|^2\right\}$ and $\widehat{\sigma}^2_{j.R_iD} = E\left\{\left|\widehat{h}_{j.R_iD}\right|^2\right\}$ are the variances of $\widehat{h}_{j.SR_i}$ and $\widehat{h}_{j.R_iD}$, respectively, with $j = 1, 2$.

$\widehat{h}^{(\tau)}_{j.SR_i}$ and $\widehat{h}^{(\tau)}_{j.R_iD}$ are the $N_t \times 1$ feedback delayed of the estimated channels vector of $\hat{h}_{j.SR_i}$ and $\hat{h}_{j.R_iD}$ for the two-hops are defined as:

$\hat{h}_{j.SR_i} = \rho\widehat{h}^{(\tau)}_{j.SR_i} + e_{fd_{j.I}}$ and $\hat{h}_{j.R_iD} = \rho\widehat{h}^{(\tau)}_{j.R_iD} + e_{fd_{j.II}}$, where $e_{fd_{j.I}}$ and $e_{fd_{j.II}}$ are the feedback error of the two-hops modelled as:

$e_{fd_{j.I}} \sim CN\left(0, \sigma^2_{fd_{j.I}}\right)$ and $e_{fd_{j.II}} \sim CN\left(0, \sigma^2_{fd_{j.II}}\right)$, which:

$$\sigma^2_{fd_{j.I}} = (1 - \rho_{j.I})\,\widehat{\sigma}^2_{j.SR_i} \, ,$$

$$\sigma^2_{fd_{j.II}} = (1 - \rho_{j.II})\,\widehat{\sigma}^2_{j.R_iD} \, ,$$

$$\rho = J_0(2\pi f_d \tau), \ \ 0 < \rho < 1 \, ,$$

ρ and $f_d\tau$ represents the time correlation coefficient and the normalized Doppler frequency, respectively.

We assume $e_{j.fd}$ and κ_j of the two hops are independent random variables, we define a new error for both hops:

$e_{j.I} = \kappa_{j.I} + e_{fd_{j.I}}$ and $e_{j.II} = \kappa_{j.II} + e_{fd_{j.II}}$,

then

$e_{j.I} \sim CN\left(0, \sigma^2_{e_{j.I}}\right)$ and $e_{j.II} \sim CN\left(0, \sigma^2_{e_{j.II}}\right)$,

where $\sigma^2_{e_{j.I}} = \sigma^2_{\kappa_{j.I}} + \sigma^2_{fd_{j.I}}$ and $\sigma^2_{e_{j.II}} = \sigma^2_{\kappa_{j.II}} + \sigma^2_{fd_{j.II}}$ are the variances of each element in the error term $e_{j.I}$ and $e_{j.II}$, respectively.

To simplify the calculation, we make the assumption:

$$\sigma^2_{\kappa_I} = \sigma^2_{\kappa_{j.I}} \ , \ \sigma^2_{\kappa_{II}} = \sigma^2_{\kappa_{j.II}} \, ,$$

$$\sigma^2_{fd_I} = \sigma^2_{fd_{j.I}} \text{ and } \sigma^2_{fd_{II}} = \sigma^2_{fd_{j.II}} \, .$$

Applying the TAS scheme in the two hops, a transmit antenna, denoted as $\widetilde{t}$, is selected at both the BS and relay. This selection is based on having the maximum sum of squared channel gains between the receiver antennas of the relay and the users. So, the selected transmit antenna criteria for the two-hop is given by:

$$\widetilde{t} = \underset{1<t<N_t}{\arg\max} |h_t|^2. \tag{1}$$

In the best relay selection (BRS) scheme, the optimal relay is selected based on maximizing the minimum SINR between

the source-to-relay link ($S - R_l$) and the relay-to-UE$_j$ link. This selection criteria is expressed as:

$$R_s = \underset{1<l<L}{\arg\max} \left\{ \min \left(\gamma_{SR_i}, \gamma_{R_i D} \right) \right\}. \tag{2}$$

A partial relay selection (PRS) scheme is performed for one hop only, wherein the relay is chosen based on the CSI of each hop. The selection criteria is given as:

$$R_s = \underset{1<l<L}{\arg\max} \left\{ \left(\gamma_{SR_i} \right) \right\}, \tag{3}$$

and

$$R_s = \underset{1<l<L}{\arg\max} \left\{ \left(\gamma_{R_i D} \right) \right\}. \tag{4}$$

In the first hop of communication, the BS transmits a superimposed signal $x = \sum_{j=1}^{2} \sqrt{P_t \alpha_j} x_j$ to the relay, where x_j are the messages of UE$_j$, P_t denotes the BS transmit power and the coefficients α_j satisfy the conditions:

$\sum_{j=1}^{2} \alpha_j = 1$ and $\alpha_1 > \alpha_2$ leading to $|h_{1.SR_i}|^2 < |h_{2.SR_i}|^2$.

The received signal by the relay is given as:

$$y_r = \left(\rho \widehat{\widetilde{h}}^{(\tau)}_{j.SR_i} + e_I \right) \sum_{j=1}^{2} \sqrt{P_t \alpha_j} x_j + n_r, \tag{5}$$

where n_r is the additive white Gaussian noise (AWGN) $n_r \sim CN\left(0, N_o\right)$.

The received SINRs for x_1 and x_2 at the relay are given, respectively by:

$$\gamma_1^{SR_i} = \frac{\alpha_1 \gamma_0 \left| \widehat{\widetilde{h}}^{(\tau)}_{1.SR_i} \right|^2}{\alpha_2 \gamma_0 \left| \widehat{\widetilde{h}}^{(\tau)}_{1.SR_i} \right|^2 + \frac{\gamma_0}{\rho_I^2} \left(\sigma_{\kappa_I}^2 + \sigma_{fd_I}^2 \right) + \frac{1}{\rho_I^2}}, \tag{6}$$

and

$$\gamma_2^{SR_i} = \frac{\alpha_2 \gamma_0 \left| \widehat{\widetilde{h}}^{(\tau)}_{2.SR_i} \right|^2}{\alpha_1 \gamma_0 \xi \left| \widehat{\widetilde{h}}^{(\tau)}_{2.SR_i} \right|^2 + \frac{\gamma_0}{\rho_I^2} \left(\sigma_{\kappa_I}^2 + \sigma_{fd_I}^2 \right) + \frac{1}{\rho_I^2}}, \tag{7}$$

where ξ denotes the residual SIC factor and $\gamma_0 = \frac{P_t}{N_0}$ is the transmit SNR.

In the second hop of communication, the RS transmits a superimposed signal $x = \sum_{j=1}^{2} \sqrt{P_r \alpha_j} x_j$ to the users, where x_j are the messages of UE$_j$, P_r denotes the relay transmit power and the coefficients α_j satisfy the conditions:

$$\sum_{j=1}^{2} \alpha_j = 1, \quad \alpha_1 > \alpha_2$$

leading to

$$|h_{1.R_i D}|^2 < |h_{2.R_i D}|^2 .$$

The signal received by the UE$_j$ users can be expressed as follows:

$$y_j = \left(\rho \widehat{\widetilde{h}}^{(\tau)}_{j.R_i D} + e_{II} \right) \sum_{j=1}^{2} \sqrt{P_r \alpha_j} x_j + n_j \,, \tag{8}$$

where P_r denotes the relay transmit power, n_j is the additive white Gaussian noise (AWGN) $n_j \sim CN\left(0, N_o\right)$.

The received SINRs for x_1 and x_2 at UE$_j$ are defined as:

$$\gamma_1^{R_i D} = \frac{\alpha_1 \gamma_0 \left| \widehat{\widetilde{h}}^{(\tau)}_{1.R_i D} \right|^2}{\alpha_2 \gamma_0 \left| \widehat{\widetilde{h}}^{(\tau)}_{1.R_i D} \right|^2 + \frac{\gamma_0}{\rho_{II}^2} \left(\sigma_{\kappa_{II}}^2 + \sigma_{fd_{II}}^2 \right) + \frac{1}{\rho_{II}^2}}, \tag{9}$$

$$\gamma_{2\to 1}^{R_i D} = \frac{\alpha_1 \gamma_0 \left| \widehat{\widetilde{h}}^{(\tau)}_{2.R_i D} \right|^2}{\alpha_2 \gamma_0 \left| \widehat{\widetilde{h}}^{(\tau)}_{2.R_i D} \right|^2 + \frac{\gamma_0}{\rho_{II}^2} \left(\sigma_{\kappa_{II}}^2 + \sigma_{fd_{II}}^2 \right) + \frac{1}{\rho_{II}^2}}, \tag{10}$$

and

$$\gamma_2^{R_i D} = \frac{\alpha_2 \gamma_0 \left| \widehat{\widetilde{h}}^{(\tau)}_{2.R_i D} \right|^2}{\alpha_1 \gamma_0 \xi \left| \widehat{\widetilde{h}}^{(\tau)}_{2.R_i D} \right|^2 + \frac{\gamma_0}{\rho_{II}^2} \left(\sigma_{\kappa_{II}}^2 + \sigma_{fd_{II}}^2 \right) + \frac{1}{\rho_{II}^2}}, \tag{11}$$

where $\gamma_0 = \frac{P_r}{N_0}$ is the transmit SNR.

3. Performance Analysis

In this section, we derive the OP and throughput expressions of the CNOMA based RS and TAS protocol in the presence of practical impairments, then SIC, CEE, and feedback delay over the Rayleigh fading channel.

3.1. Outage Probability Analysis at Far User

The OP of the CNOMA can be obtained as:

$$P_{e2e.j} = 1 - \left(1 - P_{out.j.I}\right)\left(1 - P_{out.j.II}\right), \tag{12}$$

where $P_{out.j.I}$ and $P_{out.j.II}$ are the OP for the first and second hops, respectively, for each user.

The end-2-end (e2e) OP of UE$_1$ is expressed as follows:

$$P_{e2e.1} = 1 - \left(1 - \Pr\left(\gamma_1^{SR} < \gamma_{th,1}\right)\right)\left(1 - \Pr\left(\gamma_1^{RD} < \gamma_{th,1}\right)\right). \tag{13}$$

In context of best relay selection, the e2e OP for the far user in CNOMA considering both BRS and TAS schemes is expressed as:

$$\begin{aligned} P_{e2e.1}^{BRS} \;=\; & 1 - \prod_{i=1}^{N} \left(1 - \prod_{t=1}^{N_t} \Pr\left(\gamma_1^{SR_i} < \gamma_{th,1}\right) \right) \\ & \times \prod_{i=1}^{N} \left(1 - \prod_{t=1}^{N_t} \Pr\left(\gamma_1^{R_i D} < \gamma_{th,1}\right) \right). \end{aligned} \tag{14}$$

The OP for the first and second hops for UE$_1$ can be calculated as in [31]:

$$\prod_{t=1}^{N_t} \Pr\left(\gamma_1^{SR_i} < \gamma_{th,1}\right) = \prod_{t=1}^{N_t} \left(1 - \exp \frac{\lambda_1^{SR_i}}{\sigma_{1.SR_i}^2} \right), \tag{15}$$

and

$$\prod_{t=1}^{N_t} \Pr\left(\gamma_1^{R_i D} < \gamma_{th,1}\right) = \prod_{t=1}^{N_t} \left(1 - \exp \frac{\lambda_1^{R_i D}}{\sigma_{1.R_i D}^2}\right), \quad (16)$$

where:

$$\lambda_1^{SR_i} = \frac{\frac{\gamma_{th,1}}{\rho_I^2}\left(1+\gamma_0\left(\sigma_{\kappa_I}^2+\sigma_{fd_I}^2\right)\right)}{\gamma_0\left(\alpha_1-\alpha_2\gamma_{th,1}\right)},$$
$$\lambda_1^{R_i D} = \frac{\frac{\gamma_{th,1}}{\rho_I^2}\left(1+\gamma_0\left(\sigma_{\kappa_{II}}^2+\sigma_{fd_{II}}^2\right)\right)}{\gamma_0\left(\alpha_1-\alpha_2\gamma_{th,1}\right)}.$$

Finally, by substituting Eqs. (15) and (16) into (14), the end-to-end OP of the far user can be expressed as:

$$P_{e2e.1}^{BRS} = 1 - \prod_{i=1}^{N}\left(1 - \prod_{t=1}^{N_t}\left(1 - \exp \frac{\lambda_1^{SR_i}}{\sigma_{1.r}^2}\right)\right) \times \prod_{i=1}^{N}\left(1 - \prod_{t=1}^{N_t}\left(1 - \exp \frac{\lambda_1^{R_i D}}{\sigma_1^2}\right)\right). \quad (17)$$

In context of partial relay selection, the e2e OP for the far user in CNOMA incorporating both PRS at first hop and TAS schemes is formulated as:

$$P_{e2e.1}^{PRS} = 1 - \left(1 - \prod_{i=1}^{N}\prod_{t=1}^{N_t} \Pr\left(\gamma_1^{SR_i} < \gamma_{th.1}\right)\right) \times \left(1 - \prod_{t=1}^{N_t} \Pr\left(\gamma_1^{R_s D} < \gamma_{th.1}\right)\right). \quad (18)$$

By substituting $\lambda_1^{SR_i}$ and $\lambda_1^{R_s D}$ into Eq. (18), the end-to-end OP for the far user, considering TAS and PRS, can be expressed as:

$$P_{e2e.1}^{PRS} = 1 - \left(1 - \prod_{i=1}^{N}\prod_{t=1}^{N_t}\left(1 - \exp \frac{\lambda_1^{SR_i}}{\sigma_{1.SR_i}^2}\right)\right) \times \left(1 - \prod_{t=1}^{N_t}\left(1 - \exp \frac{\lambda_1^{R_s D}}{\sigma_{1.R_s D}^2}\right)\right). \quad (19)$$

3.2. Outage Probability Analysis at Near User

We recall that the e2e OP of CNOMA for the near user is:

$$P_{e2e.2} = 1 - \left(1 - \Pr\left(\gamma_2^{SR} < \gamma_{th,2}\right)\right)\left(1 - \Pr\left(\gamma_2^{RD} < \gamma_{th,2}\right)\right). \quad (20)$$

In the best relay selection, the e2e OP for the near user in CNOMA considering both BRS and TAS schemes is expressed as:

$$P_{e2e.2}^{BRS} = \prod_{i=1}^{N}\left(1 - \left(1 - \underbrace{\prod_{t=1}^{N_t} P_{out.2}^{SR_i}}_{term\ I}\right)\left(1 - \underbrace{\prod_{t=1}^{N_t} P_{out.2}^{R_i D}}_{term\ II}\right)\right), \quad (21)$$

In this context, terms I and II denote the OP of the near user utilizing the TAS protocol in the first and second hops, respectively. Therefore, the OP of the term I for UE_2, can be expressed as:

$$\prod_{t=1}^{N_t} P_{out.2}^{SR_i} = \prod_{t=1}^{N_t} P\left(\gamma_{2\to1}^{SR_i} < \gamma_{th.1}\right) + \left[\prod_{t=1}^{N_t} P\left(\gamma_{2\to1}^{SR_i} \geqslant \gamma_{th.1}\right)\prod_{t=1}^{N_t} P\left(\gamma_2^{SR_i} < \gamma_{th.2}\right)\right]. \quad (22)$$

Each probability condition is calculated as:

$$\prod_{t=1}^{N_t} P\left(\gamma_{2\to1}^{SR_i} < \gamma_{th.1}\right) = \prod_{t=1}^{N_t}\left(1 - \exp -\frac{\lambda_2^{SR_i}}{\sigma_{2,SR_i}^2}\right), \quad (23)$$

and

$$\prod_{t=1}^{N_t} P\left(\gamma_2^{SR_i} < \gamma_{th.2}\right) = \prod_{t=1}^{N_t}\left(1 - \exp -\frac{\lambda_3^{SR_i}}{\sigma_{2,SR_i}^2}\right), \quad (24)$$

where:

$$\lambda_2^{SR_i} = \frac{\frac{\gamma_{th_1}}{\rho_I^2}\left(1+\gamma_0\left(\sigma_{\kappa_I}^2+\sigma_{fd_I}^2\right)\right)}{\gamma_0\left(\alpha_1-\alpha_2\gamma_{th_1}\right)},$$
$$\lambda_3^{SR_i} = \frac{\frac{\gamma_{th_2}}{\rho_I^2}\left(1+\gamma_0\left(\sigma_{\kappa_I}^2+\sigma_{fd_I}^2\right)\right)}{\gamma_0\left(\alpha_2-\alpha_1\xi\gamma_{th_2}\right)}.$$

After substituting Eqs. (23) and (24) into (22), the term I can be written as:

$$\prod_{t=1}^{N_t} P_{out.2}^{SR_i} = \prod_{t=1}^{N_t}\left(1 - \exp -\frac{\lambda_2^{SR_i}}{\sigma_{2,SR_i}^2}\right) + \left(1 - \prod_{t=1}^{N_t}\left(1 - \exp -\frac{\lambda_2^{SR_i}}{\sigma_{2,SR_i}^2}\right)\right) \times \prod_{t=1}^{N_t}\left(1 - \exp -\frac{\lambda_3^{SR_i}}{\sigma_{2,SR_i}^2}\right). \quad (25)$$

Similarly to the first hop, the OP of the second hop for UE_2, can be expressed as follows:

$$\prod_{t=1}^{N_t} P_{out.2}^{R_i D} = \prod_{t=1}^{N_t} P\left(\gamma_{2\to1}^{R_i D} < \gamma_{th.1}\right) + \left[\prod_{t=1}^{N_t} P\left(\gamma_{2\to1}^{R_i D} \geqslant \gamma_{th.1}\right)\prod_{t=1}^{N_t} P\left(\gamma_2^{R_i D} < \gamma_{th.2}\right)\right]. \quad (26)$$

The calculation of each probability condition is as follows:

$$\prod_{t=1}^{N_t} P\left(\gamma_{2\to1}^{R_i D} < \gamma_{th.1}\right) = \prod_{t=1}^{N_t}\left(1 - \exp -\frac{\lambda_2^{R_i D}}{\sigma_{2,R_i D}^2}\right), \quad (27)$$

$$\prod_{t=1}^{N_t} P\left(\gamma_2^{R_i D} < \gamma_{th.2}\right) = \prod_{t=1}^{N_t}\left(1 - \exp -\frac{\lambda_3^{R_i D}}{\sigma_{2,R_i D}^2}\right), \quad (28)$$

where:

$$\lambda_2^{R_i D} = \frac{\frac{\gamma_{th_1}}{\rho_{II}^2}\left(1+\gamma_0\left(\sigma_{\kappa_{II}}^2+\sigma_{fd_{II}}^2\right)\right)}{\gamma_0\left(\alpha_1-\alpha_2\gamma_{th_1}\right)},$$
$$\lambda_3^{R_i D} = \frac{\frac{\gamma_{th_2}}{\rho_{II}^2}\left(1+\gamma_0\left(\sigma_{\kappa_{II}}^2+\sigma_{fd_{II}}^2\right)\right)}{\gamma_0\left(\alpha_2-\alpha_1\xi\gamma_{th_2}\right)}.$$

Now, by substituting Eqs. (27) and (28) into (26), the term II can be obtained as:

$$\begin{aligned}\prod_{t=1}^{N_t} P_{out.2}^{R_iD} &= \prod_{t=1}^{N_t}\left(1-\exp-\frac{\lambda_2^{R_iD}}{\sigma^2_{2,R_iD}}\right)\\ &+\left(1-\prod_{t=1}^{N_t}\left(1-\exp-\frac{\lambda_2^{R_iD}}{\sigma^2_{2,R_iD}}\right)\right)\\ &\times\prod_{t=1}^{N_t}\left(1-\exp-\frac{\lambda_3^{R_iD}}{\sigma^2_{2,R_iD}}\right).\end{aligned} \tag{29}$$

Finally, the e2e OP of the UE_2 using TAS and BRS is obtained by substituting Eqs. (25) and (29) into Eq. (21).

In line with the proofs presented previously, for partial relay selection the e2e OP for the near user in CNOMA using both PRS and TAS schemes is obtained as:

$$P_{e2e.2}^{PRS} = 1-\left(1-\prod_{i=1}^{N}\prod_{t=1}^{N_t}P_{out.2}^{SR_i}\right)\left(1-\prod_{t=1}^{N_t}P_{out.2}^{R_sD}\right), \tag{30}$$

where:

$$\begin{aligned}\prod_{t=1}^{N_t} P_{out.2}^{R_sD} &= \prod_{t=1}^{N_t} P\left(\gamma_{2\to1}^{R_sD}<\gamma_{th.1}\right)\\ &+\left[\prod_{t=1}^{N_t} P\left(\gamma_{2\to1}^{R_sD}\geqslant\gamma_{th.1}\right)\prod_{t=1}^{N_t} P\left(\gamma_{2}^{R_sD}<\gamma_{th.2}\right)\right].\end{aligned} \tag{31}$$

Following the same mathematical proofs:

$$\begin{aligned}\prod_{t=1}^{N_t} P_{out.2}^{R_sD} &= \prod_{t=1}^{N_t}\left(1-\exp-\frac{\lambda_2^{R_sD}}{\sigma^2_{2,R_sD}}\right)\\ &+\left(1-\prod_{t=1}^{N_t}\left(1-\exp-\frac{\lambda_2^{R_sD}}{\sigma^2_{2,R_sD}}\right)\right)\\ &\times\prod_{t=1}^{N_t}\left(1-\exp-\frac{\lambda_3^{R_sD}}{\sigma^2_{2,R_sD}}\right).\end{aligned} \tag{32}$$

Finally, the e2e OP of the UE_2 using TAS and PRS is obtained by substituting Eqs. (25) and (32) into Eq. (30).

3.3. Asymptotic Outage Probability

In this subsection, we analyze the performance of asymptotic OP in scenarios marked by high SNR, aiming to gain a more profound understanding of the considered situations. Hence, the asymptotic OP is defined in conditions of high SNR (i.e. as $\gamma_0\to\infty$), employing the McLaurin series expansion [32] $e^{-x}\approx 1-x$.

The asymptotic expression for the OP of far user of UE_1, at high SNR regime is given for BRS and PRS scenarios, respectively, as:

$$\begin{aligned}P_{e2e.1}^{BRS,asy} &\approx 1-\prod_{i=1}^{N}\left(1-\prod_{t=1}^{N_t}\frac{\lambda_1^{SR_i}}{\sigma^2_{1.SR_i}}\right)\\ &\times\prod_{i=1}^{N}\left(1-\prod_{t=1}^{N_t}\frac{\lambda_1^{R_iD}}{\sigma^2_{1.R_iD}}\right)\end{aligned} \tag{33}$$

and

$$\begin{aligned}P_{e2e.1}^{PRS,asy} &\approx 1-\left(1-\prod_{i=1}^{N}\prod_{t=1}^{N_t}\frac{\lambda_1^{SR_i}}{\sigma^2_{1.SR_i}}\right)\\ &\times\left(1-\prod_{t=1}^{N_t}\frac{\lambda_1^{R_sD}}{\sigma^2_{1.R_sD}}\right).\end{aligned} \tag{34}$$

Similarly, the asymptotic expression for the OP of near user of UE_2, at high SNR regime is given for BRS and PRS scenarios, respectively, as follows:

$$\begin{aligned}\prod_{t=1}^{N_t} P_{out.2}^{SR_i,asy} &\approx \prod_{t=1}^{N_t}\frac{\lambda_2^{SR_i}}{\sigma^2_{2,SR_i}}\\ &+\left[\left(1-\prod_{t=1}^{N_t}\frac{\lambda_2^{SR_i}}{\sigma^2_{2,SR_i}}\right)\times\prod_{t=1}^{N_t}\frac{\lambda_3^{SR_i}}{\sigma^2_{2,SR_i}}\right],\end{aligned} \tag{35}$$

$$\begin{aligned}\prod_{t=1}^{N_t} P_{out.2}^{R_iD,asy} &\approx \prod_{t=1}^{N_t}\frac{\lambda_2^{R_iD}}{\sigma^2_{2,R_iD}}\\ &+\left[\left(1-\prod_{t=1}^{N_t}\frac{\lambda_2^{R_iD}}{\sigma^2_{2,R_iD}}\right)\times\prod_{t=1}^{N_t}\frac{\lambda_3^{R_iD}}{\sigma^2_{2,R_iD}}\right],\end{aligned} \tag{36}$$

and

$$\begin{aligned}\prod_{t=1}^{N_t} P_{out.2}^{R_sD,asy} &\approx \prod_{t=1}^{N_t}\frac{\lambda_2^{R_sD}}{\sigma^2_{2,R_sD}}\\ &+\left[\left(1-\prod_{t=1}^{N_t}\frac{\lambda_2^{R_sD}}{\sigma^2_{2,R_sD}}\right)\times\prod_{t=1}^{N_t}\frac{\lambda_3^{R_sD}}{\sigma^2_{2,R_sD}}\right].\end{aligned} \tag{37}$$

3.4. System Throughput Analysis

The system throughput is the sum of the achievable received bit rates at UE_1 and UE_2. As a result, the system throughput can be calculated as:

$$Tp_{sys} = Tp_1+Tp_2 = (1-P_{out.1})\,R_1^*+(1-P_{out.2})\,R_2^*, \tag{38}$$

where R_1^* and R_2^* are the threshold rates of UE_1 and UE_2, respectively.

4. Numerical Results

In this section, we present validation of the analysis for CNOMA scenarios provided in the previous sections.

The practical impairments, SIC, CEE, and feedback delay, are taken into account to evaluate OP and throughput over the Rayleigh fading channels. With power coefficients allocated as $\alpha_1 = 0.8$ and $\alpha_2 = 0.2$, the distance between the BS and relays set to $dsr = 1$ m and distances between the users and relays set to $dr_1 = 2$ m and $dr_2 = 1$ m.

Additionally, we have $\xi = 0.01$, $\sigma_\kappa^2 = 0.01$, and $f_d\tau = 0.02$. In the plots, different lines and markers represent analytical, simulation, and asymptotic curves. The figures demonstrate a close alignment among the analytical, asymptotic, and simulated results, providing strong confirmation of the validity of the performance analysis.

The analytical and simulation results in Figs. 2-5 compare multiple scenarios inside the downlink CNOMA system, examining the OP of UE_1 and UE_2, separately. These scenarios include the use of AS and RS to CNOMA while considering

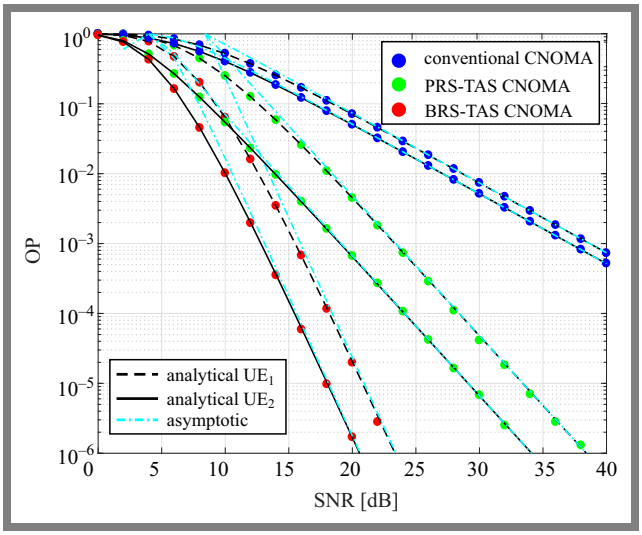

Fig. 2. OP of the downlink CNOMA scenarios under ideal conditions.

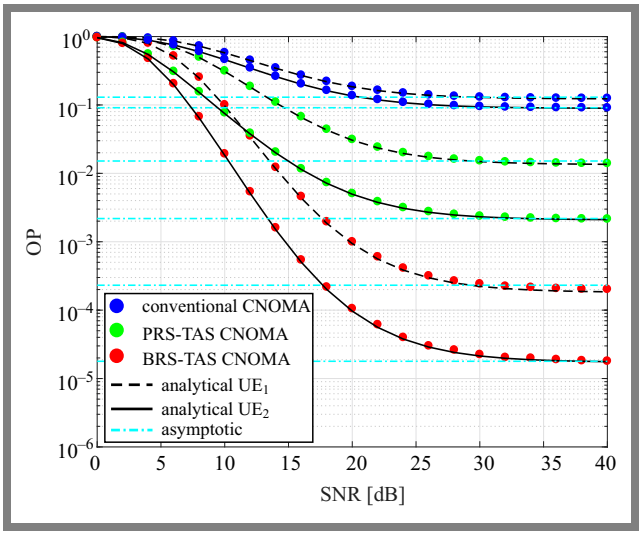

Fig. 3. OP of the downlink CNOMA scenarios under impaired conditions.

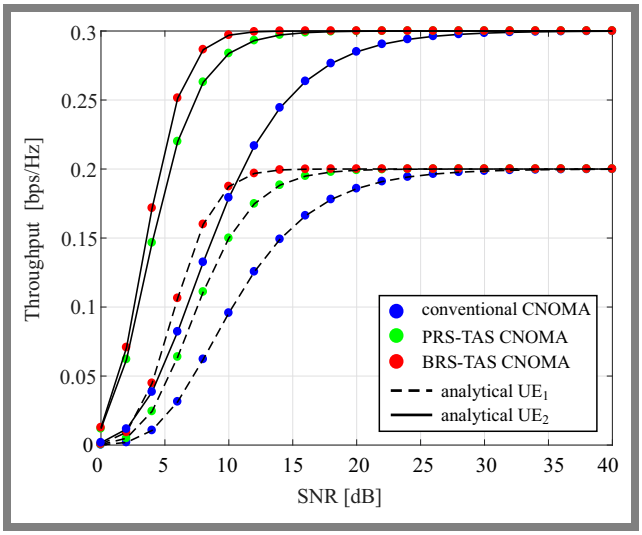

Fig. 4. Throughput of the downlink CNOMA scenarios under ideal conditions.

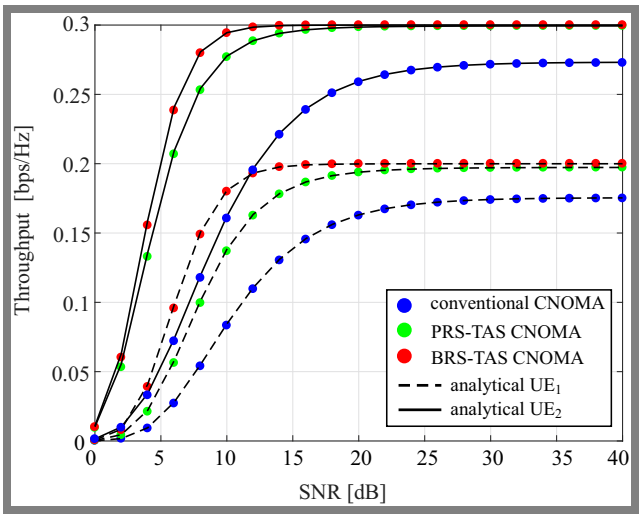

Fig. 5. Throughput of the downlink CNOMA scenarios under impaired conditions.

the presence and absence of certain impairments: $\xi = 0.01$, $\sigma_\kappa^2 = 0.01$, and $f_d\tau = 0.02$.

Figures 2 and 3 depict the OP outcomes at UE_1 and UE_2 for various scenarios, including conventional CNOMA using a single antenna, and one relay as a benchmark, and TAS integration into CNOMA-based PRS and BRS. Through analysis and simulation, it becomes evident that the performance of the CNOMA system gradually improves with the addition of more RS.

Moreover, employing the TAS protocol at both the BS and relays yields better performance than relying only on conventional CNOMA. This underscores the significance of antenna configurations at the BS and relays for enhancing CNOMA system performance.

However, as depicted in Fig. 3, there is a decline in performance attributed to the presence of imperfections in SIC, CEE and feedback delay, leading to an error floor at high SNR.

Also, Figs. 2 and 3 provide insights into the OP of UE_1, and UE_2. Notably, the CNOMA-based BRS-TAS scenario demonstrates superior performance compared to other scenarios. This enhancement is achieved by incorporating TAS at both the BS and relays, confirming the critical role of leveraging multiple antennas for improved users performance.

In addition, the asymptotic results closely align with the analytical findings, validating the accuracy and reliability of our analysis.

In Figs. 4 and 5, we present the throughput performance comparison of the studied scenarios versus SNR, considering the presence and absence of imperfections in SIC, CEE, and feedback delay. As we can see, the PRS-TAS CNOMA outperforms conventional CNOMA. While the BRS-TAS CNOMA has a positive impact compared to counterparts PRS-TAS CNOMA.

The simulation plots are well-matched with the analytical results, validating the accuracy of our analysis.

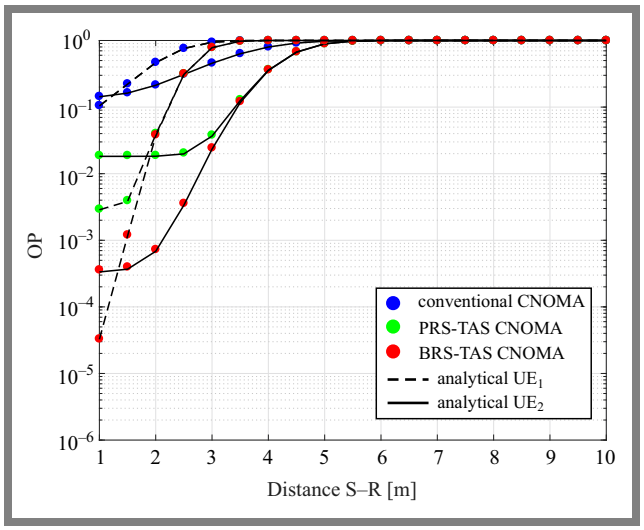

Fig. 6. The impact of distance on the OP of the downlink CNOMA under various impairments.

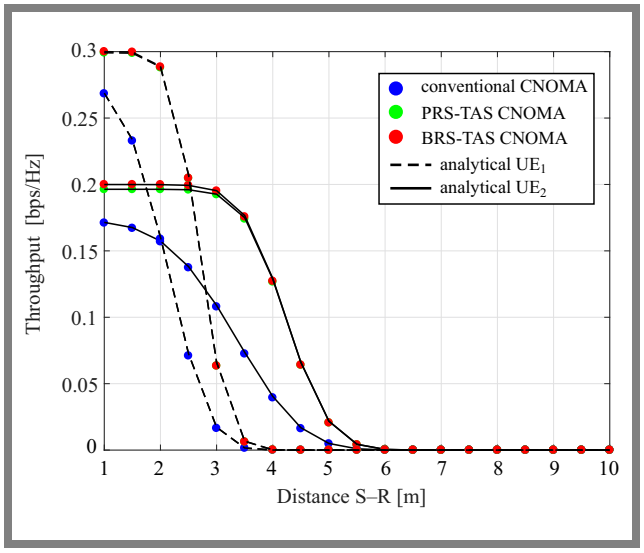

Fig. 7. The impact of distance on the throughput of the downlink CNOMA under various impairments.

Figures 6 and 7, depict the OP and throughput of the CNOMA system relative to the distance dr, assuming an SNR of 25 dB. Specifically, we compare the performance of CNOMA with BRS-TAS against conventional CNOMA.

Noteworthy observations lead to several conclusions. At $dr = 1$ m, UE_2 exhibits superior OP performance compared to the far user UE_1. However, at distances $dr = 2$ m and beyond, the performance of UE_1 surpasses that of UE_2. Furthermore, as the distance increases, UE_2 experiences outage more rapidly than UE_1.

The assessment of the influence of power allocation α_1 on CNOMA is conducted by examining the OP and throughput in Figs. 8 and 9, respectively. With a fixed SNR of 25 dB, it is evident that the performance of UE_1 improves proportionally with the increased power allocated to it.

However, this improvement comes at the expense of deteriorating performance for UE_2.

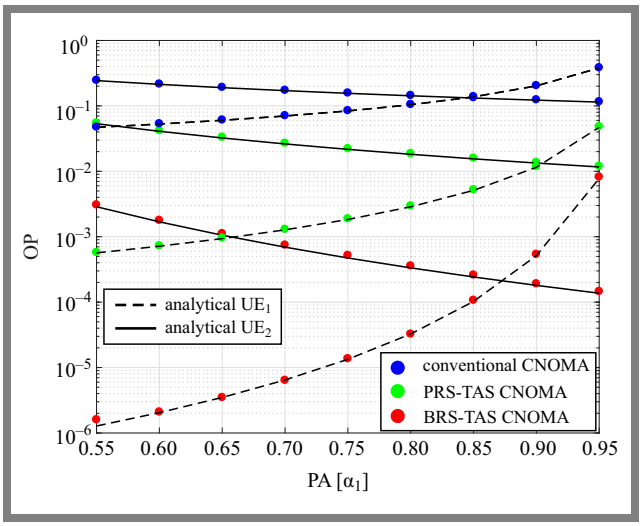

Fig. 8. The impact of PA on the OP of the downlink CNOMA under various impairments.

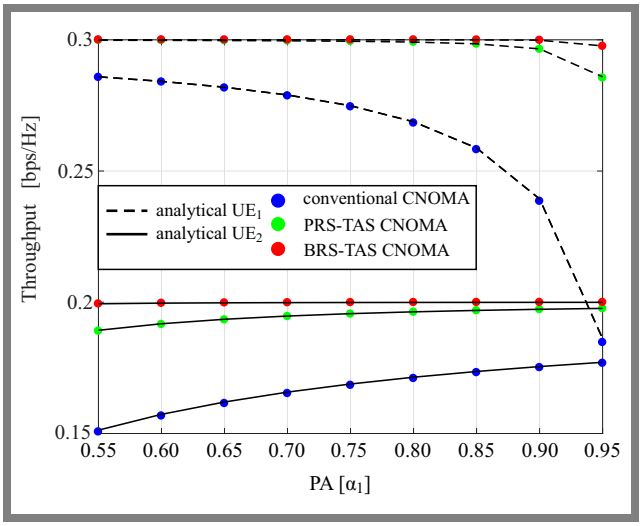

Fig. 9. The impact of PA on the throughput of the downlink CNOMA under various impairments.

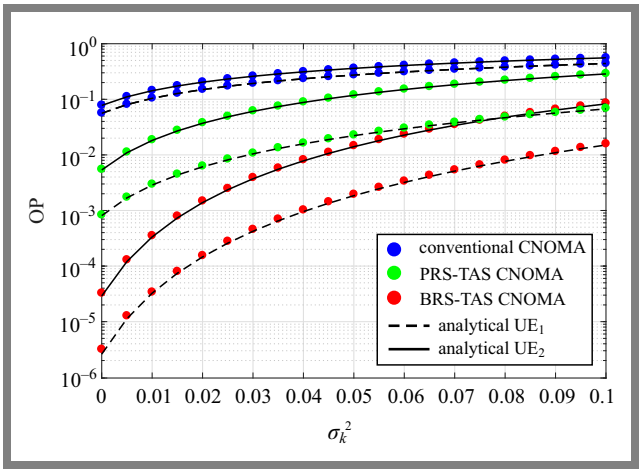

Fig. 10. The impact of CEE on the OP of the downlink CNOMA under various impairments.

Furthermore, it is notable that, across varying values of α_1, the CNOMA-based BRS-TAS consistently ensures superior performance in both OP and throughput.

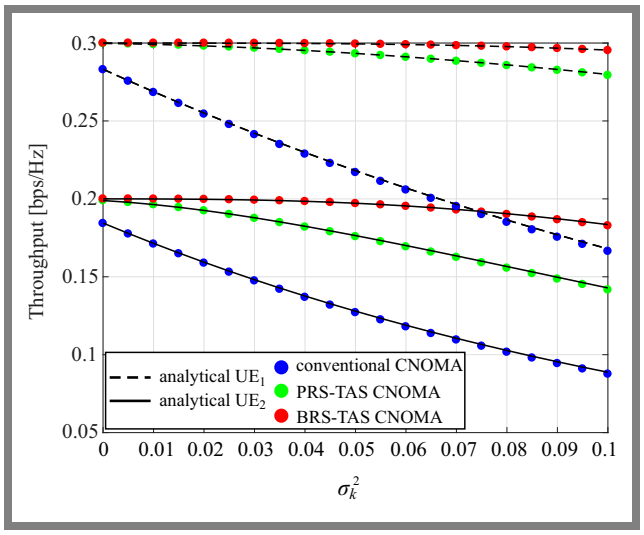

Fig. 11. The impact of CEE on the throughput of the downlink CNOMA under various impairments.

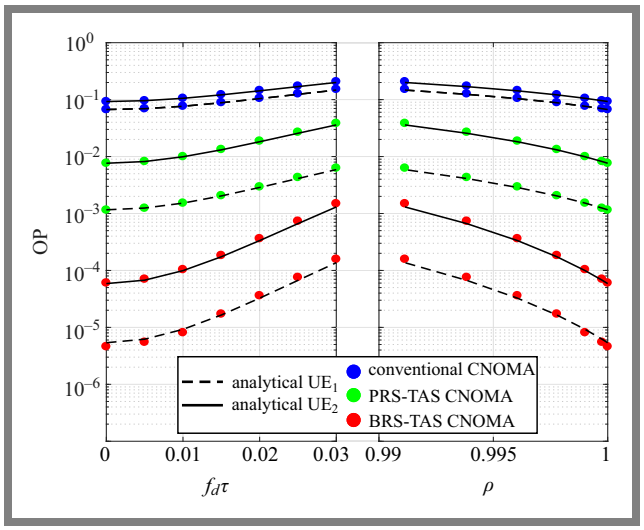

Fig. 12. The impact of ρ and $f_d\tau$ on the OP of the downlink CNOMA under various impairments.

The channel estimation error σ_κ^2 exerts a detrimental impact on the OP and throughput across all CNOMA scenarios. This adverse effect becomes more pronounced with an escalation in the degree of the error, as illustrated in Figs. 10 and 11.

Figures 12 and 13, depict the influence of feedback delay and time correlation coefficient ρ on the OP and throughput of the CNOMA system. Maintaining a constant SNR at 25 dB, it is evident that an increase in feedback delay results in a degradation of the system's performance, both in terms of OP and throughput. Conversely, higher values of ρ contribute to an enhancement in performance.

5. Conclusions

In conclusion, this paper systematically explored and compared CNOMA with TAS against the conventional CNOMA approach. The investigation considered both BRS and PRS paradigm selection among multiple relays, resulting in the

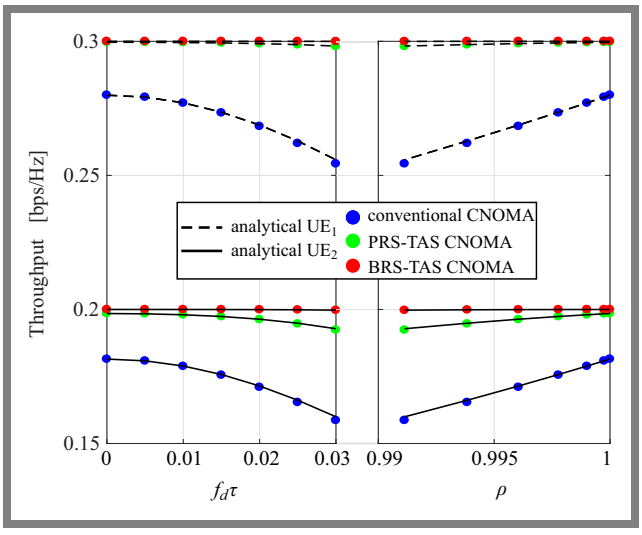

Fig. 13. The impact of ρ and $f_d\tau$ on the throughput of the downlink CNOMA under various impairments.

formulation of BRS-TAS CNOMA and PRS-TAS CNOMA schemes. Practical impairments such as SIC error, CEE, and feedback delay were comprehensively considered in both scenarios. Exact expressions for OP and throughput over Rayleigh fading channels were derived and validated through simulations and asymptotic proofs.

The findings underscored that augmenting the number of antennas at the BS or relays significantly improves the overall performance of CNOMA schemes. Furthermore, the incorporation of a relay selection paradigm yielded additional enhancements in system performance. Importantly, the adverse effects of impairments were alleviated through the strategic application of selection criteria for antennas and/or relays.

In summary, the alignment between analytical, asymptotic, and simulated results establishes the robustness and accuracy of the analytical framework presented in this study. These insights contribute to advancing the understanding of selection strategies in CNOMA systems, with implications for optimizing performance in practical communication scenarios.

References

[1] H. Tullberg *et al.*, "The METIS 5G System Concept: Meeting the 5G Requirements", *IEEE Communications Magazine*, vol. 54, no. 12, pp. 132–139, 2016 (https://doi.org/10.1109/MCOM.2016.1500799CM).

[2] W. Shin *et al.*, "Non-orthogonal Multiple Access in Multi-cell Networks: Theory, Performance, and Practical Challenges", *IEEE Communications Magazine*, vol. 55, no. 10, pp. 176–183, 2017 (https://doi.org/10.1109/MCOM.2017.1601065).

[3] B. Selim *et al.*, "Radio-frequency Front-end Impairments: Performance Degradation in Non-orthogonal Multiple Access Communication Systems", *IEEE Vehicular Technology Magazine*, vol. 14, no. 1, pp. 89–97, 2019 (https://doi.org/10.1109/MVT.2018.2867646).

[4] Z. Yang, Z. Ding, P. Fan, and N. Al-Dhahir, "A General Power Allocation Scheme to Guarantee Quality of Service in Downlink and Uplink NOMA Systems", *IEEE Transactions on Wireless Communications*, vol. 15, no. 11, pp. 7244–7257, 2016 (https://doi.org/10.1109/TWC.2016.2599521).

[5] H. Liu *et al.*, "Decode-and-forward Relaying for Cooperative NOMA Systems with Direct Links", *IEEE Transactions on Wireless Communications*, vol. 17, no. 12, pp. 8077–8093, 2018 (https://doi.org/10.1109/TWC.2018.2873999).

[6] G. Li, D. Mishra, and H. Jiang, "Cooperative NOMA with Incremental Relaying: Performance Analysis and Optimization", *IEEE Transactions on Vehicular Technology*, vol. 67, no. 11, pp. 11291–11295, 2018. (https://doi.org/10.1109/TVT.2018.2869531).

[7] A. Tregancini, *et al.*, "Performance Analysis of Full-duplex Relay-aided NOMA Systems Using Partial Relay Selection", *IEEE Transactions on Vehicular Technology*, vol. 69, no. 1, pp. 622–635, 2019 (https://doi.org/10.1109/TVT.2019.2952526).

[8] L. Zhang *et al.*, "Performance Analysis and Optimization in Downlink NOMA Systems with Cooperative Full-duplex Relaying", *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 10, pp. 2398–2412, 2017 (https://doi.org/10.1109/JSAC.2017.2724678).

[9] X. Liang *et al.*, "Outage Performance for Cooperative NOMA Transmission with an AF Relay", *IEEE Communications Letters*, vol. 21, no. 11, pp. 2428–2431, 2017 (https://doi.org/10.1109/LCOMM.2017.2681661).

[10] A.A. Hamza, I. Dayoub, I. Alouani, and A. Amrouche, "On the Error Rate Performance of Full-duplex Cooperative NOMA in Wireless Networks", *IEEE Transactions on Communications*, vol. 70, no. 3, pp. 1742–1758, 2021 (https://doi.org/10.1109/TCOMM.2021.3138079).

[11] X. Yue *et al.*, "Exploiting Full/half-duplex User Relaying in NOMA Systems", *IEEE Transactions on Communications*, vol. 66, no. 2, pp. 560–575, 2017 (https://doi.org/10.1109/TCOMM.2017.2749400).

[12] N. Guo, J. Ge, Q. Bu, and C. Zhang, "Multi-user Cooperative Non-orthogonal Multiple Access Scheme with Hybrid Full/half-duplex User-assisted Relaying", *IEEE Access*, vol. 7, pp. 39207–39226, 2019 (https://doi.org/10.1109/ACCESS.2019.2906382).

[13] V. Aswathi and A. Babu, "Full/half Duplex Cooperative NOMA under Imperfect Successive Interference Cancellation and Channel State Estimation Errors", *IEEE Access*, vol. 7, pp. 179961–179984, 2019 (https://doi.org/10.1109/ACCESS.2019.2959001).

[14] F. Khennoufa, K. Abdellatif, and F. Kara, "Bit Error Rate and Outage Probability Analysis for Multi-hop Decode-and-forward Relay-aided NOMA with Imperfect SIC and Imperfect CSI", *AEU-International Journal of Electronics and Communications*, vol. 147, art. no. 154124, 2022 (https://doi.org/10.1016/j.aeue.2022.154124).

[15] Y. Li *et al.*, "Performance Analysis of Relay Selection in Cooperative NOMA Networks", *IEEE Communications Letters*, vol. 23, no. 4, pp. 760–763, 2019 (https://doi.org/10.1109/LCOMM.2019.2898409).

[16] S. Lee *et al.*, "Non-orthogonal Multiple Access Schemes with Partial Relay Selection", *IET Communications*, vol. 11, no. 6, pp. 846–854, 2017 (https://doi.org/10.1049/iet-com.2016.0836).

[17] J. Ju *et al.*, "Performance Analysis for Cooperative NOMA with Opportunistic Relay Selection", *IEEE Access*, vol. 7, pp. 131488–131500, 2019 (https://doi.org/10.1109/ACCESS.2019.2940969).

[18] T.-T.T. Nguyen *et al.*, "New Look on Relay Selection Strategies for Full-duplex Multiple-relay NOMA over Nakagami-m Fading Channels", *Wireless Networks*, vol. 27, no. 2, pp. 3827–3843, 2021 (https://doi.org/10.1007/s11276-021-02676-1).

[19] H. Lei *et al.*, "Secrecy Outage Analysis for Cooperative NOMA Systems with Relay Selection Schemes", *IEEE Transactions on Communications*, vol. 67, no. 9, pp. 6282–6298, 2019 (https://doi.org/10.1109/TCOMM.2019.2916070).

[20] I. Umakoglu *et al.*, "BER Performance Comparison of AF and DF Assisted Relay Selection Schemes in Cooperative NOMA Systems", *2021 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom)*, Bucharest, Romania, 2021 (https://doi.org/10.1109/BlackSeaCom52164.2021.9527771).

[21] K. Sultan, "Best Relay Selection Schemes for NOMA Based Cognitive Relay Networks in Underlay Spectrum Sharing", *IEEE Access*, vol. 8, pp. 190160–190172, 2020 (https://doi.org/10.1109/ACCESS.2020.3031631).

[22] S. Sanayei and A. Nosratinia, "Antenna Selection in MIMO Systems", *IEEE Communications Magazine*, vol. 42, no. 10, pp. 68–73, 2004 (https://doi.org/10.1109/MCOM.2004.1341263).

[23] A.P. Shrestha *et al.*, "Performance of Transmit Antenna Selection in Non-orthogonal Multiple Access for 5G Systems", *Eighth International Conference on Ubiquitous and Future Networks*, Vienna, Austria, 2016 (https://doi.org/10.1109/ICUFN.2016.7536954).

[24] A. Dua, K. Medepalli, and A. J. Paulraj, "Receive Antenna Selection in MIMO Systems Using Convex Optimization", *IEEE Transactions on Wireless Communications*, vol. 5, no. 9, pp. 2353–2357, 2006 (https://doi.org/10.1109/TWC.2006.1687757).

[25] S. Menaa *et al.*, "On the Ergodic Capacity of MIMO-NOMA Systems with JTRAS Protocol under Imperfect SIC and CSI", *International Journal of Electronics*, pp. 1–20, 2023 (https://doi.org/10.1080/00207217.2023.2240077).

[26] T-N. Tran and M. Voznak, "On Secure System Performance over SISO, MISO and MIMO- NOMA Wireless Networks Equipped a Multiple Antenna Based on TAS Protocol", *EURASIP Journal on Wireless Communications and Networking*, vol. 2020, no. 1, pp. 1–22, 2020 (https://doi.org/10.1186/s13638-019-1586-y).

[27] M. Mohammadi, Z. Mobini, H.A. Suraweera, and Z. Ding, "Antenna Selection in Full-duplex Cooperative NOMA Systems", *IEEE International Conference on Communications (ICC)*, Kansas City, USA, 2018 (https://doi.org/10.1109/ICC.2018.8422356).

[28] M. Aldababsa *et al.*, "Unified Performance Analysis of Antenna Selection Schemes for Cooperative MIMO-NOMA with Practical Impairments", *IEEE Transactions on Wireless Communications*, vol. 21, no. 6, pp. 4364–4378, 2021 (https://doi.org/10.1109/TWC.2021.3129307).

[29] Y. Yu *et al.*, "Antenna Selection for MIMO Nonorthogonal Multiple Access Systems", *IEEE Transactions on Vehicular Technology*, vol. 67, no. 4, pp. 3158–3171, 2017 (https://doi.org/10.1109/TVT.2017.2777540).

[30] S. Beddiaf *et al.*, "Impact of Hardware Impairment on the Uplink SIMO Cooperative NOMA with Selection Relay under Imperfect CSI", *IEEE Access*, vol. 11, pp. 106706–106721, 2023 (https://doi.org/10.1109/ACCESS.2023.3318932).

[31] S. Singh and M. Bansal, "Outage Analysis of NOMA-based Cooperative Relay Systems with Imperfect SIC", *Physical Communication*, vol. 43, art. no. 101219, 2020 (https://doi.org/10.1016/j.phycom.2020.101219).

[32] S. Lee *et al.*, "Non-orthogonal Multiple Access Schemes with Partial Relay Selection", *IET Communications*, vol. 11, no. 6, pp. 846–854, 2017 (https://doi.org/10.1049/iet-com.2016.0836).

Menaa Saber, Ph.D. student
Department of Electrical Engineering
https://orcid.org/0000-0001-7363-7131
E-mail: menaa-saber@univ-eloued.dz
Echahid Hamma Lakhdar University, El-Oued, Algeria
https://lang.univ-eloued.dz/en/

Khelil Abdellatif, Professor
Department of Electrical Engineering
https://orcid.org/0000-0001-5545-1851
E-mail: abdellatif-khelil@univ-eloued.dz
Echahid Hamma Lakhdar University, El-Oued, Algeria
https://lang.univ-eloued.dz/en/

Multiprobe Planar Near-field Range Antenna Measurement System with Improved Performance

Samvel Antonyan[1] and Hovhannes Gomtsyan[1,2]

[1]*National Polytechnic University of Armenia, Yerevan, Armenia,*
[2]*European University of Armenia, Yerevan, Armenia*

https://doi.org/10.26636/jtit.2024.3.1624

Abstract — This article presents a novel multiprobe planar near-field range (PNFR) measurement system. The said system simplifies the overall mechanical design, making it simpler than the existing scanning probe PNFR measurements, and also significantly reduces testing time. A dielectric-based probe is introduced to reduce the antenna size, thereby improving resolution. The probe under consideration is an antipodal Vivaldi antenna offering broadband support and ensuring wideband characteristics of aerials. Numerical results for representative X-band antenna models, presented in the Matlab environment, demonstrate robust performance of the developed measurement system.

Keywords — microstrip patch antenna characteristics, multiprobe PNFR measurements, planar near-field antenna range

1. Introduction

The high degree of complexity associated with antennas operating at high frequencies, resulting from the proportional increase in their far-field, makes standard far-field measurements challenging in open space environments. To avoid these testing difficulties, a near-field antenna measurement technique has been introduced. This technique has gained popularity due to its reduced size and lab footprint.

Despite the fact that near-field measurements are typically performed on a planar, spherical, or cylindrical surface, planar or plane-rectangular near-field measurements are particularly popular due to the simplicity of processing and data acquisition [1]. Assuming that the number of near-field data points is $2^n + 1$, where n is a positive integer, the full plane transformation of the far-field can be computed in a time proportional to:

$$n \cdot a^2 \cdot \log_2(n \cdot a)\,,$$

where a is the radius of the smallest circle in which the test antenna is inscribed [2]. Earlier, other methods, such as bipolar or plane polar near-field measurements, were introduced. Nonetheless, these techniques offer limited benefits when considering the computational cost of the more math-intensive transformation procedure involved [3].

This article proposes a novel technique that leverages multiprobe plane rectangular near-field antenna measurements to improve antenna measurement test time performance and capabilities.

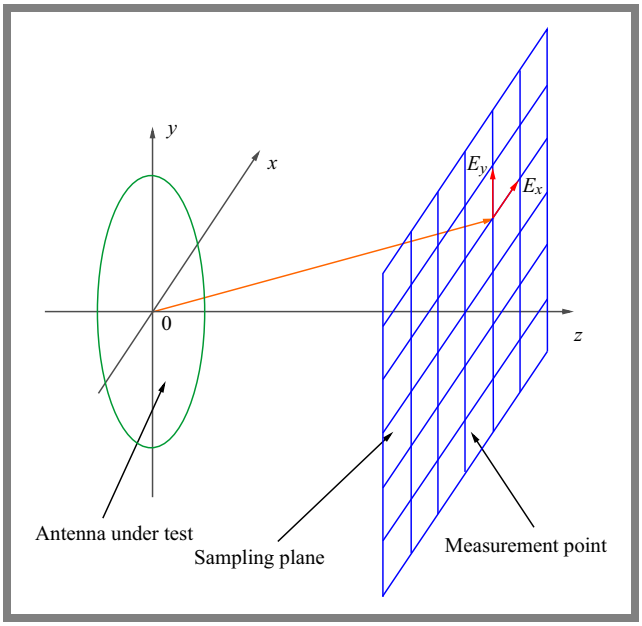

Fig. 1. Planar near-field scanning surface.

2. Problem Statement

In all of the aforementioned cases, acquisition of near field EM vectors is realized by placing the probe at a particular position, pointing it at the direction of the antenna under test (AUT) and allowing the electric field surrounding the probe to generate current. The same procedure should be repeated in two polarizations in order to be able to construct the far-field of the AUT [2]. This technique also refers to "probe scanning", where the probe needs to be moved physically within the specified surface of interest.

Acquisition of planar near-field data is usually conducted over a rectangular $x - y$ grid, as shown in Fig. 1, with a maximum near-field sample spacing of:

$$\Delta x_{max} = \Delta y_{max} = \frac{\lambda}{2}.$$

This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) License.
For more information, see https://creativecommons.org/licenses/by/4.0/

To ensure near-plane wave measurement conditions at the AUT aperture, the phase taper across the AUT should vary maximum by $\Delta\varphi = \pi/8$ or 22.5° based on its optical analogue, which was proposed in [4]. In practice, this condition can be expressed as:

$$\Delta\varphi(< \frac{\pi}{8}) \approx \frac{\pi D^2}{4\lambda R} , \tag{1}$$

where D is the aperture of the AUT, λ is the operating wavelength, and R is far-field distance.

2.1. Plane Wave (Modal) Expansion

Mathematical formulations of the planar near-field to far-field (NF/FF) system rely on the utilization of plane wave (modal) expansions through Fourier transform techniques. In theory, any form of a monochromatic wave, regardless of its arbitrary nature, can be expressed as a superposition of plane waves propagating in diverse directions, each with distinct amplitudes, but sharing a common frequency.

The primary objective of this plane wave expansion lies in the determination of the unknown amplitudes and propagation directions associated with these constituent plane waves [5]. The relationships between the near-zone E fields and far zone fields for planar systems can be represented by [4]:

$$E_x(x,y,z=0) = \frac{1}{4\pi^2}\iint_{-\infty}^{\infty} f_x(k_x,k_y)e^{-j(k_x x+k_y y)}dk_x dk_y , \tag{2}$$

$$E_y(x,y,z=0) = \frac{1}{4\pi^2}\iint_{-\infty}^{\infty} f_y(k_x,k_y)e^{-j(k_x x+k_y y)}dk_x dk_y , \tag{3}$$

where $f_x(k_x,k_y)$ and $f_x(k_x,k_y)$ represent the plane wave spectrum of the field, x and y are the components of the electric field measured over a plane surface, and k is the wave number. The far field pattern of the antenna, in terms of plane wave spectrum function f, can be found by:

$$E_\theta(r,\theta,\varphi) = j\frac{ke^{-jkr}}{2\pi r}\left(f_x\cos\varphi + f_y\sin\varphi\right) , \tag{4}$$

$$E_\varphi(r,\theta,\varphi) = j\frac{ke^{-jkr}}{2\pi r}\left(-f_x\sin\varphi + f_y\cos\varphi\right) , \tag{5}$$

The generalized methodology for identifying the far field region of an AUT involves several steps. Initially, electric field components are measured in the near field region. Subsequently, the plane wave spectrum functions, denoted as f_x and f_y, are derived by performing a direct inverse fast Fourier transform (FFT) on Eqs. (2) and (3) representing the electric field components. Finally, the far field region is computed by using Eqs. (4) and (5) for the electric field components in spherical coordinates [2].

Although near-field methods are more complicated physically and mathematically, the ability to use small distances means that it is possible to make measurements in a climate-controlled and electromagnetic-controlled environment of an antenna measurement facility, which is not possible in the case of regular far-field measurements. Potentially, this feature can also result in improvements in security, measurement accuracy, and test throughput [6].

In this paper, a multiprobe-based rectangular planar near-field range (PNFR) measurement system, together with an antipodal Vivaldi antenna (AVA)-based near-field probe, is presented. Simulation of the measurement system has been performed with the use of Altair FEKO software, and the near-field to far-field reconstruction has been competed using Matlab.

3. Probe Design

In general, the standard approach to performing a PNFR measurement involves using open-ended waveguides (OEW), as they are quite homogenous in terms of gain across wide frequency ranges. Nevertheless, designing OEWs for higher frequencies can be challenging due to the need to maintain a small footprint to satisfy the $\lambda/2$ sampling point criteria. One of the possible solutions could consists in implementing antennas on dielectrics, which are gaining in popularity these days. The benefit of this approach is that the size of the antenna can be reduced by the ε_r, i.e. the dielectric constant of the substrate [7].

It is also worth noting that one of the other requirements of the PNFR measurement system is that its bandwidth must be sufficient to cover wide frequency ranges. To satisfy all these scenarios, AVA-based near-field probes are designed to be linearly polarized and to operate over a wide bandwidth with high gain [8]. The size of these antennas is quite small due to their dielectric-based substrate. Therefore, they can be easily used in a PNFR lattice. In addition, AVA antennas offer a good return loss and are characterized by minimum signal distortion.

A basic Vivaldi antenna consists of a feed line, usually of the microstrip or stripline type, and the radiating structure. Many Vivaldi antenna designs with different radiating structures have been described in the literature. The exponentially shaped type is the most widely used, as it can provide the broadest band solution [9]. The structure of an AVA antenna is shown in Fig. 2. It includes two main parts, namely the feed line and the radiation flares.

The flares are designed to have the shape of electrical curves, as these configurations provide wide broadband characteristics. The equation for this tapered slot is [9]:

$$y(x) = \pm\left(C_1 e^x + C_2\right) , \tag{6}$$

where C_1 and C_2 are given by:

$$C_1 = \frac{y_2(x) - y_1(x)}{e^{ax_2} - e^{ax_1}} , \tag{7}$$

$$C_2 = \frac{y_1(x)e^{ax_2} - y_2(x)e^{ax_1}}{e^{ax_2} - e^{ax_1}} , \tag{8}$$

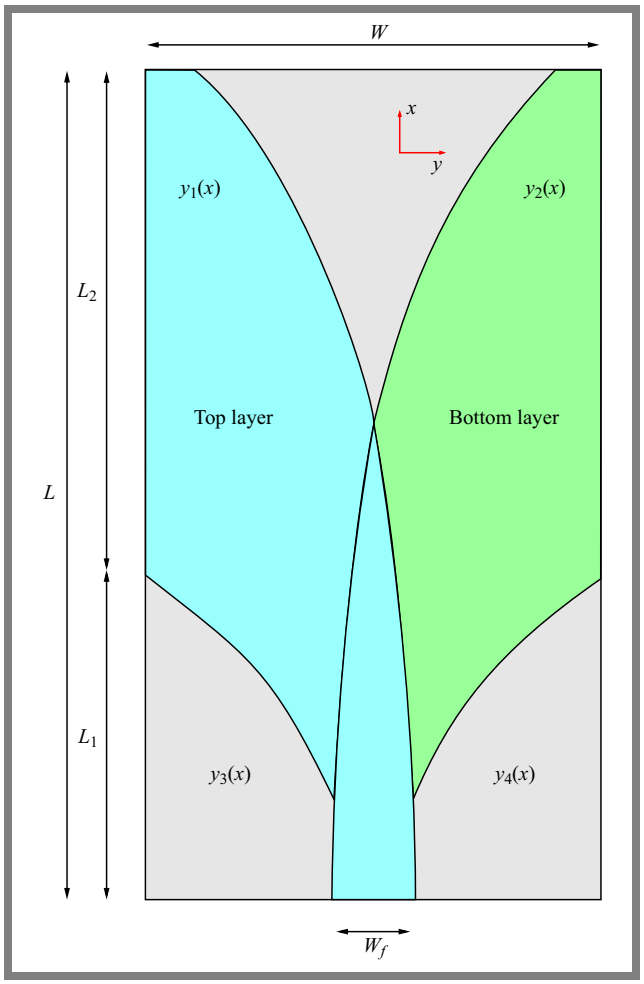

Fig. 2. AVA antenna design.

In these equations, C_1 and C_2 are constants, a is the exponential curve rate, x_1, y_1, x_2, and y_2 are the start and end points of the exponential curve [10] and are defined by:

$$\begin{cases} y_1(x) \rightarrow x_1 = 0, \quad x_2 = L, \quad y_1 = -\dfrac{W_f}{2}, \quad y_2 = \dfrac{W}{2}, \\ y_2(x) \rightarrow x_1 = 0, \quad x_2 = L, \quad y_1 = \dfrac{W_f}{2}, \quad y_2 = -\dfrac{W}{2}, \\ y_3(x) \rightarrow x_1 = 0, \quad x_2 = L_1, \quad y_1 = \dfrac{W_f}{2}, \quad y_2 = \dfrac{W}{2}, \\ y_4(x) \rightarrow x_1 = 0, \quad x_2 = L_1, \quad y_1 = -\dfrac{W_f}{2}, \quad y_2 = \dfrac{W}{2}. \end{cases} \tag{9}$$

The minimum operating frequency range f_{min}, the thickness of substrate h, and its dielectric ε_r constant, as well as length L of the antenna structure are calculated using the following equation [8]:

$$L = \frac{c}{2 f_{min} \sqrt{\varepsilon_{eff}}}, \tag{10}$$

where:

$$\varepsilon_{eff} = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2} \cdot \left(1 + 12 \frac{h}{W}\right), \tag{11}$$

The AVA layout is designed at $f_{min} = 10$ GHz on a substrate with the dielectric constant $\varepsilon_r = 4.6$ and height $h = 10^{-3}$ m. From Eq. (9), the aerial's parameters are: $W = 0.0125$ m, $L = 0.024$ m, $W_f = 2.3 \cdot 10^{-3}$ m for such the line with 50 Ω impedance is achieved.

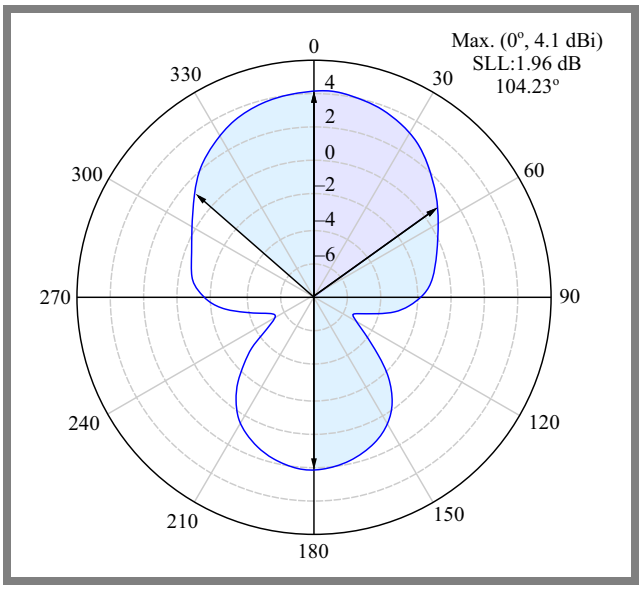

Fig. 3. AVA radiation pattern at 10 GHz in θ plane.

The coefficients of Eq. (9) are obtained during the simulation by optimizing reflection and radiation characteristics. Consequently, the end points of curves $y_3(x)$ and $y_4(x)$ are determined at $L_1 = 9.3 \cdot 10^{-3}$ m, and the equations for top and bottom tapered slots are:

$$\begin{cases} y_1(x) = 17 \cdot 10^{-5}\, e^{150x} - 17 \cdot 10^{-5}, \\ y_2(x) = -\left(17 \cdot 10^{-5}\, e^{150x} - 17 \cdot 10^{-5}\right), \\ y_3(x) = -\left(2 \cdot 10^{-3}\, e^{200x} + 0.3 \cdot 10^{-3}\right), \\ y_4(x) = 2 \cdot 10^{-3}\, e^{200x} + 0.3 \cdot 10^{-3}. \end{cases} \tag{12}$$

As shown in Fig. 3, the beamwidth in θ plane is: $2\,\theta_{0.5}$ =104.23° with a gain of 4.104 dBi.

4. Results and Discussions

The PNFR measurement setup for the operating frequency f_{op} =10 GHz (X-band) was developed with the use of the designed AVA-based near-field probe. A 1×4 phased array antenna is tested as AUT and it is defined in the FEKO environment, for far-field analysis. Results of the design's analysis are then compared with measurements concerning the NF-FF conversion.

4.1. AUT NF Results with PNFR System

As shown in Fig. 4, 17 near field AVA probes have been used in order to create the PNFR measurement setup. The distance between each probe is 3.75 mm, which is sufficient to meet sampling point requirement, as the AUT's operating wavelength is $\lambda_{op}/2$ =14.98 mm >3.75 mm [2]. The AUT is placed at the center of the setup to capture E field data at the $X = 0$ position, where the field strength is assumed to be at its highest value.

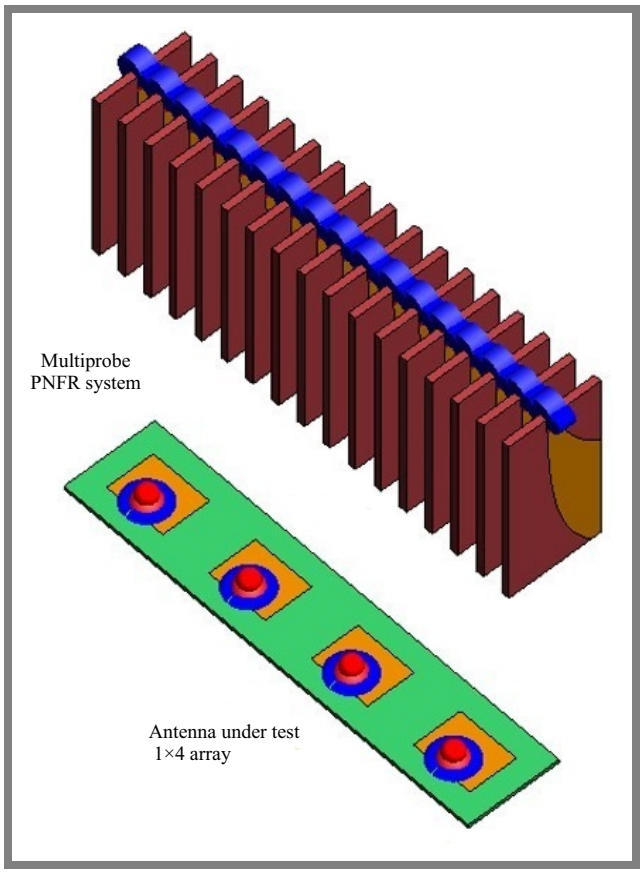

Fig. 4. PNFR measurement system using AVA probes.

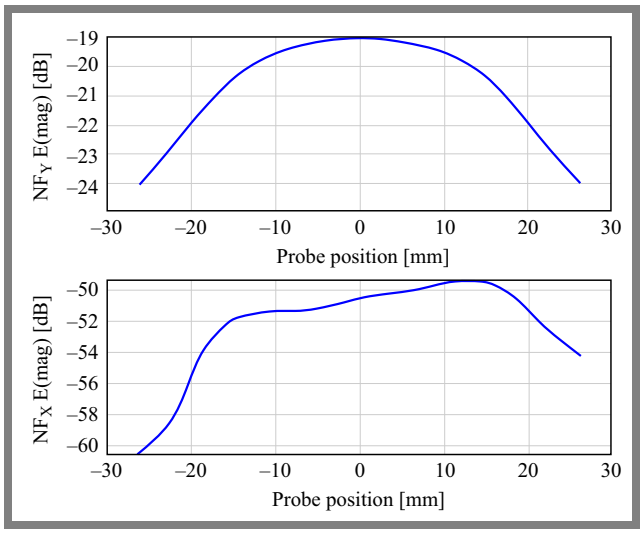

Fig. 5. Measurement results of: a) Ey field and b) Ex field.

The distance between the AUT and the farthest probe is 87.47 mm, which satisfies the radiative near-field distance criteria [6]. To obtain both X and Y polarized electric field components, the AUT was rotated by 90°. The collected E field data is shown in Fig. 5.

4.2. NF–FF Conversion

Resolution of the far-field pattern can be increased by adding artificial data sampling points (with zero value) at the outer extremities of the near-field distribution. This method increases the number of sample points without changing spacing between measurement sample, thus resulting in fixed wavenumber limits. The additional wavenumber spectrum points are within the original wavenumber limits. This leads to increased resolution in the computed far-field patterns [2]. Using the Matlab environment, the plane wave spectrum of both fields is solved by adding artificial data sampling points and applying spline interpolation, as shown in Fig. 6a-b, by using Eqs. (2) and (3).

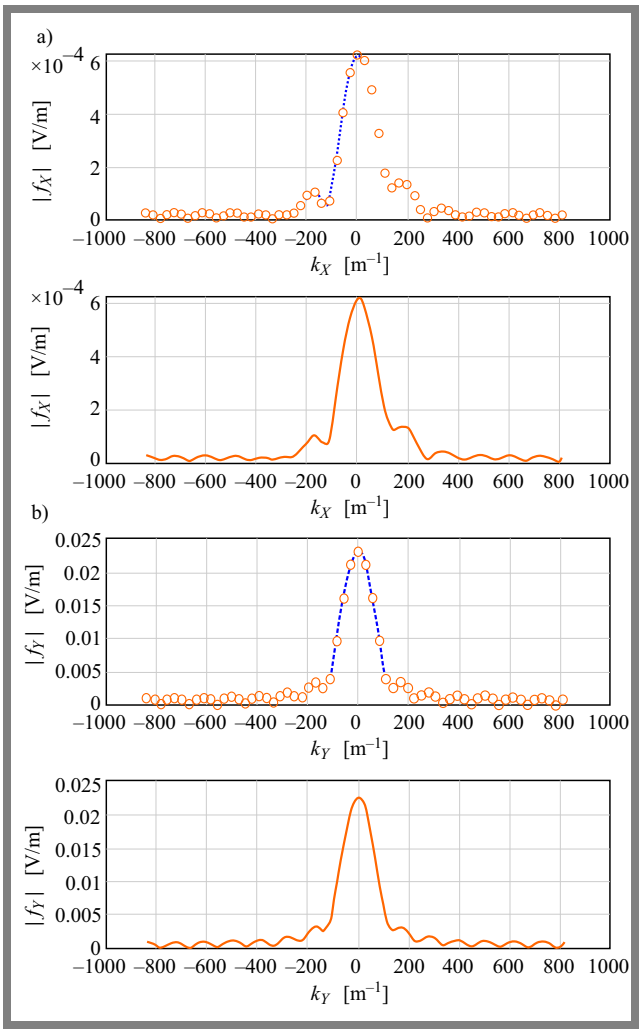

Fig. 6. Interpolated plane wave spectrum for a) x axis and b) y axis.

As mentioned in subsection 2.1, the far-field pattern of the antenna can be determined in terms of plane wave spectrum functions f_x and f_y by using Eqs. (4) and (5). The reconstructed far-field for a 1×4 phased array AUT using such functions is illustrated in Fig. 7. As one may notice, beamwidth in θ plane is $2\,\theta_{0.5} \approx 25.55°$.

4.3. AUT Design Calculation Results with Ideal Scanning Probe Approach

The procedure discussed in subsection 4.1 has been repeated for 51 ideal data sampled results analyzed in the FEKO environment. Based on the gathered dataset, a 2D plot of the analyzed ideal E field is shown in Fig. 8.

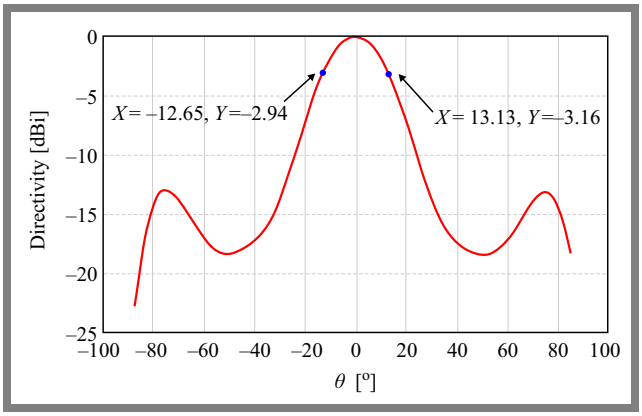

Fig. 7. Reconstructed electric field in θ plane.

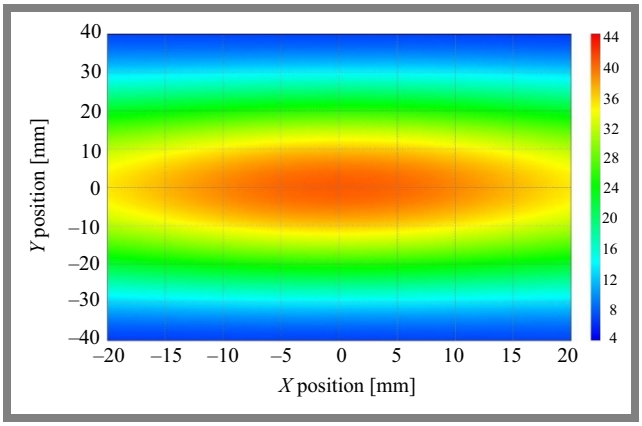

Fig. 8. 2D plot of ideal E in [V/m] in near-field of antenna under test.

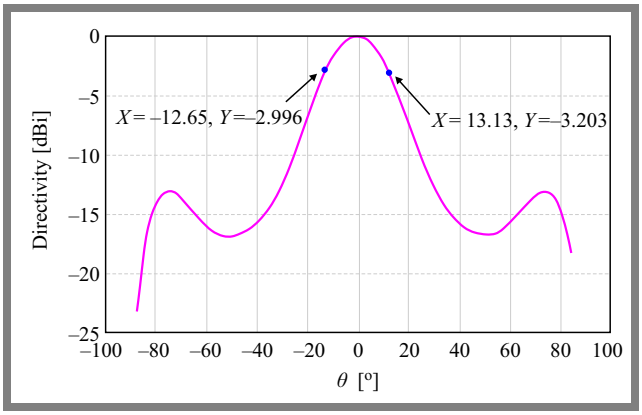

Fig. 9. Reconstructed far-field based on ideal continuous near-field E data in θ plane.

As a result of plane modal expansion and after applying Eqs. (4) and (5), the reconstructed far-field of AUT with the ideal data is shown in Fig. 9.

As one may see, the beamwidth in θ plane equals $2\,\theta'_{0.5} \approx 25.5^\circ$ and is similar to the obtained results, i.e. those that have been reconstructed based on data captured by near-field probes: $\Delta 2\,\theta_{0.5} \approx 2\,\theta_{0.5} - 2\,\theta'_{0.5} \approx 0.05^\circ$.

5. Conclusion

This paper presents the design of a novel multiprobe planar near-field range (PNFR) measurement system, coupled with an antipodal Vivaldi antenna (AVA)-based near-field probe. The multiprobe PNFR system offers a simplified mechanical structure compared to traditional scanning probe methods, leading to significant improvements in terms of testing time and efficiency.

The probe design outlines the process of designing AVA antennas, considering such factors as dielectric substrate properties and geometric configurations to optimize performance. Through simulation and optimization, the AVA antennas demonstrate decent characteristics, including compact size, wide bandwidth, and minimum signal distortion, making them suitable for integration with the PNFR system.

17 dielectric-based AVA probes have been used to characterize a 1×4 phased array as an antenna under test. Near-field probe data of the AUT and ideal near-field data from the FEKO environment were captured for a comparative analysis of AUT's directivity. The captured near-field data was converted to far-field information using an NF-FF reconstruction algorithm implemented in the Matlab environment. Far-field representation of the captured data shows that the difference between the ideal and probe-captured methods is $\Delta 2\,\theta_{0.5} \approx \pm 0.05^\circ$. This small difference proves the validity of the measurement system.

Results from the multiprobe PNFR measurements prove the effectiveness of the system in capturing near-field data with a high degree of precision. Comparison with the ideal scanning probe approach further validates the accuracy and reliability of the multiprobe PNFR method in reconstructing far-field patterns of antennas.

References

[1] A.D. Yaghjian, "An Overview of Near-field Antenna Measurements", *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 1, pp. 30–45, 1986 (https://doi.org/10.1109/TAP.1986.1143727).

[2] C.A. Balanis, *Antenna Theory Analysis and Design*, 3rd ed., John Wiley & Sons, USA, 2005 (ISBN: 9780471667827).

[3] R.G. Yaccarino, Y. Rahmat-Samii, and L.I. Williams, "The Bi-polar Planar Near-field Measurement Technique, Part II: Near-field to Far-field Transformation and Holographic Imaging Methods", *IEEE Transactions on Antennas and Propagation*, vol. 42, no. 2, pp. 196–204, 1994 (https://doi.org/10.1109/8.277213).

[4] S. Gregson, J. McCormick, and C. Parini, *Principles of Planar Near-field Antenna Measurements*, 2nd ed., SciTech Publishing, UK, 634 p., 2023 (ISBN: 9781839536991).

[5] D. Paris, W. Leach, and E. Joy, "Basic Theory of Probe Compensated Near-field Measurements", *IEEE Transactions on Antennas and Propagation*, vol. 26, no. 3, pp. 373–379, 1978 (https://doi.org/10.1109/TAP.1978.1141855).

[6] *IEEE Recommended Practice for Near-Field Antenna Measurements*, IEEE, 2012 (https://doi.org/10.1109/IEEESTD.2012.6375745).

[7] K.E. Kedze, H. Wang, Y.B. Park, and I. Park, "Substrate Dielectric Constant Effects on the Performances of a Metasurface-based Circularly Polarized Microstrip Patch Antenna", *International Journal of Antennas and Propagation*, vol. 2022, art. no. 3026677, 2022 (https://doi.org/10.1155/2022/3026677).

[8] J. Fisher, "Design and Performance Analysis of a 1–40 GHz Ultra-Wideband Antipodal Vivaldi Antenna", *German Radar Symposium GRS 2000*, Berlin, Germany, 2000.

[9] C.B. Hien, H. Shirai, and D.N. Chien, "Analysis and Design of Antipodal Vivaldi Antenna for UWB Applications", *Fifth International Conference on Communications and Electronics (ICCE)*, Danang, Vietnam, 2014 (https://doi.org/10.1109/CCE.2014.6916735).

[10] A.S. Dixit and S. Kumar, "A Survey of Performance Enhancement Techniques of Antipodal Vivaldi Antenna", *IEEE Access*, vol. 8, pp. 45774–45796, 2020 (https://doi.org/10.1109/ACCESS.2020.2977167).

Samvel Antonyan, M.Sc.
Institute of Information and Telecommunication Technologies and Electronics
https://orcid.org/0009-0007-4523-2090
E-mail: samwellantonyan@gmail.com
National Polytechnic University of Armenia, Yerevan, Armenia
https://polytech.am/en/home/

Hovhannes Gomtsyan, Ph.D.
Institute of Information and Telecommunication Technologies and Electronics
https://orcid.org/0009-0003-2479-7923
E-mail: hovhannes.gomcyan@polytechnic.am
National Polytechnic University of Armenia, Yerevan, Armenia
https://polytech.am/en/home/
European University of Armenia, Yerevan, Armenia
https://www.eua.am

Analyzing Performance of THz Band Graphene-Based MIMO Antenna for 6G Applications

Noor S. Asaad, Adham M. Saleh, and Mahmod A. Alzubaidy

Ninevah University, Mosul, Iraq

https://doi.org/10.26636/jtit.2024.3.1518

Abstract — In this paper, a compact 2×2 hexagon ring-shaped MIMO antenna operating at the terahertz band is proposed for future 6G wireless communication applications. The antenna is designed using graphene, due to its unique high-speed transmission capabilities. DGS and NL decoupling approaches are applied to enhance isolation between the two radiating elements. A parametric study is performed to investigate the significance of using these methods. Performance in terms of different metrics is studied using the CST Microwave Studio simulator. The final outcomes show that the proposed MIMO antenna achieves 23 dB of isolation, 0.004859 of ECC, 0.004 bits/sec/Hz of CCL, and efficiency of 98%.

Keywords — 6G, DGS, graphene, MIMO antenna, NL, THz band

1. Introduction

According to [1], [2], 6G will be the first wireless technology to transmit data at terabits per second and is intended to be launched in the future, between 2027 and 2030 [3]. For the development of 6G communication technologies, terahertz bands are advised to be used which are located between microwave and infrared ranges, and their spectrum varies from 0.1 to 10 THz [4]–[6].

It is critical to recognize that graphene, which works as conducting material in antennas, has numerous advantages at the THz band, such as high carrier mobility enabling very fast transport of charges and resulting in enhanced performance of the antenna. On the other hand, the main difference between graphene and metal conductors lies in the level of losses at the THz band. Graphene exhibits low losses at the terahertz band which, in turn, improves antenna efficiency [3]–[4].

Furthermore, graphene may be deposited onto substrates using different techniques, such as chemical vapor deposition (CVD), double self-assembly (DSA), and offers good adhesion strength. These methods open the way to develop lightweight, flexible, and efficient electronic devices [5].

Multiple-input multiple-output (MIMO) antennas are one of the most significant developments in mobile communications achieved in recent years. In this type technology, several antennas are placed at fixed locations on the transmitter and receiver, apart from one another [7]. The main benefit of using a MIMO antenna consists in the provision of high data rates and reduction of latency by transmitting data simultaneously over several channels. Therefore, it is essential to mention here that MIMO technology at the THz band can be considered the backbone of 6G communication systems [8]–[10]. In multi-antenna technologies, the isolation between the radiating elements is affected considerably by the separation between them. Good isolation may be obtained by increasing the distance between the elements. Low envelope correlation coefficient (ECC) values, as well as high diversity gain (DG) and channel capacity loss (CCL) values may be achieved together with high isolation [4].

There are many papers aiming to improve MIMO performance at the THz band. For example, the authors in [11] designed a proximity-coupled graphene patch-based 2×2 MIMO antenna. No decoupling method was mentioned in this design to enhance performance. On the other hand, the authors of [12] presented an elliptical-shaped microstrip feed super wide band (SWB) 2×2 MIMO antenna for the terahertz band. In this design, L-shaped decoupling was added to the ground plane to improve performance.

In [13], the authors proposed a 4×4 MIMO antenna with SWB at THz frequencies. The isolation between elements was improved by applying three different fractal antenna configurations. Furthermore, paper [4] suggested a graphene MIMO patch antenna using E-shaped metamaterial unit cells to reduce coupling by placing it between the patches. Article [14] presented a 4×4 MIMO antenna for the THz band, where the decoupling between the elements was achieved by placing the elements at different orientations.

It can be clearly seen from these studies that there is a lack of investigations concerning the effect of other decoupling methods in MIMO antennas and of their impact on performance parameters. Therefore, in this paper, a graphene-based MIMO antenna operating at the THz band for 6G applications is presented. The design consists of 2×2 monopole antennas. To increase antenna performance and improve isolation between the radiating elements, defected ground structure (DGS) and neutralization line (NL) techniques are employed.

Performance characteristics of the single-element antenna and MIMO antennas are investigated to determine applicability for 6G applications. Finally, due to the difficulty of fabricating the proposed design, CST Microwave Studio (CST), High-Frequency Structure Simulator (HFSS), and Advanced Design

This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) License.
For more information, see https://creativecommons.org/licenses/by/4.0/

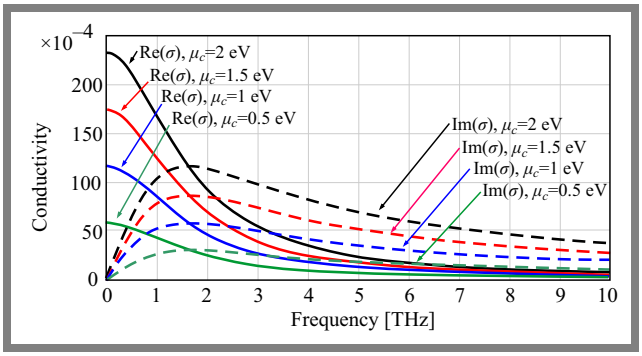

Fig. 1. Overall graphene conductivity vs. frequency for different chemical potential values.

System (ADS) software RF simulators are used to simulate the proposed solution.

2. Antenna Design

Graphene is a 2D carbon film with its thickness being qual to that of a single atomic layer [15]. In antenna applications, two aspects of its surface conductivity may be considered: intra-band, which covers the frequency range below 5 THz, and inter-band, covering frequencies above 5 THz [13]. The parameters of graphene in these two categories depend on chemical potential, frequency of operation, temperature and relaxation time, as shown by the following equations [6]:

$$\sigma_{intra} = -j\frac{e^2 k_B T}{\pi h^2(\omega - j\tau^{-1})}\left[\frac{\mu c}{k_B T} + 2\ln(e^{-\frac{\mu c}{k_B T}} + 1)\right], \quad (1)$$

$$\sigma_{inter} = \frac{je^2}{4\pi h}\ln\frac{2|\mu c| - h(\omega - j\tau^{-1})}{2|\mu c| + h(\omega - j\tau^{-1})}, \quad (2)$$

where T represents temperature, h represents reduced Planck's constant, K_B is Boltzmann constant, μ_c stands for chemical potential in eV, e is electron charge, τ represents the relaxation time and its value is 0.1 ps at room temperature, and, finally, ω is the frequency of operation.

The total intra-band and inter-band surface conductivity may be calculated as:

$$\sigma_s = \sigma_{inter} + \sigma_{intra}\,. \quad (3)$$

The overall conductivity of the graphene layer is shown in Fig. 1 with different chemical potentials. It is clearly seen that changing the chemical potential has a high impact on overall surface conductivity. On the other side, the chemical potential affects the carrier density by applying an external gate voltage [6]. This property can be very helpful in controlling the resonant frequency in antenna design processes.

3. Simulation Results

3.1. Single Antenna Design

The proposed shape with one radiating element is shown in Fig. 2. It consists of a hexagon ring that uses teflon with a dielectric constant ε_r of 2.1 and thickness h of 10 μm. The hexagon ring is printed above the dielectric surface while the ground plane is located underneath it. The overall dimensions of the presented element are 53×35×10 μm^3. Graphene is used as a conducting material due to its unique features, such as reducing antenna losses and high-speed conductivity within the THz band. The thickness of graphene is set at 35 nm. The antenna is also fed by a microstrip line with a size of 15.13×3.60 μm^2.

The final dimensions of the proposed structure are listed in Tab. 1. The simulation target is determined to design an antenna that covers the terahertz band with a resonant frequency of approximately 3 THz. The design is simulated with different values of the applied chemical potential (from 0 to 2 eV) to achieve the resonant frequency. The simulated reflection coefficient is illustrated in Fig. 3. It can be clearly noted that the operating frequency is affected by changing the chemical potential and varies from 2.62 to 3.67 THz with $S_{11} \leqslant -10$ dB by applying $\mu_c = 2$ eV.

After that, the equivalent circuit of the proposed antenna is simulated by the ADS software to enhance the results obtained from CST and HFSS, as shown in Fig. 4. A comparison between the simulated S_{11} obtained from the three different

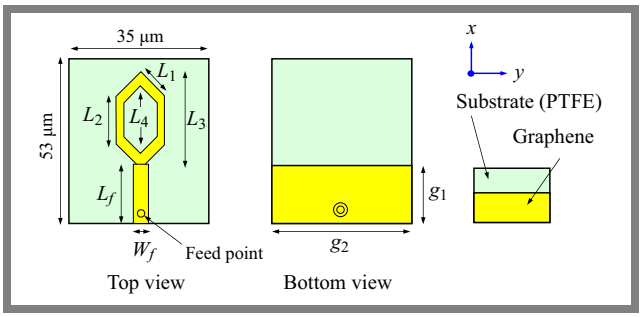

Fig. 2. Details of the hexagon ring antenna design.

Tab. 1. Dimensions of one radiating element in [μm].

Parameter	L_1	L_2	L_3	L_4
Value	8.50	12.02	24.04	16.83
Parameter	g_1	g_2	w_f	L_f
Value	14.70	35	3.60	15.13

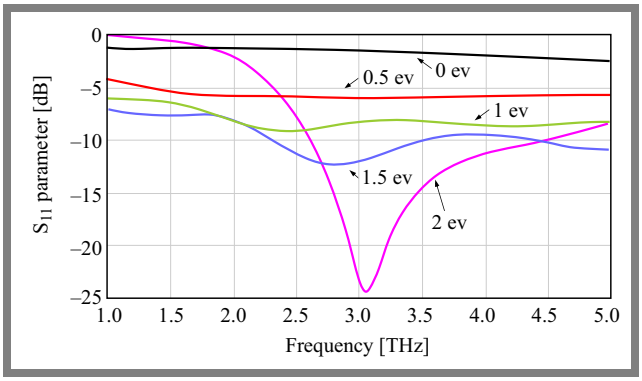

Fig. 3. Simulated result of S_{11} for a single hexagon ring antenna for different chemical potential values.

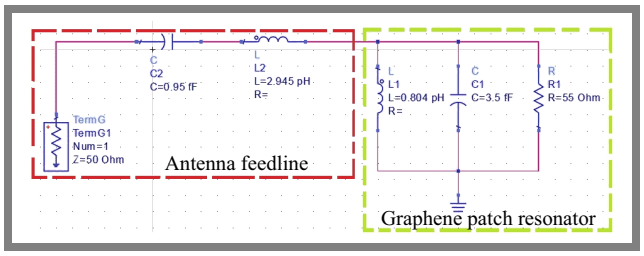

Fig. 4. Equivalent circuit model for the proposed single antenna.

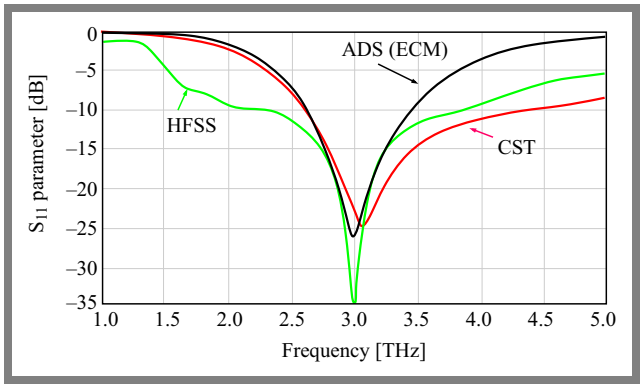

Fig. 5. Simulated S_{11} for the single element antenna with CST, HFSS, and ADS equivalent circuit.

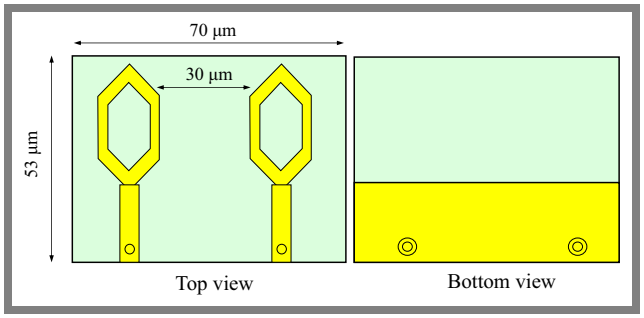

Fig. 6. Dimensions of the proposed 2×2 MIMO antenna.

simulators is illustrated in Fig. 5. As one may notice, the figure shows a perfect agreement between the obtained results.

3.2. Two-Element MIMO Antenna Design

In this part, a 2×2 MIMO hexagon ring antenna is designed as a single structure, as shown in Fig. 6. The distance between the two antenna elements is 30 µm and the dimensions of the ground plane equal 70×14.70×0.035 µm^3. The total size of the presented antenna is 53×70×10 µm^3.

The equivalent circuit of the proposed two-element MIMO antenna from the ADS simulator is presented in Fig. 7. Figure 8 shows the simulated S_{11} of the 2×2 MIMO antenna from the three different simulation software instances. It is clear from CST results that the antenna works from 2.47 to 3.44 THz, with a resonant frequency of 2.8 THz. It has been also observed that the resonant frequency has been shifted to a lower value compared to the single antenna case. This stems from mutual coupling between the two radiating elements.

The simulated transmission coefficient S_{21} is shown in Fig. 8. This plot proves that mutual coupling S_{21} between the two radiating elements is less than –15 dB for the 2.02 to 2.5 THz

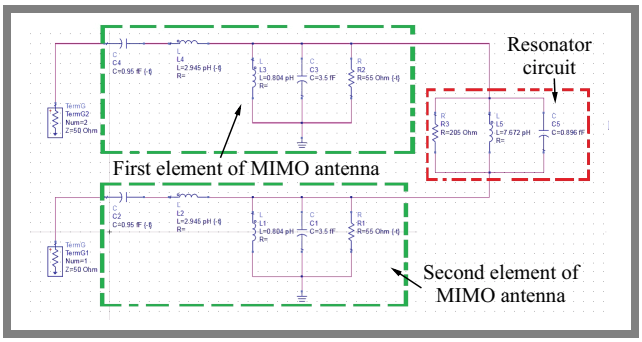

Fig. 7. Equivalent circuit model of the proposed 2×2 MIMO configuration.

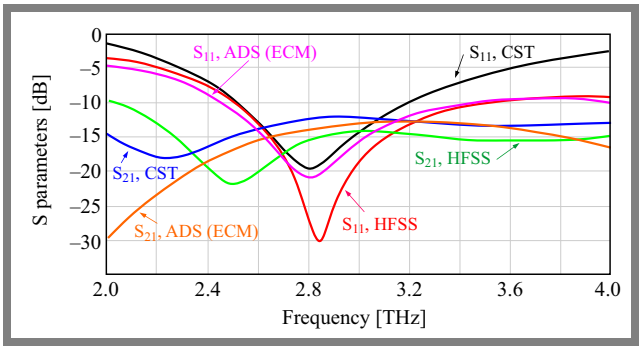

Fig. 8. S parameter values for a 2×2 MIMO antenna simulated with CST, HFSS, and ADS software.

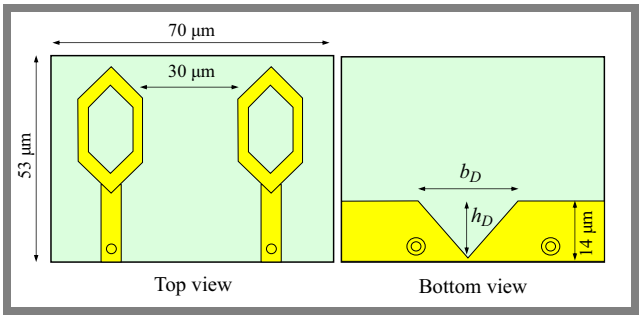

Fig. 9. Proposed geometry of a 2×2 MIMO hexagon ring antenna with DGS.

range. As the frequency ranges for S_{11} and S_{21} are not the same, mutual coupling should be reduced below –15 dB for the entire operating range. Finally, a good agreement between the results of CST, HFSS and ADS is achieved.

3.3. MIMO Antenna with DGS Decoupling Approach

Defected ground structure (DGS) is an etched-out area on the ground plane of a microstrip PCB. It is a beneficial method that is used to improve the gain and the bandwidth of printed antennas. It can be also used to reduce coupling between radiating elements in multi-antenna technologies, due to its unique feature of inserting virtual inductances and capacitances into the structure of the MIMO antenna. Hence, it may act as a stop band filter. A defective area on the ground plane is represented by a triangular shape in the middle between the two elements, as shown in Fig. 9.

The effects of height h_D and base b_D of the defective area are studied to achieve the lowest results in terms of mutual

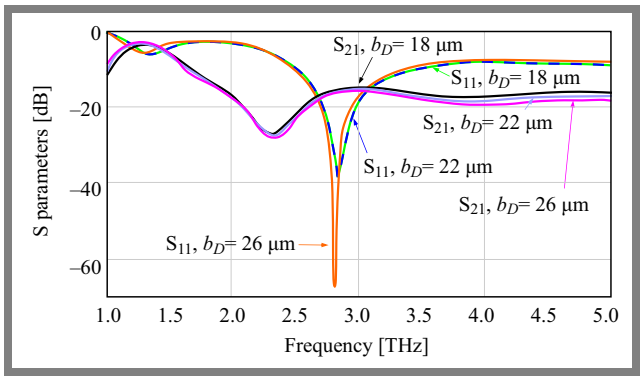

Fig. 10. Simulated S_{11} and S_{21} for a 2×2 MIMO hexagon ring antenna with DGS: $h_D = 14$ μm and variable b_D.

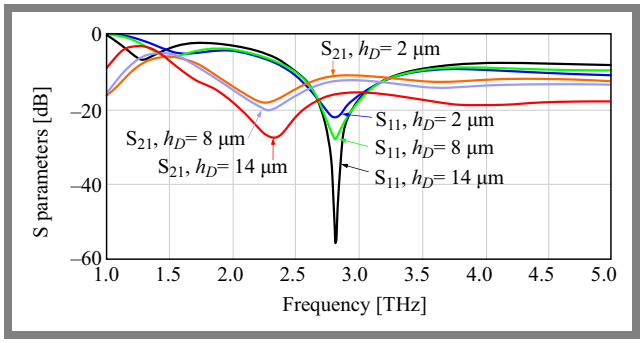

Fig. 11. Simulation of S_{11} and S_{21} for a 2×2 MIMO hexagon ring antenna with DGS: $b_D = 26$ μm and variable h_D.

coupling. The analysis starts by setting height h_D to the maximum value of 14 μm and changing the length of base b_D. The initial value of b_D is fixed at 18 μm and then different lengths of b_D are tested.

Figure 10 illustrates the reflection and transmission coefficients obtained for different values of b_D. One may notice that the frequency band of the reflection coefficient S_{11} is slightly affected by changing the length of b_D and the magnitude of S_{11} is reduced significantly by increasing b_D. The best value is obtained at $b_D = 26$ μm with $|S_{11}|$ equaling –67 dB.

On the other hand, isolation is improved by increasing the length of b_D and the antenna achieved an isolation value of more than 15 dB for the entire frequency band. Another study concerning the parameters of the defected area consists in assuming a constant value of the base ($b_D = 26$ μm) and changing height h_D. Three different h_D values are tested and the outcomes of the simulated reflection and transmission coefficients are presented in Fig. 11.

The same dependence as described in the first parametric study is noticed in this case as well, i.e. an increase in h_D leads to a reduction in the magnitude of S_{11} and in enhancing the isolation between the two antennas. From this simulation, we can conclude that the optimum dimensions for the defected area equal $h_D = 14$ and $b_D = 26$ μm.

3.4. MIMO Antenna with Three DGSs

This section investigated the use of the DGS technique in a MIMO antenna, in three different regions. The first cut has a triangular shape and is located in the middle between the

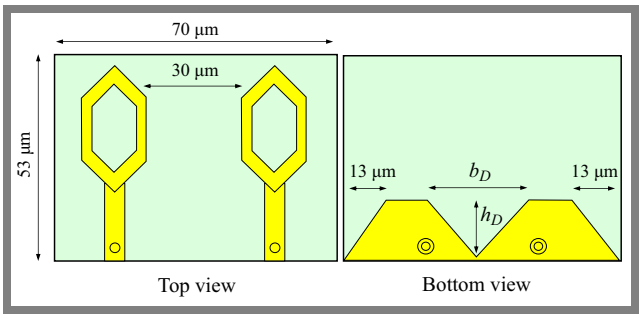

Fig. 12. A 2×2 MIMO hexagon ring antenna with three DGS.

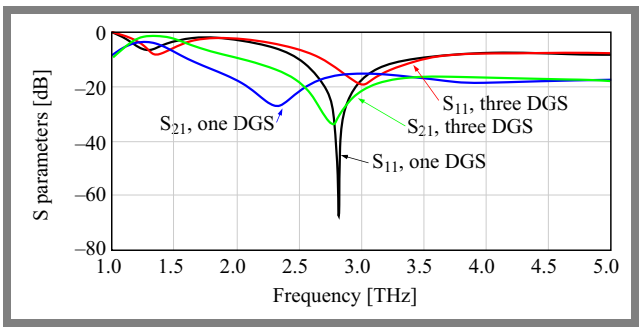

Fig. 13. Simulated results of S_{11} and S_{21} for a 2×2 MIMO hexagon ring antenna with one and three DGSs.

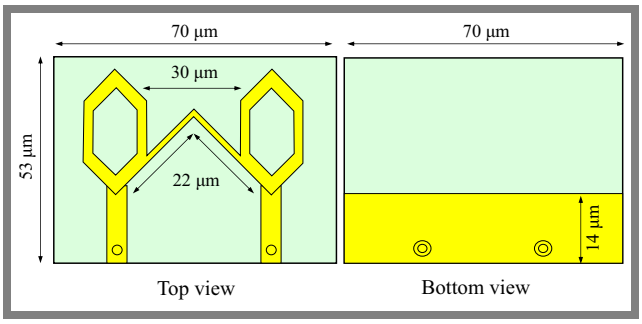

Fig. 14. A 2×2 MIMO hexagon ring antenna with NL.

two radiating elements with dimensions $b_D = 26$ μm and $h_D = 14$ μm. The second and third DGSs have the shape of a half triangle, as shown in Fig. 12.

The comparison between the simulated results of S_{11} and S_{21} and the three DGSs are shown in Fig. 13. By applying the three DGSs, the resonant frequency of the reflected coefficient is shifted to the upper band and the achieved bandwidth is widened slightly. The mutual coupling band is also shifted to the upper band and the antenna achieves a lower mutual coupling value within the operating frequency band, i.e. 22 dB at the resonant frequency, compared to –17 dB for a solution with one DGS.

3.5. MIMO Antenna Design Based on NL Decoupling Approach

Coupling in MIMO antennas may be reduced by adding a neutralization line between the radiating elements. In this technique, the current in one position is neutralized by using a sample with an inverted phase, so it can cancel the current from the other radiating element.

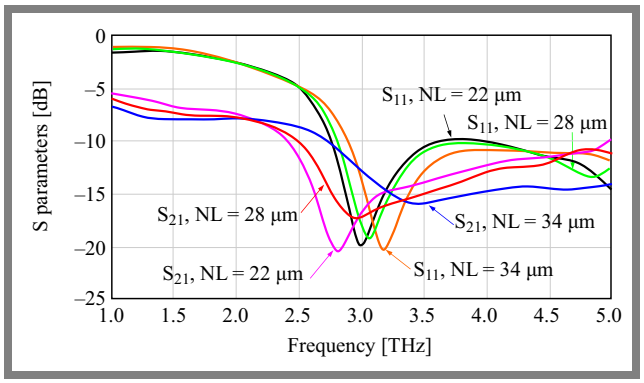

Fig. 15. Simulated outcomes of S_{11} and S_{21} for a 2×2 MIMO hexagon ring antenna with NL.

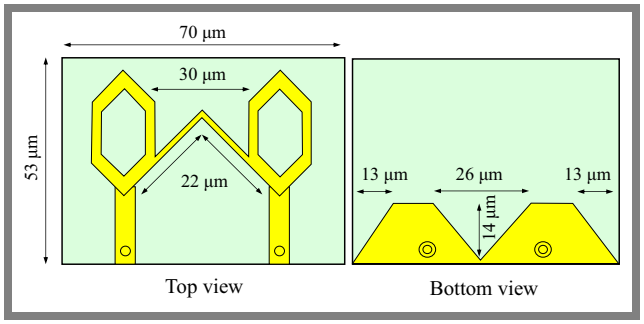

Fig. 16. A 2×2 MIMO hexagon ring antenna with NL and with DGS.

The technique is incorporated by using NL in the form of a triangle with a total length of 22 µm, as shown in Fig. 14. During simulation, the length of NL was varied to achieve the optimum (lowest) coupling value. S_{11} and S_{21} results are illustrated in Fig. 15. It may be noted that two different phenomena appeared by adding NL. The first one consists in the fact that the resonant frequency is shifted toward the upper frequency band, while the other is that a decrease in the length of NL leads to shifting the resonant frequency towards the lower frequency band. Coupling between the elements is reduced with a longer neutralization line. Isolation of more than 15 dB is obtained around the resonant frequency with NL of 22 µm.

To further analyze the obtained results, the dependence between the length of NL and the wavelength inside the substrate λ_d is simulated by determining the ratio between these values. The following equations are used to calculate λ_d inside the substrate:

$$\varepsilon_{eff} = \frac{\varepsilon_r + 1}{2}, \tag{4}$$

$$\lambda_d = \frac{\lambda_o}{\sqrt{\varepsilon_e}}, \tag{5}$$

where λ_o represents wavelength in free space, λ_d is wavelength inside the substrate, ε_{eff} stands for effective permittivity and ε_r represents relative permittivity.

The obtained NL-to-λ_d ratio is 98% at 2.8 THz, at the resonant frequency without using any decoupling methods. This ratio proves high correlation and shows that NL is working properly.

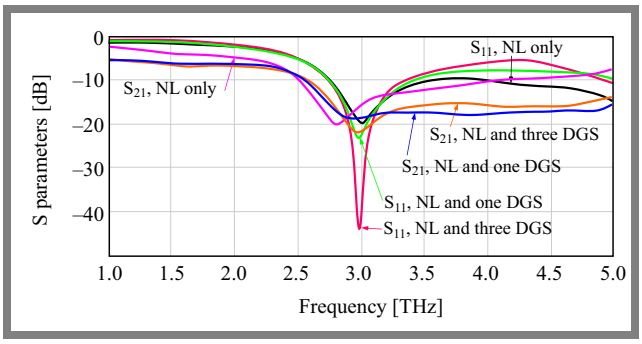

Fig. 17. Simulated results of S_{11} and S_{21} for a 2×2 MIMO hexagon ring antenna with NL for three different cases.

3.6. MIMO Antenna with DGSs and NL Decoupling Methods

Finally, the proposed design of a 2×2 MIMO hexagon ring antenna was tested by combining the NL decoupling method with DGS, as shown in Fig. 16. The simulated values of reflection coefficient S_{11} and transmission coefficient S_{21} are compared, as shown in Fig. 17.

The resonant frequency is 3 THz in all simulations and the magnitude is significantly reduced, reaching –45 dB for the scenario with NL and three DGSs. Coupling is also reduced at the resonant frequency and reaches –23 dB at 3 THz.

4. Performance Evaluation

To assess the performance of the proposed design, a range of parameters, such as channel capacity loss (CCL), total active reflection coefficient (TARC), diversity gain (DG) and envelope correlation coefficient (ECC) is determined. ECC-related results for all six cases are presented in Fig. 18a – in all of the scenarios ECC is below 0.1 at the resonant frequency.

DG values (Fig. 18b) simulated for all the cases equal approximately 10 dB at the resonant frequency, while CCL (Fig. 18c) is approximately zero at the resonant frequency which is below 0.4 bit/s/Hz, i.e. reaches the limit for a MIMO antenna. Finally, TARC is determined (Fig. 18d) for different phase angles at 30° steps, covering the range from 0 to 180°. TARC is below 6 dB at the resonant frequency and in drops, in some cases, below 10 dB.

Simulations of the proposed antenna's total efficiency for all three cases are summarized in Fig. 19a. The highest efficiency of 98% is achieved in the case of NL and DGSs, compared with all other cases at the operating frequency of 3 THz.

The realized gain (Fig. 19b) at the resonant frequency is approximatley 2 dB for all the cases. Finally, far field radiation patterns are plotted in 3D for the case of NL and DGSs in Fig. 20, showing the antenna's omnidirectional pattern.

Table 2 summarizes the parameters of similar antennas from the literature and compares them with the proposed design. It is clear from this comparison that the proposed design is characterized by the smallest physical size. This can be very helpful in increasing the number of radiating elements when adopting massive MIMO.

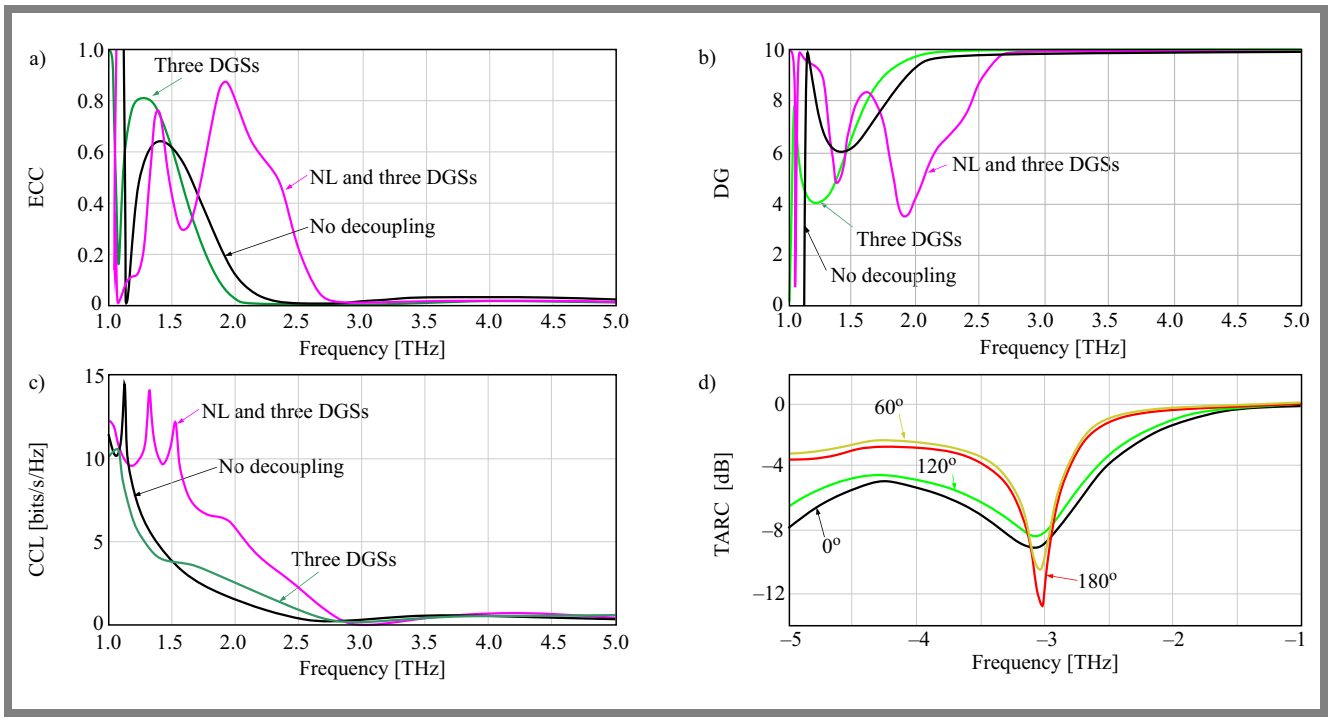

Fig. 18. Simulations of: a) ECC, b) DG, c) CCL, and d) TARC of the proposed antenna.

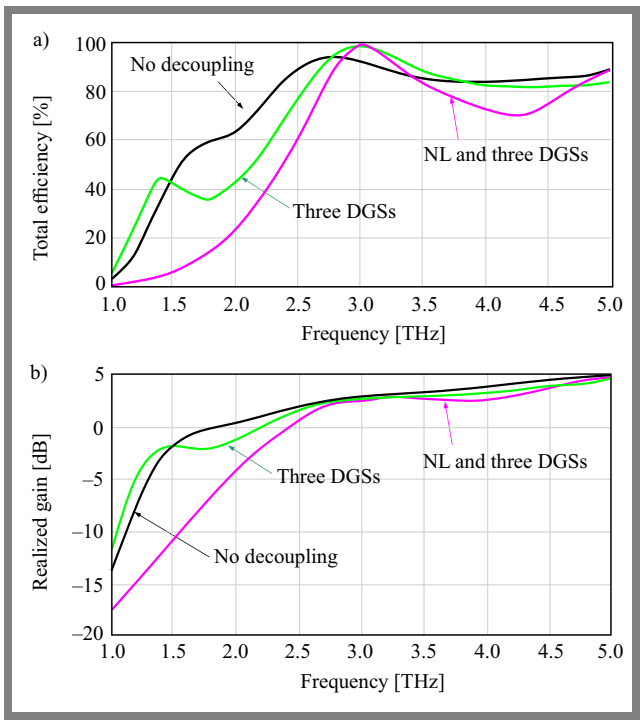

Fig. 19. Total efficiency a) and realized gain b).

On the other hand, the antenna achieves high efficiency compared with other works. Furthermore, the coupling of the design is below –15 dB for the entire band and –23 dB at the resonant frequency. This means that only 0.5% of the applied power will be coupled to the adjusting antenna at the resonant frequency.

This comparison reveals that this MIMO antenna offers promising performance in the terahertz band.

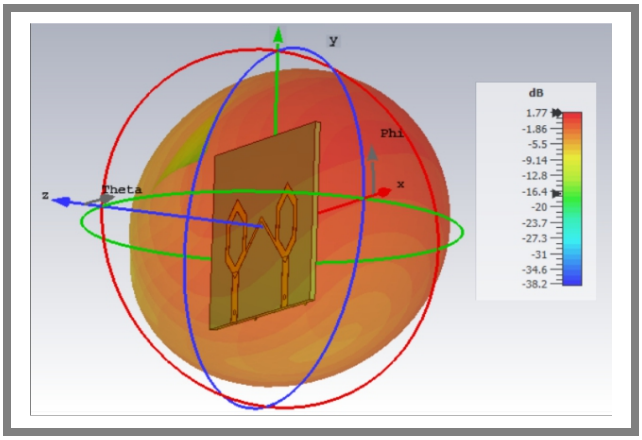

Fig. 20. Simulated 3D radiation pattern for the case with NL and DGSs.

5. Conclusion

This paper presents a two-element compact MIMO antenna utilizing graphene, intended for operating in the THz band for use in upcoming 6G applications. Mutual coupling inside the antenna structure has been minimized by applying two different decoupling approaches. It has been noticed the best isolation value is achieved when the two design approaches are combined together. The value obtained for the combined decoupling approach is –23 dB at the resonant frequency of 2.8 THz. The fundamental characteristics of the proposed MIMO antenna have been studied in terms of ECC, DG, TARC, CCL, antenna gain, efficiency, and antenna radiation pattern. The outcomes show that the antenna may serve as a good candidate for deployment in future 6G applications.

Tab. 2. Comparison with other works.

Ref.	Resonant frequency [THz]	Antenna size [μm^2]	Mutual coupling [dB]	CCL [bits/sec/Hz]	Antenna efficiency	DG [dB]	ECC
[4]	1.9	120×90	–54	0.0014	–	9.99	0.000023
[11]	1.82	60×40	–25	–	80%	9.95	–
[12]	0.33–10	1000×1400	–25	0.25	70%	9.99	0.0015
[13]	1.42	125×125	–20	0.31	–	9.98	0.0022
[14]	7.1–13	100×100	–18	–	–	9.97	0–0.5
This work	2.8	70×35	–23	0.004	98%	9.99	0.004859

References

[1] S. Elmeadawy and R.M. Shubair, "Enabling Technologies for 6G Future Wireless Communications: Opportunities and Challenges", 2020 (https://doi.org/10.48550/arXiv.2002.06068).

[2] F. Tariq *et al.*, "A Speculative Study on 6G", *IEEE Wireless Communications*, vol. 27, no. 4, pp. 118–125, 2019 (https://doi.org/10.1109/MWC.001.1900488).

[3] M.S. Swetha, M.S. Muneshwara, A.S.M. Hegde, and Z. Lu, "6G Wireless Communication Systems and Its Applications", in: *Machine Learning and Mechanics Based Soft Computing Applications. Studies in Computational Intelligence*, vol. 1068, pp. 271–288, 2023 (https://doi.org/10.1007/978-981-19-6450-3_25).

[4] S.A. Khaleel, E.K.I. Hamad, N.O. Parchin, and M.B. Saleh, "Programmable Beam-steering Capabilities Based on Graphene Plasmonic THz MIMO Antenna via Reconfigurable Intelligent Surfaces (RIS) for IoT Applications", *Electronics*, vol. 12, no. 1, 2023 (https://doi.org/10.3390/electronics12010164).

[5] A.B. Suriani *et al.*, "Synthesis, Transfer and Application of Graphene as a Transparent Conductive Film: a Review", *Bulletin of Materials Science*, vol. 43, art. no. 310, pp. 1–14, 2020 (https://doi.org/10.1007/s12034-020-02270-9).

[6] S.A. Khaleel, E.K.I. Hamad, N.O. Parchin, and M.B. Saleh, "MTM-Inspired Graphene-based THz MIMO Antenna Configurations Using Characteristic Mode Analysis for 6G/IoT Applications", *Electronics*, vol. 11, no. 14, 2022 (https://doi.org/10.3390/electronics11142152).

[7] M.H. Reddy and D. Sheela, "MIMO Antenna Design and Optimization with Enhanced Bandwidth for Wireless Applications", *Journal of Telecommunication and Information Technology*, no. 4, pp. 22–26, 2020 (https://doi.org/10.26636/jtit.2020.145520).

[8] B.-Y. Duan, "Evolution and Innovation of Antenna Systems for Beyond 5G and 6G", *Frontiers of Information Technology and Electronic Engineering*, vol. 21, no. 1, 2020 (https://doi.org/10.1631/FITEE.2010000).

[9] T.O. Olwal, P.N. Chuku, and A.A. Lysko, "Antenna Research Directions for 6G a Brief Overview Through Sampling Literature", *7th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Coimbatore, India, 2021 (https://doi.org/10.1109/ICACCS51430.2021.9441781).

[10] K.K. Wong, K.F. Tong, Y. Zhang, and Z. Zheng, "Fluid Antenna System for 6G: When Bruce Lee Inspires Wireless Communications", *Electronics Letters*, vol. 56, no. 24, pp. 1288–1290, 2020 (https://doi.org/10.1049/el.2020.2788).

[11] G. Varshney, S. Gotra, V.S. Pandey, and R.S. Yaduvanshi, "Proximity-coupled Two-port Multi-input-multi-output Graphene Antenna with Pattern Diversity for THz Applications", *Nano Communication Networks*, vol. 21, art. no. 100246, 2019 (https://doi.org/10.1016/j.nancom.2019.05.003).

[12] G. Saxena, Y.K. Awasthi, and P. Jain, "High Isolation and High Gain Super-wideband (0.33-10 THz) MIMO Antenna for THz Applications", *Optik*, vol. 223, art. no. 165335, 2020 (https://doi.org/10.1016/j.ijleo.2020.165335).

[13] S. Das, D. Mitra, and S.R.B. Chaudhuri, "Fractal Loaded Planar Super Wide Band Four-element MIMO Antenna for THz Applications", *Nano Communication Networks*, vol. 30, art. no. 100374, 2021 (https://doi.org/10.1016/j.nancom.2021.100374).

[14] R.H. Abd and H.A. Abdulnabi, "Design of Graphene-based Multi-input Multi-output Antenna for 6G/IoT Applications", *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 31, no. 1, pp. 212–221, 2023 (https://doi.org/10.11591/ijeecs.v31.i1.pp212-221).

[15] G.W. Hanson, A.B. Yakovlev, and A. Mafi, "Excitation of Discrete and Continuous Spectrum for a Surface Conductivity Model of Graphene", *Journal of Applied Physics*, vol. 110, no. 11, 2011 (https://doi.org/10.1063/1.3662883).

Noor S. Asaad, M.Sc.
Department of Communication Engineering
https://orcid.org/0000-0000-0000-0000
E-mail: noor.sabah2021@stu.uoninevah.edu.iq
Ninevah University, Mosul, Iraq
https://uoninevah.edu.iq/en

Adham M. Saleh, Ph.D.
https://orcid.org/0000-0003-4589-7491
E-mail: adham.saleh@uoninevah.edu.iq
Ninevah University, Mosul, Iraq
https://uoninevah.edu.iq/en

Mahmod A. Alzubaidy, Assistant Professor
https://orcid.org/0000-0003-4115-9491
E-mail: mahmod.mahmod@uoninevah.edu.iq
Ninevah University, Mosul, Iraq
https://uoninevah.edu.iq/en

Increasing Parallelism in Forward-backward Distributed Algorithm for Finding Strongly Connected Components of Directed Graphs

Dominik Ryżko

Warsaw University of Technology, Warsaw, Poland

https://doi.org/10.26636/jtit.2024.3.1693

Abstract — The paper proposes a modification of the existing distributed forward-backward algorithm for finding strongly connected components in directed graphs. The modification aims at improving the parallelism of the algorithm by increasing the branching factor while dividing the workload. Instead of randomly picking the pivot vertex, a heuristic technique is used which allows more sub-tasks to be generated, on average, for the subsequent step of the algorithm. The work describes suitable algorithm modifications and presents empirical results, proving suitability of the approach in question.

Keywords — distributed computation, graph algorithms, strongly connected components

1. Introduction

Directed graphs are an elegant formalism suitable for representing several abstract notions and natural phenomena. For real life problems, such graphs contain a lot of detailed information, so their size becomes extremely large. Therefore, there is a need to decompose such structures into meaningful sub-graphs which can be interpreted and processed independently.

One of the most important potential decompositions is the detection of strongly connected components (SCCs), i.e. maximal subsets of vertices in which a directed path exists from each vertex to any other vertex.

SCC detection has several practical applications, including in telecommunications [1], [2], radiation transport solvers [3], distributed formal verification [4], distributed reasoning [5], [6] analysis of citation graphs [7], [8] page rank calculations [9], and social networks [10]. The number, size and ordering of SCCs may provide valuable information about the data represented by the graph. Consequently, one may gain a deeper insight into the specific phenomenon modeled by the data.

Various SCC detection algorithms have been proposed in literature so far. Along with the growing availability of parallel computation architectures, scientific interest has shifted recently towards those algorithms that are able to effectively utilizing such capabilities. In order to take advantage of the efficiency of distributed processors, the algorithms have to split the workload into several sub-tasks which can be computed separately. The degree of such decomposition can be measured by the branching factor of a tree of tasks generated at each step of a recursive algorithm.

In the case of several parallel algorithms used for finding SCCs, selection of a pivot, i.e. vertex from which the decomposition begins, is an important step impacting the algorithm branching process. Typically, a random vertex is selected, with such an approach usually rendering satisfactory results.

In this work, we introduce a heuristic pivot selection method improving parallelism by increasing the average branching factor of the algorithms.

The remainder of the paper is structured as follows. Section 2 describes state of the art in the field of SCC detection, with the emphasis placed on parallel algorithms and pivot selection strategies. In Section 3, basic concepts used in the article are defined. A description of heuristics and algorithm modifications is presented in Section 4. Section 5 presents empirical results from experiments conducted with the use of the algorithms in question. Finally, conclusions and future work are presented in Section 6.

2. Related Work

To date, several algorithms for finding SCCs have been proposed. Classical approaches based on depth-first search [11], [12] have been proven to work in linear time with respect to the number of edges. This group of algorithms is difficult to parallelize and, therefore, is not useful in multi-core architectures. However, some limited results have been achieved in this area [13]. Also, several distributed algorithms for speeding up the process have been proposed and successfully refined [14]–[18]. The most important of these algorithms will be described in the subsequent sections.

With the proliferation of GPUs observed in recent years, multiple algorithms rely on their computational power to improve efficiency and inherent parallelism [19], [20].

This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) License.
For more information, see https://creativecommons.org/licenses/by/4.0/

2.1. Forward-backward Algorithm

One of the most popular approaches has the form of the forward-backward (FW-BW) algorithm proposed originally by Fleischer *et al.* [21] and further refined by other researchers [3], [22]. It is based on the computation of forward and backward closures of a selected node, referred to as a pivot. One can prove that the intersection of the closures is a single SCC and that the algorithm may be recursively called for three disjointed sets, i.e. vertices in the forward closure, but not in the backward closure, vertices in the backward closure but not in the forward closure, and the remaining vertices which did not belong to any of the closures.

Pivot selection is an important step in the above algorithm, since it exerts a direct impact on how many sub-problems can be generated in the subsequent step. To date, in most of the approaches random pivot detection is used. Several other methods have been proposed as well.

Barnat and Moravec [23] introduced an approach to prevent some trivial strongly connected components from being selected as pivots. In this method, the pivot is chosen out of a set of candidates computed with the use of the back-level edge concept. A back-level edge in a cycle is an edge that leads from a vertex with some distance from the initial vertex to a vertex with an equal or smaller distance from the initial vertex. The solution is to use destination vertices of the back-level edges as candidate pivots.

Another approach to pivot selection can be found in [18], where a multipivot algorithm is proposed. The idea is to sample s vertices and to start the reachability procedure simultaneously for each of them. Once this is done, the following sets are computed:

$A =$ vertices reached from vertex set $\{1, 2, ..., s-1\}$,

$B =$ vertices reached from vertex s,

$C =$ vertices that reach and are reached from s.

Finally, we can recursively compute, in parallel, the strongly connected components of:
$V - (A \cup B)$, $A - B$, $B - (A \cup C)$, and $(A \cap B) - C$.

This paper proposes yet another heuristic pivot selection technique. The advantage of the proposed approach is that it involves practically no computational or memory overhead.

Alternative improvements to the FW-BW approach have been proposed using the idea that a spanning tree can be employed instead of the breadth first search (BFS) tree. A promising version of such an approach has been shown by Ji *et al.* [24], with a prototype called iSpan, consisting of a parallel, relaxed synchronization design of spanning trees for detecting large and small SCCs, combined with fast trims for small SCCs.

The trimming phase is an important improvement to the FW-BW algorithm and is a subject of a separate field of research seeking further improvements (e.g. [25]), where FB-AI-Trim, a parallel SCC algorithm with selective trimming is proposed.

2.2. OBF Algorithm

OBF [23] is another parallel SCC detection algorithm which identifies a number of independent sub-graphs (called OBF slices) in $O(n+m)$ time. The slices are then decomposed using the FW-BW algorithm. The algorithm starts with a complete set of vertices and selection of the initial vertex.

One OBF slice is computed is one step. Then, the identified slice is decomposed and the remaining part of the graph is analyzed, in parallel, for extraction of the next slice. The OBF algorithm has been improved to create the recursive OBF approach, in which OBF is applied recursively to the extracted slices [26].

2.3. Coloring Algorithm

Yet another approach to SCC detection is based on the concept of coloring [27]. Such a technique starts with a set of ordered colors. Initially, all vertices are assigned different colors. Then, they are propagated along the edges, so that a higher color overwrites a smaller one. As the result, all vertices in a single SCC have the same color. Therefore, the edges between different colors can be removed and the resulting sub-graphs can be processed independently.

In contrast to the FW-BW algorithm, the coloring algorithm allows to split the graph into more parts. However, this implies a higher overhead.

2.4. Graph Generation

This work shows that the process of generating a graph to test the performance of an algorithm is of crucial significance. Ideally, an approach would be preferred which outputs uniformly random graphs while controlling the key parameters of the graph, i.e. its size, density, number of SCCs etc.

$G(n,p)$ and $G(n,M)$, as proposed by Erdos and Renyi [28], are some of the most widely used graph generation methods which can be harnessed for non-oriented graphs, as well as DAGs. Other methods include: layer-by-layer [29], fan-in/fan-out [30], and random orders [31].

While these algorithms provide some means for predefining general graph properties, such as the number of vertices and edges, for the case described in this work we also wanted to control other parameters, with a particular emphasis placed on the number of SCCs. To this end, a custom graph generation method has been employed.

3. Basic Concepts

We start with the definition of a directed graph:

Definition 1. A *directed graph* G is a pair (V, E), where V is a set of vertices, and $E \subseteq V \times V$ is a set of directed edges. If $(u, v) \in E$, then v is called immediate successor of u, and u is called immediate predecessor of v.

Definition 2. A *directed path* in a digraph is a sequence of vertices in which there is a (directed) edge pointing from each vertex in the sequence towards its successor in the sequence. A simple path is the one with no repeated vertices.

A *strongly connected component* (SCC) of a directed graph G is defined in the following way:

Algorithm 1 DetectSCCs

```
1: if G has no more then one vertex then return
2:     TRIM G in forward direction
3:     if G is not empty then
4:         TRIM G in backward direction
5:         select vertex v from G
6:         MARK Pred(G, v) and Desc(G, v) in G
7:         SCC(G, v) = Pred(G, v) ∩ Desc(G, v)
8:         Do in parallel
9:         detectSCCs(Pred(G, v) − SCC(G, v))
10:        detectSCCs(Desc(G, v) − SCC(G, v))
11:        detectSCCs(Rem(G, v))
12:    end if
13: end if
```

Definition 3. A *strongly connected component* of a directed graph G is a maximal set of vertices $C \subseteq V$ such that for every pair of vertices u and v, there is a directed path from u to v and a directed path from v to u.

A back-level edge is a concept used by one of the pivot selection algorithms presented in Section 2.

Definition 4. Let $G = (V, E)$ be a graph with the source vertex s. An edge $(u, v) \in E$ is called the *back-level edge* if and only if $d(u) >= d(v)$, where $d(u)$ and $d(v)$ are the distances from s to u and s to v, respectively. Vertex u is called the start vertex of the back-level edge, while v is called destination vertex.

4. Increasing Parallelism in FW-BW SCC Detection Algorithm

This section presents improvements to the SCC detection algorithms. We start with a description of the standard FW-BW algorithm and then discuss the pivot selection problem.

4.1. FW-BW Algorithm

Let us assume that we intend to discover all strongly connected components in a graph by means of computations performed asynchronously by a set of nodes. Following the work performed by McLendon [3], we can apply, to this end, a distributed algorithm performed at each node participating in the process (Algorithm 1).

The trim operations are performed in order to remove those vertices which do not belong to the SCC. The forward trim begins with all vertices that have no ancestors and removes them together with all their edges. After the removal, some other vertices may have no ancestors as well, and they are removed too. The process continues until no more vertices can be removed. More formally, the trim iteratively removes vertices whose in-degree is non-zero and whose out-degree is zero.

By analogy, the reverse trim performs the same operation in the other direction.

4.2. Pivot Selection

The problem of pivot selection is important for the described algorithm. Ideally, such a pivot should be picked each time that produces four sub-graphs at the end of the algorithm step, allowing for the highest possible branching of the recursion. However, if no ancestors or predecessors exist (or there are a few of them and all belong to the intersection of closures) for the particular SCC detected at the current step, then the branching factor will be smaller and fewer sub-processes will be started.

In most of the algorithms proposed so far, the pivot was selected randomly. The argument was that the optimal pivot selection process requires depth-first search post-ordering, which is rather difficult to perform in parallel. In this paper, we bring forward the observation that a useful heuristic may be applied by using the information already computed as one of the algorithm steps.

First, the reachability procedure is performed for two out of three sets generated by the FW-BW algorithm. This returns us the breadth first partial ordering over the vertices in these sub-graphs. This information can be used to avoid selecting, as pivots, those vertices which are likely to belong to an SCC with no descendants or predecessors. These will most likely be the vertices reached at the beginning or at the end of the process.

Let us analyze the example presented in Fig. 1. The sub-graph represents all vertices reachable from v_2. The possible order of visiting the vertices by the breadth-first search algorithm is shown in Fig. 2.

Since some of these vertices will belong to the SCC (darker color), they will be excluded from the set passed on to the next algorithm step, meaning they should not be considered for the pivot selection process.

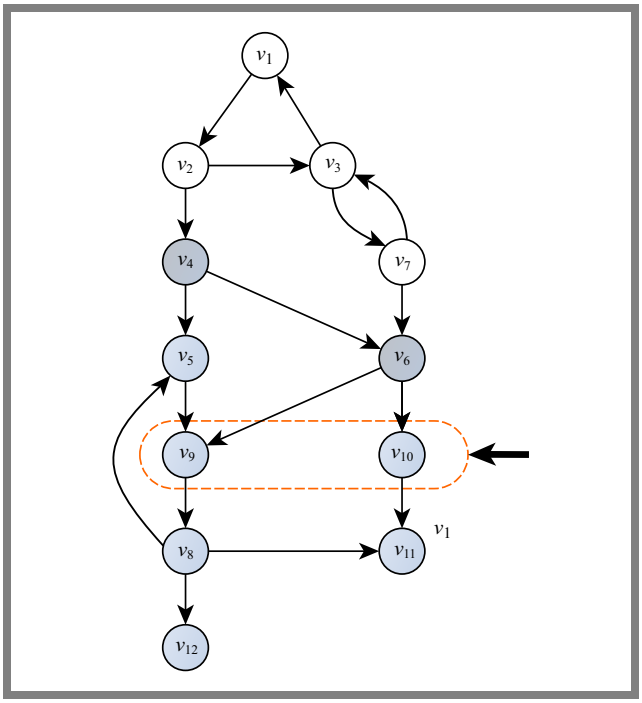

Fig. 1. Example graph.

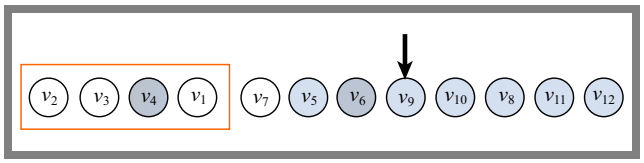

Fig. 2. Order of vertices visited by the breadth-first search algorithm.

Vertices located in the middle of the graph (after exclusion of SCC vertices), after breadth-first search layering, are most likely to have a sizable set of predecessors and ancestors. This layer is marked in Fig. 1. On the other hand, picking a vertex from the top or bottom layers (e.g. v_4 or v_{12}) leaves us with one sub-task empty.

Following the observation described above, we should ideally compute the middle layer of the relevant sub-graph and pick a vertex out of it. However, since a precise identification of these vertices can introduce some computational overhead, we have tested a simplified solution.

We assign sequential numbers to the vertices during the reachability procedure and then the index of the pivot is calculated as:

$$idx = \big(size(ForwardClosure) - size(SCC)\big) \text{ div } 2$$

or

$$idx = \big(size(BackwardClosure) - size(SCC)\big) \text{ div } 2,$$

respectively.

If the index points to a vertex which belongs to the SCC, which is possible with such a simplification, a random vertex is chosen.

Algorithm 2 presents the final version of the SCC detection including the improvements described above.

5. Experimental Results

To test the usability of the heuristic technique described in the Section 4, a specific test environment has been created. The standard FW-BW algorithm has been implemented using the Scala Akka library for parallel computations [32]. Then, the modification of pivot selection has been added. The test environment has been run on a multi-core workstation. It is important to note that the branching factor can be measured independently of the physical computational environment. It impacts the execution times only.

5.1. Graph Generation

To analyze the performance of the algorithm on graphs with different numbers and sizes of the SCCs, a dedicated method for test graph generation has been adopted. First, a spanning tree is generated over the initial set of vertices. Then, the vertices are divided equally (if possible) into a predefined number of sets and a predefined number of vertices is added within each set only. Finally, some extra edges are added between the sets.

Such an approach allows us to control the graph size, its density, as well as the number and size of SCCs. Due to the random character of some of the generation algorithm steps, some of the graph parameters, e.g. number of SCCs, may vary when given the same algorithm parameters. However, these variations are very small and do not affect the results of the experiments.

5.2. Experiments

In order to test the modified algorithm, it has been run for graphs of various sizes and SCC counts. It has been benchmarked against the standard algorithm. Figure 3 shows a comparison of the average branching factor for a graph containing 100k vertices, 400k edges and various SCC counts.

The scalability of the new algorithm that may be adapted to the growing graph size has been tested and is shown in Fig. 4.

Figure 5 shows the run time (in log scale) as a function of the graph size (vertices/SCCs), when run on a 4-core computer system. While the branching factor is independent of the environment, the run time still shows some improvement potential when a system with more cores is used. But the improvement is visible even on a 4-core system.

5.3. Experiment Outcomes

Next, the algorithm has been run for various graph sizes and various number of SCCs within the graph. The results prove the initial assumption that it is capable of increasing the branching factor. It scales according to graph size and the number of SCCs for a fixed number of vertices. As practically no additional computational overhead is present, the execution times are also consistently better.

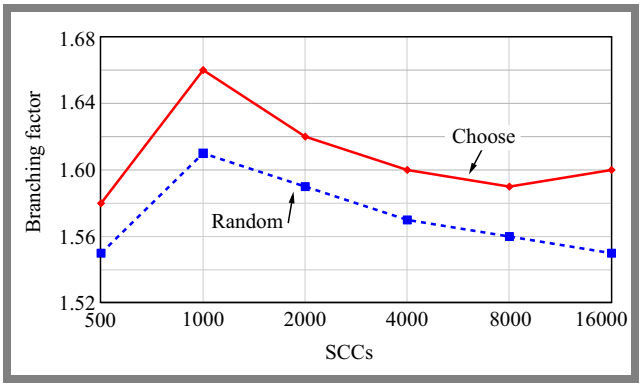

Fig. 3. Branching factor as a function of the number of SCCs.

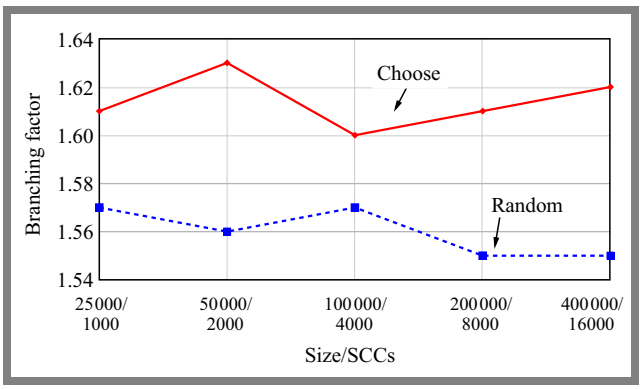

Fig. 4. Branching factor as a function of the size of the graph.

Algorithm 2 ImprovedDetectSCCs

Input: *pivot*

1: **if** G has no more then one vertex **then** return
2: TRIM G in forward direction
3: **if** G is not empty **then**
4: TRIM G in backward direction
5: **if** $pivot = \emptyset$ **then** select random vertex v from G
6: MARK $Pred\,(G, v)$ and $Desc\,(G, v)$ in G
7: $SCC(G, v) = Pred\,(G, v) \cap Desc\,(G, v)$
8: $FwdPivot = Desc\left[\left(size(Desc) - size(SCC)\right) \text{ div } 2\right]$
9: **if** $FwdPivot \in SCC$ **then** $FwdPivot = \emptyset$
10: $BckPivot = Pred\left[\left(size(Pred) - size(SCC)\right) \text{ div } 2\right]$
11: **if** $BckPivot \in SCC$ **then** $BckPivot = \emptyset$
12: **Do in parallel**
13: $detectSCCs\,(Pred(G, v) - SCC(G, v), BckPivot)$
14: $detectSCCs\,(Desc(G, v) - SCC(G, v), FwdPivot)$
15: $detectSCCs\,(Rem(G, v))$
16: **end if**
17: **end if**
18: **end if**
19: **end if**
20: **end if**

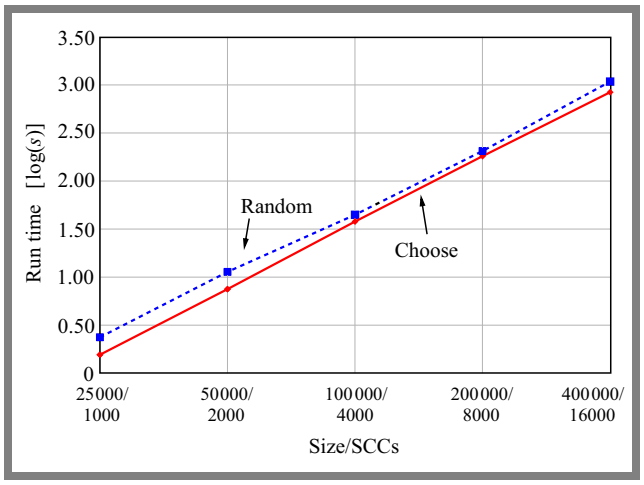

Fig. 5. Run time as a function of the size of the graph.

6. Conclusions

This paper presents a potential improvement to the pivot selection process being part of the FW-BW algorithm used for SCC decomposition. It has been shown that the branching factor of the algorithm may be increased based on the information already computed in the previous algorithm step, hence without introducing significant overhead.

Future work should involve blending this approach with other pivot selection techniques, such as the multipivot or degree-based methods which are also mentioned herein. It could be also interesting to find out if the heuristic introduced here may be relied upon to solve other problems involving pivot selection e.g. clique enumeration etc.

References

[1] B. Sandeep and A. Ferreira, "Complexity of Connected Components in Evolving Graphs and the Computation of Multicast Trees in Dynamic Networks", *Second International Conference, ADHOC-NOW2003*, Montreal, Canada, 2003 (https://doi.org/10.1007/978-3-540-39611-6_23).

[2] S. Raghavan, "Twinless Strongly Connected Components", in: *Perspectives in Operations Research*, pp. 285–304, 2006 (https://doi.org/10.1007/978-0-387-39934-8_17).

[3] W. McLendon III, B. Hendrickson, S.J. Plimpton, and L. Rauchwerger, "Finding Strongly Connected Components in Distributed Graphs", *Journal of Parallel and Distributed Computing*, vol. 65, no. 8, pp. 901–910, 2005 (https://doi.org/10.1016/j.jpdc.2005.03.007).

[4] H. Garavel, R. Mateescu, and I.M. Smarandache, "Parallel State Space Construction for Model-checking", *Proc. of the 8th International SPIN Workshop on Model Checking of Software (SPIN'01)*, vol. 2057, pp. 200–216, 2001 (https://doi.org/10.1007/3-540-45139-0_14).

[5] D. Ryzko, "Reasoning in Multi-agent Systems with Random Knowledge Distribution", in: *Advances in Databases and Information Systems*, pp. 310–317, 2012 (https://doi.org/10.1007/978-3-642-33074-2_23).

[6] P. Cholewinski, "Reasoning with Stratified Default Theories", *Proc. of Third International Conference on Logic Programming and Nonmonotonic Reasoning (LPNMR'95)*, pp. 273–286, 1995 (https://dl.acm.org/doi/10.5555/646397.691152).

[7] V. Batagelj, "Efficient Algorithms for Citation Network Analysis", *ArXiv*, 2003 (https://doi.org/10.48550/arXiv.cs/0309023).

[8] Y. Kajikawa *et al.*, "Creating an Academic Landscape of Sustainability Science: An Analysis of the Citation Network", *Sustainability Science*, vol. 2, pp. 221–231, 2007 (https://doi.org/0.1007/s11625-007-0027-8).

[9] H. Yang *et al.*, "Strongly Connected Components Based Efficient PPR Algorithms", *Jisuanji Xuebao/Chinese Journal of Computers*, vol. 40, pp. 584–600, 2017.

[10] N.P. Hummon and P. Doreian, "Computational Methods for Social Network Analysis", *Social Networks*, vol. 12, no. 4, pp. 273–288, 1990 (https://doi.org/10.1016/0378-8733(90)90011-W).

[11] R. Tarjan, "Depth First Search and Linear Graph Algorithms", *SIAM Journal on Computing*, vol. 1, no. 2, pp. 146–160, 1972 (https://doi.org/10.1137/0201010).

[12] H.N. Gabow, "Path-based Depth First Search Strong and Biconnected Components", *Information Processing Letters*, vol. 74, no. 3–4, pp. 107–114, 2000 (https://doi.org/10.1016/S0020-0190(00)00051-X).

[13] E. Renault, A. Duret-Lutz, F. Kordon, and D. Poitrenaud, "Parallel Explicit Model Checking for Generalized Buchi Automata", in: *Tools and Algorithms for the Construction and Analysis of Systems*, pp. 613–627, 2015 (https://doi.org/10.1007/978-3-662-46681-0_56).

[14] H. Gazit and G.L. Miller, "An Improved Parallel Algorithm that Computes the BFS Numbering of a Directed Graph", *Information Processing Letters*, vol. 28, no. 2, pp. 61–65, 1988 (https://doi.org/10.1016/0020-0190(88)90164-0).

[15] R. Cole and U. Vishkin, "Faster Optimal Parallel Prefix Sums and List Ranking", *Information and Computation*, vol. 81, no. 3, pp. 334–352, 1989 (https://doi.org/10.1016/0890-5401(89)90036-9).

[16] J. Barnat, J. Chaloupka, and J. van de Pol, "Improved Distributed Algorithms for SCC Decomposition", *Electronic Notes in Theoretical Computer Science*, vol. 198, no. 1, pp. 63–77, 2008 (https://doi.org/10.1016/j.entcs.2008.02.001).

[17] W. Schudy, "Randomized Algorithms for Graph Problems", 2007.

[18] W. Schudy, "Finding Strongly Connected Components in Parallel Using $O(\log^2 n)$ Reachability Queries", *Proceedings of the Twentieth Annual Symposium on Parallelism in Algorithms and Architectures*, pp. 146–151, 2008 (https://doi.org/10.1145/1378533.1378560).

[19] J. Jaiganesh and M. Burtscher, "A High-performance Connected Components Implementation for GPUs", *Proceedings of the 27th International Symposium on High-Performance Parallel and Distributed Computing*, pp. 92–104, 2018 (https://doi.org/10.1145/3208040.3208041).

[20] T. Ben-Nun, M. Sutton, S. Pai, and K. Pingali, "Groute: An Asynchronous Multi-GPU Programming Model for Irregular Computations", *Proceedings of the 22nd ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming*, pp. 235–248, 2017 (https://doi.org/10.1145/3018743.3018756).

[21] L. Fleischer, B. Hendrickson, and A. Pinar, "On Identifying Strongly Connected Components in Parallel", *Proceedings of Parallel and Distributed Processing Symposium*, pp. 505–511, 2000 (https://doi.org/10.1007/3-540-45591-4_68).

[22] S. Hong, N.C. Rodia, and K. Olukotun, "On Fast Parallel Detection of Strongly Connected Components (SCC) in Small-world Graphs", *Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis*, 2013 (https://doi.org/10.1145/2503210.2503246).

[23] J. Barnat and P. Moravec, "Parallel Algorithms for Finding SCCs in Implicitly Given Graphs", in: *Formal Methods: Applications and Technology*, pp. 316–330, 2007 (https://doi.org/10.1007/978-3-540-70952-7_22).

[24] Y. Ji, H. Liu, and H.H. Huang, "iSpan: Parallel Identification of Strongly Connected Components with Spanning Trees", *SC18: International Conference for High Performance Computing, Networking, Storage and Analysis*, Dallas, USA, 2022 (https://doi.org/10.1109/SC.2018.00061).

[25] D. Niewenhuis and A.L. Varbanescu, "Efficient Trimming for Strongly Connected Components Calculation", *Proceedings of the 19th ACM International Conference on Computing Frontiers*, pp. 131–140, 2022 (https://doi.org/10.1145/3528416.3530247).

[26] J. Barnat, J. Chaloupka, and J. van de Pol, "Distributed Algorithms for SCC Decomposition", *Journal of Logic and Computation*, vol. 21, no. 1, pp. 23–44, 2011 (https://doi.org/10.1093/logcom/exp003).

[27] S. Orzan, "On Distributed Verification and Verified Distribution", Ph.D. thesis, Free University of Amsterdam, 2004 (https://research.vu.nl/en/publications/on-distributed-verification-and-verified-distribution).

[28] P. Erdos and A. Renyi, "On Random Graphs I", *Publicationes Mathematicae Debrecen*, vol. 6, pp. 290–297, 1959

[29] T. Tobita and H. Kasahara, "A Standard Task Graph Set for Fair Evaluation of Multiprocessor Scheduling Algorithms", *Journal of Scheduling*, vol. 5, no. 5, pp. 379–394, 2002 (https://doi.org/10.1002/jos.116).

[30] R.P. Dick, D.L. Rhodes, and W. Wolf, "TGFF: Task Graphs for Free", *Proceedings of the 6th International Workshop on Hardware/Software Codesign*, pp. 97–101, 1998 (https://doi.org/10.1109/HSC.1998.666245).

[31] P. Winkler, "Random Orders", *Order*, vol. 1, pp. 317–331, 1985 (https://doi.org/10.1007/BF00582738).

[32] Akka, (http://akka.io).

Dominik Ryżko, Ph.D.
Institute of Computer Science
https://orcid.org/0000-0001-5632-0861
E-mail: dominik.ryzko@pw.edu.pl
Warsaw University of Technology, Warsaw, Poland
https://eng.pw.edu.pl/

A Generalized Learning Approach to Deep Neural Networks

Francesca Ponti[1], Fabrizio Frezza[1], Patrizio Simeoni[2], and Raffaele Parisi[1]

[1] *"Sapienza" University of Rome, Rome, Italy,*
[2] *South East Technological University, Carlow, Ireland*

https://doi.org/10.26636/jtit.2024.3.1454

Abstract — Optimization of machine learning architectures is essential in determining the efficacy and the applicability of any neural architecture to real world problems. In this work a generalized Newton's method (GNM) is presented as a powerful approach to learning in deep neural networks (DNN). This technique was compared to two popular approaches, namely the stochastic gradient descent (SGD) and the Adam algorithm, in two popular classification tasks. The performance of the proposed approach confirmed it as an attractive alternative to state-of-the-art first order solutions.
Due to the good results presented in the case of shallow DNN, in the last part of the article an hybrid optimization method is presented. This method consists in combining two optimization algorithms, i.e. GNM and Adam or GNM and SGD, during the training phase within the layers of the neural network. This configuration aims to benefit from the strengths of both first- and second-order algorithms. In this case a convolutional neural network is considered and its parameters are updated with a different optimization algorithm. Also in this case, the hybrid approach returns the best performance with respect to the first order algorithms.

Keywords — deep neural networks, machine learning, optimization

1. Introduction

In recent years, neural deep learning (NDL) [1]–[5] has been acquiring popularity in many applicative fields of engineering, such as system identification and control, robotics, communication systems and signal processing [6]. As a matter of fact, deep neural networks (DNNs) are able to handle large amounts of data that would be hard to treat by more traditional techniques.

A neural network (NN) is generally composed by a set of parameters (the *weights*) that are updated iteratively in order to establish the desired correspondence between the input and the output. This phase is called *training* and it is performed by means of properly selected optimization algorithms.

In general, optimization consists in the minimization or maximization of a function subject to specific constraints on its variables [7]. In the NN context, optimization corresponds to finding the optimal weights $\mathrm{bf}w^*$ that minimize a given function $E(\mathrm{bf}w)$ (the *loss function*):

$$\mathrm{bf}w^* = \underset{\mathrm{bf}w}{\operatorname{argmin}}\, E(\mathrm{bf}w)\ . \tag{1}$$

Several choices for the objective function $E(\mathrm{bf}w)$ are possible. Due to the high degree of non-linearity of $E(\mathrm{bf}w)$, minimization is performed by employing iterative techniques and it is not always guaranteed that the obtained minima are optima (*local minima*).

Optimization methods can be roughly grouped in two main classes [8], [9]: first and second-order methods.

First order methods are the most commonly applied in practice. They involve the use of the first derivative of the objective function with respect to the weights in order to choose the direction of movement in the search space.

First-order methods are quite popular in machine learning, since they have a low computational cost and are usually easy to implement. However, as highlighted in [10], first-order methods have several shortcomings, since they depend on hyper-parameter selection, they are extremely hard to tune and parallel computing opportunities are often limited.

In order to find a more effective alternative to first-order methods and to overcome their limitations, second-order methods [7] have been often employed in traditional shallow neural networks. They resulted to work well in several machine learning tasks [10]–[13]. The standard second-order technique is the *Newton's method*, that makes use of the Hessian matrix, formed by the second-order derivatives of the error functional, and has a quadratic rate of convergence. However, the Newton's method requires exact computation of the Hessian matrix, which is computationally expensive.

For this reason quasi-Newton's methods have been introduced [14]. They are based on different approximations of the inverse of Hessian matrix, with the aim of combining the rate of convergence of Newton's method with the scalability typical of first-order methods.

Moreover, another issue that affects machine learning, and in particular deep learning algorithms, are saddle points, which lead to reaching non-optimal minimum points.

Recently, many studies are proving that in many cases second order methods are more efficient than first order methods at avoiding saddle points [15], [16] reversing the traditional hypothesis that proposes as more performing the first order methods [17].

In this work we propose a general learning framework that belongs to the class of quasi-Newton's methods. This approach was introduced as a powerful learning algorithm for feedfor-

This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) License.
For more information, see https://creativecommons.org/licenses/by/4.0/

ward and recurrent NNs [18] and it was shown to belong to the class of generalized Newton's methods (GNMs).

We introduce its application to DNNs and compare it to most common optimization techniques in two deep-learning benchmark classification tasks. Moreover, due to the good performances achieved for the classification of a deep shallow neural network, an hybrid optimization approach is proposed to train a convolutional neural network (CNN). This method consists in combining two optimization algorithms, i.e. GNM and Adam or GNM and SGD, during the training phase within the layers of the neural network.

The proposed hybrid optimization approach aims to combine the speed of convergence of the second order optimization method under examination, i.e. GNM, with the low computational complexity of first order method that is required to train convolutional layers.

Indeed, as already stated, the algorithm proposed was tested in a feed-forward neural networks in the past years, when deep learning had not yet taken hold.

Then, testing the performance of existing optimization algorithms that have not been previously used on deep neural networks brings novelty by exploring unexplored territory, potentially improving performance, enabling comparisons against state-of-the-art methods, broadening applicability, and driving methodological advancements.

This paper contains a review of SGD and Adam methods in Section 2. In the same section, the description of the generalized Newton method under examination is reported. The experimental results obtained in two classification tasks, that consider MNIST and SUSY dataset, are described in Section 3. Section 4 describes the hybrid approach proposed and shows its result in case of a CNN. The conclusions are reported in Section 5.

2. Existing Optimization Algorithms

2.1. Stochastic Gradient Descent

Gradient descent (GD) [19] is a very popular iterative first-order method. At the generic k-th iteration it is represented by the following equation:

$$\mathrm{bf}w_{k+1} = \mathrm{bf}w_k - \eta \nabla_{\mathrm{bf}w_k} E \,, \tag{2}$$

where the *gradient* of the error functional with respect to the weights $\nabla_{\mathrm{bf}w_k} E$ is employed to compute the movement *in the weight space*. η (the *step length* or *learning rate*) is usually selected empirically and determines the rate of convergence of the loss function. This parameter represents how much the innovation term $\nabla_{\mathrm{bf}w} E$ influences the new weights with respect to the old ones.

The stochastic gradient descent (SGD) [20] is a modification of the GD based on an approximation of the actual gradient. Application of the SGD to NNs led to the well-known back-propagaton algorithm [21], [22]. One of the main advantages of SGD is its computational efficiency. It processes training data in small batches, making it well-suited for large datasets. SGD also has a low memory footprint since it only requires storing a small batch of data at a time. Additionally, SGD can converge faster compared to batch gradient descent as it quickly updates the model parameters using each batch of data.

However, SGD has some limitations that need to be considered. One major drawback is its susceptibility to getting stuck in local minima due to the high variance in the stochastic gradients. This can lead to suboptimal solutions.

SGD also requires careful tuning of the learning rate, as a too high learning rate can cause divergence and a too low learning rate can result in slow convergence. Additionally, SGD may struggle with handling noisy or sparse gradients, requiring the use of additional techniques such as learning rate schedules or momentum.

2.2. Adam Optimizer

One of the most popular first-order optimization algorithms in DNNs is the Adam optimizer [23]. Adam is an adaptive learning rate method, in the sense that it adapts the learning rate for different parameters by computing the first and second moments of the weight gradients.

The Adam weight update formula is expressed as follows:

$$w_k = w_{k-1} - \eta \frac{\hat{m}_{k+1}}{\sqrt{\hat{v}_k} + \epsilon} \tag{3}$$

where w_k is the weight at the k-th iteration and $\hat{m}_k$ and $\hat{v}_k$ the bias corrected estimators of the moving averages.

The advantages of the Adam optimizer lie in its adaptive learning rate, efficiency in handling sparse gradients, momentum optimization, robustness to hyperparameters, and wide applicability. These features contribute to its effectiveness and popularity in the field of deep learning.

However, there are some drawbacks to consider. Adam requires additional memory and computation resources due to its storage and update of past gradients and squared gradients. It can be sensitive to the choice of learning rate, and setting it improperly may lead to convergence issues.

2.3. Generalized Newton's Method

Here we briefly recall the main concepts related with the GNM presented in [18]. As stated in [18], a feedforward NN is a sequence of layers, each one composed by a linear and non-linear block [2], [4], [5] (see Fig. 1).

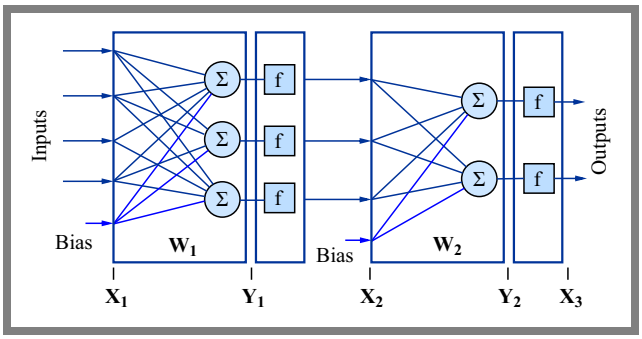

Fig. 1. Neural network representation.

At the k-th iteration of the learning procedure, in the forward propagation phase the output of the linear block of the generic n-th layer ($n = 1, \ldots, L$) is computed with the following equation:

$$\mathbf{Y}_k^n = \mathbf{X}_k^n \mathbf{W}_k^n \tag{4}$$

In Eq. (4) each row of matrices $\mathbf{X}_k^n$ and $\mathbf{Y}_k^n$ refers to the generic learning pattern, while the generic column of matrix $\mathbf{W}_k^n$ contains the weights leading to the generic output neuron.

The non-linear block computes the output of the non-linear part, that is the input to the next $(n+1)$-th layer:

$$\mathbf{X}_k^{n+1} = [f\{\mathbf{Y}_k^n\}\ \mathbf{1}], \tag{5}$$

where f represents the activation function and $\mathbf{1}$ is the column of the bias inputs. $\mathbf{X}_k^{L+1}$ is the *global* output of the network.

Considering the generic n-th layer ($n = 1, \ldots, L$) of the NN, the GNM employs the following formula to compute the new weights at the next $(k+1)$-th iteration:

$$\mathbf{W}_{k+1}^n = \mathbf{W}_k^n - \eta(\mathbf{X}_k^n)^+ \nabla_{\mathbf{Y}_k^n} E, \tag{6}$$

where $(\mathbf{X}_k^n)^+$ is the *pseudoinverse* of matrix $\mathbf{X}_k^n$.

Introducing the estimated correlation matrix of the inputs to the n-th layer $\Phi_k^n = (X_k^n)^T X_k^n$ and relating the *gradient matrix in the neuron space* $\nabla_{\mathbf{Y}_k^n} E$ with the *gradient matrix in the weight space* $\nabla_{\mathbf{W}_k^n} E$ it can be shown that:

$$\mathbf{W}_{k+1}^n = \mathbf{W}_k^n - \eta(\mathbf{\Phi}_k^n)^+ \nabla_{\mathbf{W}_k^n} E. \tag{7}$$

Equation (7) represents a *modified Newton's method* applied to the generic n-th layer. In [18] it is shown that, extending the previous formula to the whole network and comparing to the Newton's formula, the local Hessian matrix H_k (that is the Hessian related with the single layer) is approximated by $diag(\Phi_k)$. Since Φ_k is a diagonal matrix, this approximation assumes that the second-order partial derivatives with respect to the weights belonging to different layers are zero, that is neurons with respect to different layers are decoupled.

Moreover, it is well known that the Hessian matrix may not satisfy the property of positive definiteness. In contrast, the matrix Φ_k satisfies always the property of semipositive definiteness and almost always the property of positive definiteness.

This feature is important as it ensures that the algorithm assumes a descent direction. Indeed, it represents one of the main advantages of the considered algorithm. However, as it will be described in Section 4, the weights update Eq. (6) depends on the pseudoinverse of the input vector of each layer of the neural network. This type of operation could be computationally expensive in case of convolutional neural networks, in which tensors are involved.

This limitation is overcomed by the proposal of a hybrid optimization approach that combines the benefits and potential of both first and second-order methods to achieve higher performance.

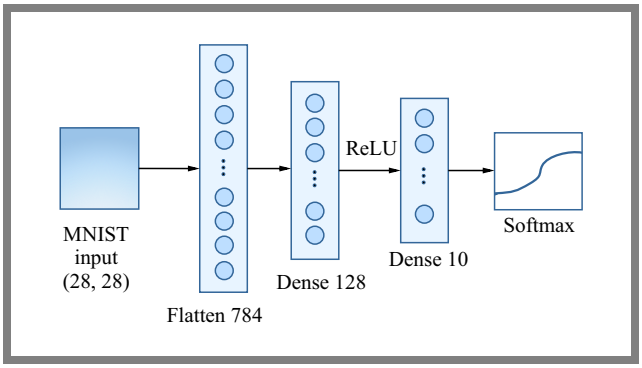

Fig. 2. Neural network architecture for MNIST classification.

3. Experimental Results

In this section we describe the results obtained in two typical experimental tasks considering two shallow NNs.

3.1. MNIST Dataset Classification Task

The first classification task used the MNIST dataset [24]. MNIST is a dataset of 60 000 squared 28×28 pixel grayscale images of handwritten single digits between 0 and 9 [25]. The aim of the network is to classify an image of a handwritten digit into one of 10 classes corresponding to the integer values from 0 to 9. A neural network with three layers was considered (Fig. 2).

The first layer in this network, called *flatten layer*, transforms the images from a two-dimensional array (28×28 pixels) to a one-dimensional array (28×28 = 784 pixels). This layer has no parameters to learn, it only reshapes the data.

In cascade with the flatten layer, the network is made of a sequence of two dense or *fully connected* layers. The first dense layer has 128 nodes (or neurons), with activation function *ReLU*. The last layer returns a logit array of length 10 and uses the activation function *softmax*. Each node of the last layer returns the probability that the current image belongs to one of the 10 classes.

The problem taken into account is an example of *multi-class classification task*. In order to evaluate the performance, the categorical cross entropy loss function has been considered:

$$E = -\sum_{i=1}^{10} y_i log(\hat{y}_i), \tag{8}$$

where $\hat{y}_i$ is s the i-th output of the network while y_i is the corresponding target value.

The generalized Newton's method (GNM) [18] was compared to the SGD and the Adam optimizer. The training process was iterated 20 times, considering a batch size of different lengths. Figure 3 shows the results obtained with batches of length 32, 64, and 128 in the case of the GNM optimizer. The batch size of 32 reached the lowest loss value and was selected in all subsequent tests. Same results were obtained with the SGD [19], [20] and Adam [23] optimizers.

In all cases weights were initialized by using a random uniform distribution between –1 and 1. The learning rate η was empirically selected. In particular, η was set to 0.001 for the

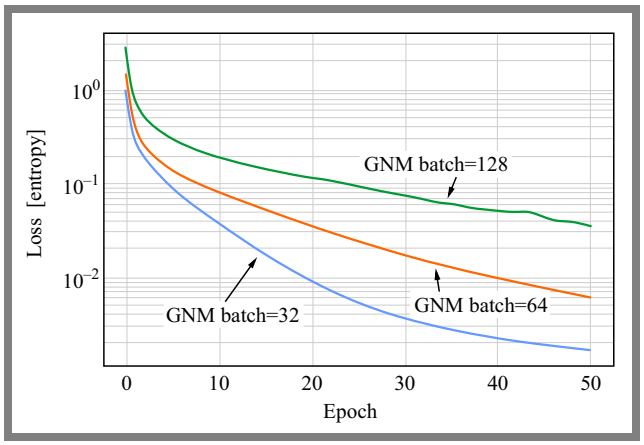

Fig. 3. MNIST dataset: GNM loss function comparison with respect to batch size.

Adam and SGD algorithms, while for the GNM approach η = 16 was considered. All simulations were carried out using the TensorFlow library [26]. Of course, since the GNM optimizer is not included in the default library, it has been defined and integrated into the eager execution. In eager execution TensorFlow does not build graphs, and each operation is executed by Python immediately. This feature makes it adapt to research and experimentation.

Conversely, the graph execution builds graph as representation of tensor computation and evaluates the model. The selection of the number of epochs, for the current case and all the other cases under examination, was established basing on two main factors: the convergence, and computational speed. In particular, the number of epochs are selected based on observing when the loss performance metric stabilizes in terms of convergence, with the goal to ensure that the model has converged and reached its optimal performance.

Figure 4 shows the loss function as a function of the number of epochs in a semilogaritmic scale.

As expected, the GNM loss curve has a faster convergence with respect to the first order optimizers, i.e. Adam and SGD.

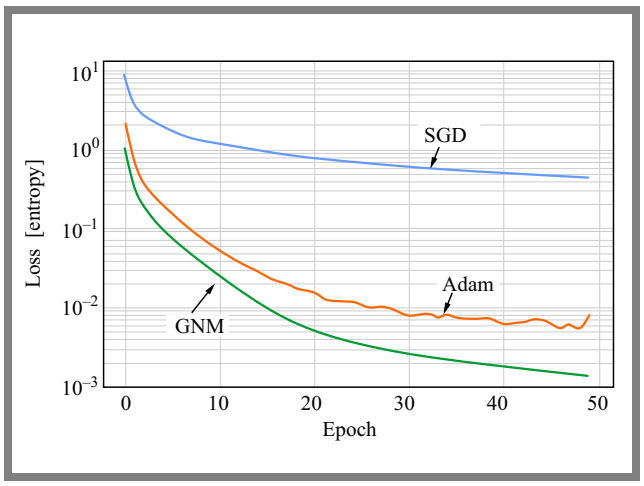

Fig. 4. MNIST dataset: GNM, SGD, and Adam loss function comparison with respect to epochs in semilogarithmic scale.

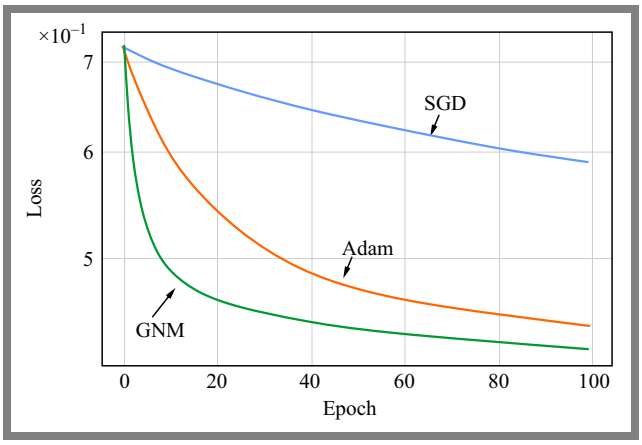

Fig. 5. SUSY dataset: GNM, SGD, and Adam loss function comparison with respect to epochs in semilogarithmic scale.

3.2. SUSY Dataset Classification Task

The second classification task is based on the SUSY dataset [27] and consists in recognizing the presence of supersymmetric particles in collisions at high energy colliders.

More specifically, the problem is a binary classification task where the goal is to establish whether each data point, generated via Monte Carlo simulations and characterized by 18 features (or kinematic variables), represents a signal *potential collision*, labeled with 1, or just *background*, labeled with 0. The first eight features are direct measurements of the kinematic features of final state particles in the accelerator.

The other ten features are high-level features derived from the first eight features and help in discriminating between the two classes. The whole dataset consists of 5 million points, of which 60 000 were considered as training set.

In this work, we propose a deep-learning architecture with the target of estimating the probability that an event is from a signal or a background process. The same approach was considered in [27]. The architecture considered is a neural network with four layers, having 256 units in the first layer, 128 in the second, 32 in the third layer, and one single unit in the last layer. This last layer uses a sigmoidal activation function while the other layers use the hyperbolic tangent function.

For each learning algorithm, the learning rate was selected by considering the best performance in terms of learning rate and weight initialization. Also in this experiment, weights were initialized by a random uniform distribution in the range [–1, 1]. The learning rate selected for the Adam optimizer was 0.001, for the gradient descent was 0.1, and 500 000 for the generalized Newton method.

In order to justify this last choice, from Eq. (6) we observe that in this case the learning rate depends on the condition number of the input matrix, i.e. on the order of magnitude of the inputs.

Figure 5 shows the average loss function of the described algorithms with respect to the epochs.

Also in this case, the GNM loss function was characterized by higher rate of convergence with respect to the Adam and SGD optimizers.

4. Hybrid Optimization Approach

In this section the results obtained with an hybrid optimization approach are reported.

As described in the previous chapter, the GNM has shown the best performances with respect to Adam and SGD optimizers in terms of rate of convergence. Due to the good results obtained in the case of shallow deep neural networks, in this chapter an hybrid optimization method is presented to train convolutional neural networks.

This method consists in combining two optimization algorithms, i.e. GNM and Adam or GNM and SGD, during the training phase within the layers of the Neural Network. In this case a convolutional neural network is considered and its parameters are updated with a different optimization algorithm.

In particular, the convolutional layers are updated with first order optimization algorithms, while the last fully connected layers are updated with the GNM method. This choice aims to combine the speed of convergence of the second order optimization method under examination, i.e. GNM, with the low computational complexity of first order method that is required to train convolutional layers.

In fact, as already described, the weight update formula of the GNM method depends on the pseudoinverse of the input matrix of the layer. It follows that in the case of convolutional neural networks, where the trainable variables are tensors, it can be computationally expensive apply the Eq. (7) to update the weights. Conversely, for the last fully connected layers of the neural network this operation can be easily computed.

To the aim to test the performances of the GNM algorithm in the hybrid optimization approach proposed, a simple CNN to classify MNIST dataset has been build with a similar approach considered in [28].

As shown in the Fig. 6, the considered CNN is composed by a convolutional layer with 32 filters of dimension 3×3 with activation function ReLU, followed by a max pooling operation with filters of 2×2 dimension. Then, the fully connected layers are applied: after the flatten layer, a dense layer of 128 units is present. Finally, the last classification layer of 10 units follows with activation function softmax.

In Fig. 7 the performances of the hybrid approach considering SGD and GNM are compared with respect the performance obtained by considering the single GD optimizer to update all the parameters of the neural network. In this case, two different

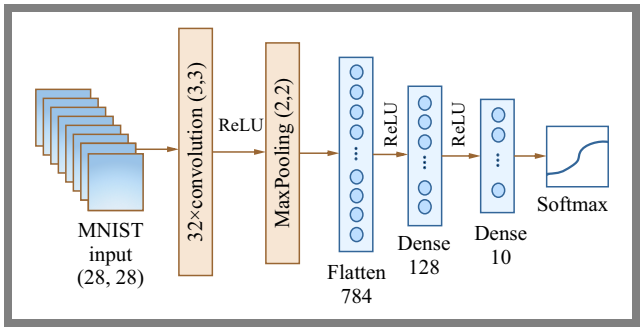

Fig. 6. MNIST dataset: convolutional neural network architecture.

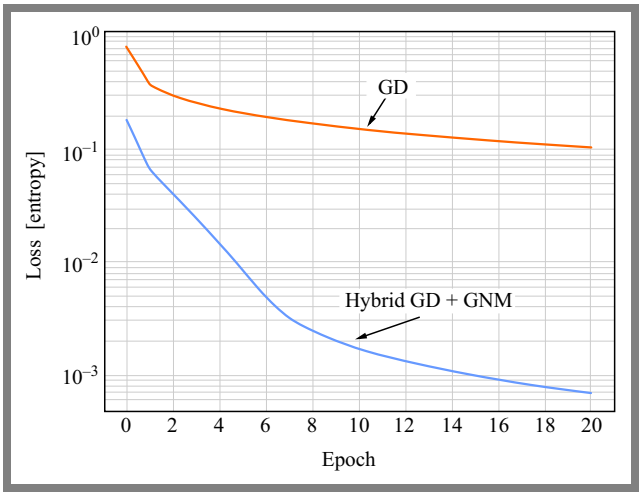

Fig. 7. MNIST dataset: comparison between fully GD and hybrid GD and GNM.

learning rates have been selected for the convolutional block and the fully connected layers of the neural network.

In particular, after a tuning activity aimed at obtaining the optimal η value, the learning rate for the convolutional layers trained with SGD has been fixed to 0.001, while the learning rate to update the fully connected layers and trained with GNM has been fixed to 17.

Indeed, as already highlighted, in case of GNM, the learning rate parameter depends on the condition number of the input matrix, i.e. on the order of magnitude of the inputs, see Eq. (6).

As can be seen from Fig. 7, the hybrid approach has shown the best performances in terms of rate of convergence in reaching the lower minima. This confirms the initial hypothesis, i.e. the combination of the second order method convergence rate with the scalability of first order methods would produce better performances.

In Fig. 8 the performances of the hybrid approach considering the combination of Adam and GNM optimizer are compared with respect to applying the single Adam optimizer to whole neural network. Also in this case, two different learning rates have been selected for the convolutional block and the fully connected layers of the neural network.

In particular, after a tuning activity, the learning rate for the convolutional layers trained with Adam has been fixed to 0.001, while the learning rate to update the fully connected layers and trained with GNM has been fixed to 17.

Also in this case, the hybrid approach has shown the best performances in terms of rate of convergence to reach the lower minima.

In Fig. 9 the performances of the hybrid approach considering Adam and SGD in combination with GNM in the fully connected layers are compared. One may notice, the combination of Adam and GNM optimizer shows the best performances in terms of rate of convergence in reaching the lower minima.

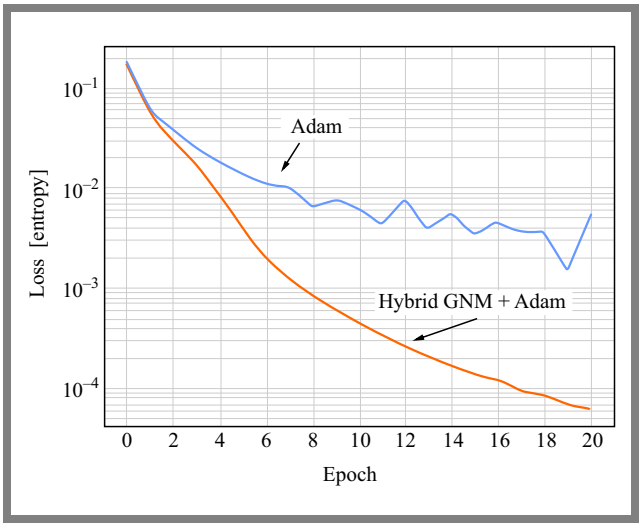

Fig. 8. MNIST dataset: comparison between fully Adam, hybrid Adam, and GNM.

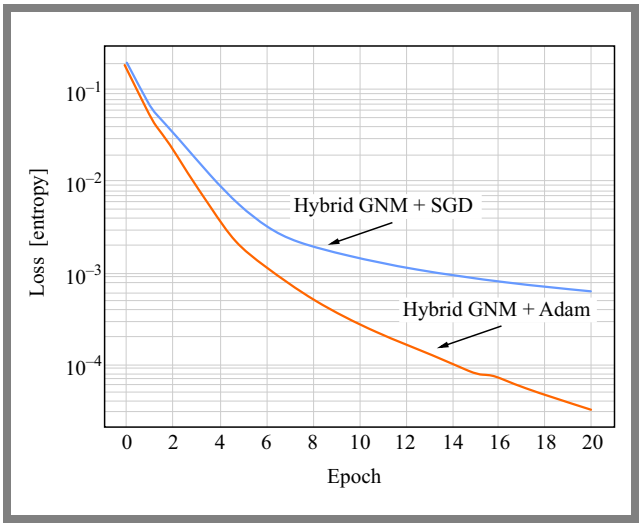

Fig. 9. MNIST dataset: comparison between hybrid approach using Adam and SGD in the convolutional layer.

5. Conclusions

In this work we presented a generalized optimization algorithm for feedforward deep neural networks and compared its performance to those of most popular algorithms in machine learning. Two popular classification tasks were considered, namely based on the MNIST and SUSY datasets. In both cases, the proposed approach showed optimal results in terms of speed of convergence, thus offering a viable alternative to most used first order methods in deep learning problems.

Given the good performances obtained with the two classification tasks considered, an hybrid optimization approach has been proposed. This new approach is configured as a powerful optimization method, as it combines the high rate of convergence of the GNM method with the scalability typical of first order ones. Also this case and as expected, the hybrid approach has shown an higher rate of convergence with respect of the standard one that considers only Adam or SGD optimizer for the training phase.

As future studies, it is essential to explore the performance of the optimizer on more diverse and challenging datasets, as well as with larger-scale deep learning architectures. This will help validate its effectiveness in real-world scenarios and provide insights into its generalizability.

Additionally, incorporating comparative studies with other state-of-the-art optimizers commonly used in deep learning can offer a more comprehensive understanding of its strengths and limitations.

In conclusion, the tested optimizer has demonstrated promising results in training deep shallow neural networks, showcasing its potential to improve convergence and achieve higher accuracy.

By addressing future challenges and exploring its application to real-world problems, the proposed method holds considerable prospects for advancing the field of deep learning and finding practical solutions in diverse domains.

References

[1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning", *Nature*, vol. 521, pp. 436–444, 2015 (https://doi.org/10.1038/nature14539).

[2] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016 (http://www.deeplearningbook.org).

[3] Y. Bengio, Y. LeCun, and G. Hinton, "Deep Learning for AI", *Communications of the ACM*, vol. 64, no. 7, pp. 58–65, 2021 (https://doi.org/10.1145/3448250).

[4] C.M. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press, 502 p., 1996 (ISBN: 9780198538646).

[5] R.D. Reed and R.J. Marks, *Neural Smithing: Supervised Learning in Feedforward Artificial Neural Networks*, MIT Press, 1999 (https://doi.org/10.7551/mitpress/4937.001.0001).

[6] R. Parisi, E.D. Di Claudio, G. Orlandi, and B.D. Rao, "Fast Adaptive Digital Equalization by Recurrent Neural Networks", *IEEE Transactions on Signal Processing*, vol. 45, no. 11, pp. 2731–2739, 1997 (https://doi.org/10.1109/78.650099).

[7] R. Battiti, "First- and Second-order Methods for Learning: Between Steepest Descent and Newton's Method", *Neural Computation*, vol. 4, no. 2, pp. 141–166, 1992 (https://doi.org/10.1162/neco.1992.4.2.141).

[8] L. Bottou, F.E. Curtis, and J. Nocedal, "Optimization Methods for Large-scale Machine Learning", *SIAM Review*, vol. 60, no. 2, pp. 223–311, 2018 (https://doi.org/10.1137/16M1080173).

[9] J. Nocedal and S.J. Wright, *Numerical Optimization*, Springer, 664 p., 2006 (https://doi.org/10.1007/978-0-387-40065-5).

[10] A.S. Berahas, M. Jahani, P. Richtárik, and M. Takác, "Quasi-Newton Methods for Machine Learning: Forget the Past, Just Sample", *arXiv*, 2019 (https://doi.org/10.48550/ARXIV.1901.09997).

[11] D. Goldfarb, Y. Ren, and A. Bahamou, "Practical Quasi-Newton Methods for Training Deep Neural Networks", *arXiv*, 2020 (https://doi.org/10.48550/arXiv.2006.08877).

[12] A.S. Berahas, R. Bollapragada, and J. Nocedal, "An Investigation of Newton-Sketch and Subsampled Newton Methods", *Optimization Methods and Software*, vol. 35, no. 4, pp. 661–680, 2020 (https://doi.org/10.1080/10556788.2020.1725751).

[13] A.S. Berahas and M. Takác, "A Robust Multi-batch L-BFGS Method for Machine Learning", *Optimization Methods and Software*, vol. 35, no. 1, pp. 191–219, 2020 (https://doi.org/10.1080/10556788.2019.1658107).

[14] J.E. Dennis, Jr. and J.J. Moré, "Quasi-Newton Methods, Motivation and Theory", *SIAM Review*, vol. 19, no. 1, pp. 46–89, 1977 (https://doi.org/10.1137/1019005).

[15] Z. Yao *et al.*, "ADAHESSIAN: An Adaptive Second Order Optimizer for Machine Learning", *arXiv*, 2020 (https://doi.org/10.48550/ARXIV.2006.00719).

[16] R. Anil *et al.*, "Scalable Second Order Optimization for Deep Learning", *arXiv*, 2020 (https://doi.org/10.48550/ARXIV.2002.09018).

[17] J.D. Lee *et al.*, "First-order Methods Almost Always Avoid Saddle Points", *arXiv*, 2017 (https://doi.org/10.48550/ARXIV.1710.07406).

[18] R. Parisi, E.D. Di Claudio, G. Orlandi, and B.D. Rao, "A Generalized Learning Paradigm Exploiting the Structure of Feedforward Neural Networks", *IEEE Transactions on Neural Networks*, vol. 7, no. 6, pp. 1450–1460, 1996 (https://doi.org/10.1109/72.548172).

[19] S. Ruder, "An Overview of Gradient Descent Optimization Algorithms", *arXiv*, 2016 (https://doi.org/10.48550/ARXIV.1609.04747).

[20] H. Robbins and S. Monro, "A Stochastic Approximation Method", *The Annals of Mathematical Statistics*, vol. 22, no. 3, pp. 400–407, 1951 (https://doi.org/10.1214/aoms/1177729586).

[21] R. Rojas, *Neural Networks. A Systematic Introduction*, Springer, 504 p., 2006 (https://doi.org/10.1007/978-3-642-61068-4).

[22] D.E. Rumelhart and J.L. McClelland, "Learning Internal Representations by Error Propagation", in: *Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Foundations*, MIT Press, pp. 318–362, 1987 (ISBN: 9780262291408).

[23] D.P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization", *arXiv*, 2014 (https://doi.org/10.48550/ARXIV.1412.6980).

[24] J.D. Lee *et al.*, "Basic Classification: Classify Images of Clothing" (https://www.tensorflow.org/tutorials/keras/classification).

[25] Y. LeCun, C. Cortes, and C.J.C. Burges, *The MNIST Database of Handwritten Digits*, 2012 (http://yann.lecun.com/exdb/mnist/).

[26] M. Abadi *et al.*, "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems", *arXiv*, 2016 (https://doi.org/10.48550/ARXIV.1603.04467).

[27] P. Baldi, P. Sadowski, and D. Whiteson, "Searching for Exotic Particles in High-energy Physics with Deep Learning", *Nature Communications*, vol. 5, art. no. 4308, 2014 (https://doi.org/10.1038/ncomms5308).

[28] D.-Y. Ge *et al.*, "Design of High Accuracy Detector for MNIST Handwritten Digit Recognition Based on Convolutional Neural Network", *2019 12th International Conference on Intelligent Computation Technology and Automation (ICICTA)*, Xiangtan, China, 2019 (https://doi.org/10.1109/ICICTA49267.2019.00145).

Francesca Ponti, Ph.D.
Department of Information Engineering, Electronics and Telecommunications
https://orcid.org/0000-0003-1855-7280
E-mail: francesca.ponti@uniroma.it
"Sapienza" University of Rome, Rome, Italy
https://www.uniroma1.it/en/

Fabrizio Frezza, Prof.
Department of Information Engineering, Electronics and Telecommunications
https://orcid.org/0000-0001-9457-7617
E-mail: fabrizio.frezza@uniroma1.it
"Sapienza" University of Rome, Rome, Italy
https://www.uniroma1.it/en/

Patrizio Simeoni, Ph.D.
Department of Electronic and Communications Engineering
https://orcid.org/0009-0008-7365-7624
E-mail: patrizio.simeoni@setu.ie
South East Technological University, Carlow, Ireland
https://www.setu.ie

Raffaele Parisi, Prof.
Department of Information Engineering, Electronics and Telecommunications
https://orcid.org/0000-0002-6062-0274
E-mail: raffaele.parisi@uniroma1.it
"Sapienza" University of Rome, Rome, Italy
https://www.uniroma1.it/en/

A Review of Isolation Techniques for 5G MIMO Antennas

Parminder Kaur, Shivani Malhotra, and Manish Sharma

Chitkara University, Punjab, India

https://doi.org/10.26636/jtit.2024.3.1490

Abstract — This paper offers an analysis of mutual coupling reduction techniques used in MIMO antennas designed for sub-6 GHz, 28 GHz, and 28/38 GHz dual frequency bands which are allocated to 5G technology. The said techniques take into account size, gain, isolation, and all diversity-related parameters, such as envelope correlation coefficient (ECC), directive gain (DG), and channel capacity loss (CCL). A review of current technologies is presented in the paper too. The isolation techniques are studied in detail and comparisons between the various works are drawn. Finally, the best isolation technique suitable for specific bands, applications and different port numbers is determined.

Keywords — 5G, isolation, MIMO antenna, mutual coupling

1. Introduction

MIMO (multiple-input multiple-output) is a wireless communication technique capable of simultaneous transmission and reception of many signals over a single frequency channel. Such a method boosts data transfer speeds, expands network capacity, and improves signal quality by using numerous antennas at the transmitter and receiver ends alike. The benefits of MIMO technology include the following:

- higher data rates due to simultaneous broadcasting and receiving of multiple data streams,
- better signal quality, as many antennas allow to minimize the effects of multipath fading and signal distortion,
- longer range achieved by decreasing signal degradation caused by buildings and multipath interference. The MIMO technology may increase the range of wireless networks,
- increased network capacity due to ability to serve, simultaneously, a higher number of users and devices.

While the MIMO approach offers highly advanced features, it also suffers from some drawbacks, as it adds complexity to wireless systems by requiring more antennas and by relying on complicated signal processing techniques, making the solution more challenging to design and deploy. It also adds more cost to the system due to the additional hardware and software needed when relying on the MIMO technology.

Last but not least, it is subject to interference or suffers from mutual coupling between MIMO radiating elements, a phenomenon which degrades signal quality and lowers data throughput rates. Interference caused by other wireless devices using the same frequency band is a factor too.

"Mutual coupling" is a term that refers to a phenomenon in which the radiating antennas used in a MIMO array interfere with each other, degrading the performance of the entire system. This interference occurs due to the electromagnetic waves emitted by one antenna and affecting the other aerials in the array. The causes of mutual coupling in MIMO can be attributed to several factors, such as the placement of antennas in the array, the distance between them, the type, and the orientation of the antennas used.

Three types of mutual coupling may be distinguished [1]:

1) Near-field coupling – the type of coupling that occurs between antennas when they are positioned close to each other. It is characterized by strong electromagnetic fields that can cause interference and signal distortion.
2) Surface wave coupling occurs when MIMO antennas are placed on a conductive surface, such as a metallic plate or a ground plane. This type of coupling is characterized by the propagation of surface waves that may deform the signal. A dispersed current is produced on the ground plane while the antenna element is fed at the feed point. Some energy will be distributed and interference will be produced as a result of this current flowing to the surrounding antenna components.
3) Free space coupling occurs when MIMO antennas are located in the free space and are not subject to any interference or distortion caused by external factors. When electromagnetic waves are radiated outward from the excitation element, some of the energy will be linked to the nearby antennas, thus inducing undesired current.

There are several ways to reduce mutual coupling in MIMO systems. Some of them include increasing the distance between adjacent antennas, using aerials with different radiation patterns and impedance characteristics, exploiting polarization diversity, implementing isolation techniques – such as adding absorptive materials or placing a shield between the MIMO antennas – and optimizing the antenna design by deploying aperture coupling or antenna decoupling techniques, as well as by using frequency-selective structures.

Common types of isolation enhancement techniques used in MIMO include the following.

The polarization isolation technique involves using antennas with different polarization orientations to reduce interference. For instance, if one antenna transmits a horizontally polarized wave, the other antenna may transmit a vertically polarized

This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) License.
For more information, see https://creativecommons.org/licenses/by/4.0/

signal. This helps reduce signal crosstalk and improves the isolation factor.

The spatial isolation approach involves designing a suitable physical layout of the antennas to help reduce interference. For instance, spacing the antennas further apart or using directional antennas may increase the spatial separation between the antennas, thus reducing mutual interference.

The frequency isolation technique relies on antennas that operate in different frequency bands, which helps reduce mutual coupling and also improves isolation. Interference between antennas is highest when they operate at the same frequency, but frequency-isolated antennas can help mitigate this problem.

In the beamforming technique, the transmitted radiation patterns of the antennas are shaped to reduce interference. Beamforming can be used to steer transmissions away from interfering antennas – an approach which can improve the isolation factor.

Due to its specific usage conditions, the antenna switching technique is also capable of helping improve isolation between antennas.

Overall, these isolation enhancement techniques are vital to the performance of MIMO communication systems, as they reduce signal crosstalk and improve the channel's signal-to-noise ratio, thus leading to higher throughput and data rates.

The specific MIMO isolation or decoupling design approaches may be categorized into two groups: external decoupling techniques and internal decoupling techniques [1]. External decoupling approaches are further divided into neutralization lines (NL), triple line, slots, stubs, parasitic decoupling elements (PDE), defective ground structures (DGS), and those using metamaterials, i.e. relying on the insertion of external structures into the patch or the ground to counteract the mutual coupling effect.

Internal decoupling techniques include space diversity and orthogonal polarization or multi-polarization about polarization diversity.

The remaining techniques used in MIMO antennas are described below.

Neutralization line is a metallic slit used to pass the electromagnetic waves between the elements of the antenna to reduce mutual coupling or improve isolation between the elements. It also improves the bandwidth and reduces antenna size. Electromagnetic bandgap (EBG) structures are made of metallic or dielectric materials with periodic arrangements for the transmission of electromagnetic waves. The periodic structures help achieve independent resonance and attain multiple band gaps, hence reducing mutual coupling and providing high efficiency.

Defected ground structure (DGS) is a type of ground plane etched with defects, slots, or slits to improve efficiency, isolation, and achieve a wide bandwidth, while the antenna with a dielectric resonator is characterized by high radiation efficiency, wide impedance bandwidth, high gain, good isolation, small form factor, and low losses.

Metamaterials have special electromagnetic, isotropic, anisotropic, chiral, photonic frequency-selective surface-based, non-linear, and tunable properties. These features offer enhancements in terms of diversity gain, envelope correlation coefficient (ECC), and wide bandwidth.

Complimentary split ring resonators (CSRRs) consist of two concentric rings with slots opposite each other. They are used to enhance isolation and efficiency and improve diversity gain as well as to reduce the size of the antenna.

The remaining methods relied upon to improve performance and isolation include slots and stubs used to achieve impedance matching, wide bandwidth, high efficiency, and higher gain. Additional parasitic elements around the primary antenna elements mitigate the coupling effect.

The self-isolation method does not involve any external decoupling. The antennas are isolated by maintaining a longer distance between them. Such a method is usually used at sub-6 GHz and mmWave bands. The self-isolation characteristic is often combined with different orientation of antenna elements to achieve spatial and polarization diversities.

In MIMO systems, spatial diversity is used to increase the capacity and reliability of wireless communication through the use of multiple antennas. By transmitting independent signals via each of the antennas, a MIMO system can effectively use the available radio spectrum, reduce errors caused by fading, and increase network capacity.

The polarization diversity technique involves using multiple antennas with different polarizations to receive the same signal simultaneously. Thanks to such an approach, a reduction in signal fading and an improvement in overall signal quality may be achieved. The polarization diversity approach is combined with spatial diversity to further improve signal quality and increase the system's capacity in noisy and fading wireless channels.

2. MIMO Antenna Performance Metrics

This section outlines S-parameters and most important diversity metrics for MIMO antennas with their typical acceptable values summarized in Tab. 1 [1], [2].

The envelope correlation coefficient (ECC) represents the relationship between the isolation factor and the number of antenna elements and in perfect circumstances is equal to zero. To determine the ECC between any number of elements, one can use the radiation field pattern and S-parameters as [3]:

$$\varrho_e = \frac{\left| \iint\limits_{4\pi} \left[\overrightarrow{F_1}(\theta,\phi) \times \overrightarrow{F_2^*}(\theta,\phi) \right] d\Omega \right|^2}{\iint\limits_{4\pi} \left| \overrightarrow{F_1}(\theta,\phi) \right|^2 d\Omega \left| \overrightarrow{F_2}(\theta,\phi) \right|^2 d\Omega}, \tag{1}$$

$$ECC_{(m,n)} = \frac{|S_{mm}^* S_{mn} + S_{nm}^* S_{nm}|^2}{\left(1 - |S_{mm}|^2 - |S_{mn}|^2\right)\left(1 - |S_{nm}|^2 - |S_{nn}|^2\right)} \tag{2}$$

Diversity gain (DG) is a parameter which defines the quality and reliability of a MIMO antenna in wireless systems. Hence,

Tab. 1. S parameter-dependent MIMO design metrics.

Parameter	Value
ECC	< 0.5
DG	<10 dB
TARC	< 0 dB
MEG	–3 dB < MEG < –12 dB
CCL	< 0.4 b/s/Hz

DG must be high (approx. 10 dB) within the acceptable frequency band. DG is also S parameter-dependent and can be calculated, for two ports, as:

$$DG = 10\sqrt{1 - ECC^2}\,. \tag{3}$$

The lower the value of ECC, the better the diversity gain. DG can be optimized by selecting an appropriate number of antennas, spacing, polarizations, and beamforming techniques, among other factors [3].

The total active reflection coefficient (TARC) measures reflected power about incident power and should ideally be zero for MIMO antennas. This metric for a 2-port antenna is calculated from S-parameters and expressed in decibels as [4]:

$$TARC = \frac{\sqrt{(S_{11} + S_{12})^2 + (S_{21} + S_{22})^2}}{\sqrt{2}}\,. \tag{4}$$

Channel capacity loss (CCL) is the highest amount of information that may be sent across a communication link with a low channel loss factor. For MIMO systems, a typical CCL value equals 0.4 bits/s/Hz. The formula for CCL using S-parameters for two ports [4] is:

$$C_{loss} = -\log_2(\sigma^D)\,, \tag{5}$$

where:

$$\sigma^D = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}\,, \tag{6}$$

$$\sigma_{11} = 1 - \left[|S_{11}|^2 + |S_{12}|^2\right], \tag{7}$$

$$\sigma_{22} = 1 - \left[|S_{22}|^2 + |S_{21}|^2\right], \tag{8}$$

$$\sigma_{12} = -\left[S_{11}^* S_{12} + S_{21}^* S_{12}\right], \tag{9}$$

$$\sigma_{21} = -\left[S_{22}^* S_{21} + S_{12}^* S_{21}\right]. \tag{10}$$

The mean effective gain (MEG) measure is defined as the average gain of a MIMO antenna array over all possible propagation channels, taking into account the statistical properties of the channel. Higher MEG values indicate better overall signal strength and improved system performance. MEG can be estimated based on the mutual effect between antenna power gain patterns and the statistical properties of incident radio waves through its vertical and horizontal components which depend on the orientation or polarization. MEG can be calculated as [5]:

$$G_e = \int_0^{2\pi}\int_0^{\pi}\left[\frac{XPR}{1+XPR}\,G_\theta(\theta,\varphi)\,P_\theta(\theta,\varphi) + \frac{1}{1+XPR}\,G_\varphi(\theta,\varphi)\,P_\varphi(\theta,\varphi)\right]\sin\theta\;d\theta\,d\varphi\,, \tag{11}$$

where XPR is the cross polarization, $G_\theta(\theta,\varphi), P_\theta(\theta,\varphi)$ are θ components of the antenna power gain pattern and θ components of the angular density functions of incoming plane waves, respectively. $G_\varphi(\theta,\varphi), P_\varphi(\theta,\varphi)$, are the φ components of the antenna power gain pattern and φ components of the angular density functions of incoming plane waves, respectively. Also, $G_e = G_\theta(\theta_s,\varphi_s)$, which shows that MEG corresponds to the directive gain of the antenna in the (θ_s,φ_s) direction, when incoming signals are focused in the (θ_s,φ_s) direction.

The dependence of MEG on S-parameters is as follows:

$$MEG_m = 1 - |S_{mm}|^2 - |S_{mn}|^2, \tag{12}$$

$$MEG_n = 1 - |S_{nn}|^2 - |S_{nm}|^2. \tag{13}$$

3. Review of MIMO Antenna Designs

The 5G frequency spectrum may be split into the lower spectrum, covering frequencies below 6 GHz (also known as sub-6 GHz) and the higher spectrum, covering frequencies over 6 GHz. Therefore, the review of the isolation enhancement techniques and the diversity parameters is based on the aforementioned taxonomy.

3.1. Review of Sub-6 GHz Antenna Design Features

For communicating in the sub-6 GHz band, the design presented in paper [6] provides maximum spatial diversity efficiency of more than 68% and a directive gain over 9.94 dB. The MIMO system with two and four elements showcased in paper [7] did not use any decoupling structures. Instead, the orthogonal orientation of elements is used, which provides high isolation for both systems at all operating frequencies with an efficiency of approximately 85% and a directive gain of 10 dB. In [8], the authors proposed a compact two-element antenna operating in the 5.5–9 GHz band for 5G applications, in which the parasitic element approach has been employed for enhancing isolation and the stepwise modification technique was used to improve bandwidth-related parameters.

In article [9], a dual-band antenna array is proposed with its structure consisting of an L-shaped feeding strip and a Z-shaped radiation strip at the initial stage. Such a design was then modified by adding a rectangular strip as a parasitic element to increase the bandwidth and with an L-shaped strip to make it resonate at 3.4 GHz. Further, a rectangular slit is cut on an additional L-shape strip creating a modified Z-shape that covers the 4.9 GHz band.

Tab. 2. Comparative analysis of decoupling techniques for sub-6 GHz MIMO antennas.

Ref.	Size [mm]	No. of elements	Decoupling technique	Isolation [dB]	Gain [dBi]	ECC	CCL
[6]	55×50×28	4	Spatial diversity, rotational diversity, orthogonal diversity	< –26 11 10	6	< 0.10 0.39 0.28	NC
[7]	23.5×26.5×1.6	4	Polarization diversity	< –30.5/18.5	NC	< 0.001	< 0.4
[8]	40×42×1.6	2	Parasitic element, orthogonal element approach	< –20	6.15	NC	NC
[9]	150×75×6.2	4	Parasitic element approach	< –16.5	4.7/5	< 0.01	NC
[10]	14.9×7×0.8	4	Open loop ring resonator	< –15	4.2	< 0.02	NC
[12]	50×50×1.6	4	Slots	< –11	4.1	0.01	NC

A dual-band antenna with an open loop ring resonator as a feeding element and T-shaped radiating elements is described in [10]. It operates in the 3.3–3.84 GHz and 4.61–5.91 GHz bands. The surface current for 3.5 GHz is directed from the right-hand side of the element towards the center of the substrate and the electric field is routed in the outward direction. The opposite phenomenon occurs for 5.5 GHz, making the proposed antenna of the dual-band type.

A multiband eight-port MIMO aerial is presented in paper [11]. It operates in n2, n3, n39, n65-66 and n77-79 bands. The elements achieve self-isolation through the spatial diversity technique. Another high efficiency 4×4 MIMO antenna for the 3.5 GHz band is presented in article [12]. A circular slot is etched on the antenna unit cell with a radius of 129.79° and with its center coinciding with the center of the ground plane. Each unit cell contributes to the formation of a circular disc in the ground plane. The structure not only provides good isolation, but also offers a peak gain of 4.0 dBi in the 3.4–3.8 GHz frequency range.

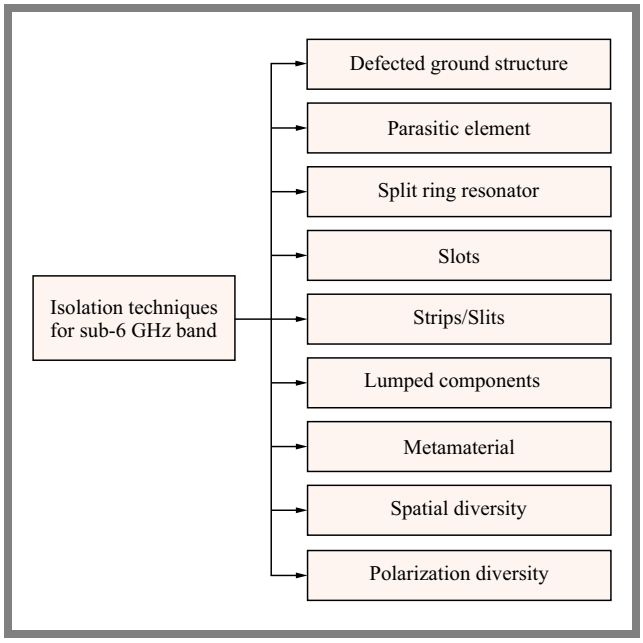

Fig. 1. Isolation techniques for sub-6 GHz band.

The isolation techniques which have been described in the literature are summarized in Fig. 1, while Tab. 2 illustrates the parameters of the described antennas using the individual decoupling techniques.

From Tab. 2, one may conclude that the hybrid of spatial and orthogonal polarization offers maximum isolation and gain for 4-port MIMO antennas while maintaining the expected values of all other parameters.

3.2. MIMO Antennas for 28.0 GHz

The 28 GHz MIMO configuration described in [13] consists of three circular rings surrounded by an infinity-shaped shell. The design offers a peak gain of 6.1 dBi and efficiency of 92%. It is also characterized by spatial and pattern-related diversity required for 5G-based communication. Next, the design was extended to a 1×4 array with a power divider providing a peak gain of 23.5 dBi and a wide bandwidth of 23.58–28.32 GHz.

A 4×4 microstrip patch antenna is proposed in [14], where one element comprises three stubs with a radial structure and stubs being tapered at the outer end. The radiating element is also extended to form an array to provide gain enhancement in the 28–38 GHz range. The array achieves 18.65 dBi of maximum gain with a return loss of –12 dB in the 28–38 GHz range.

A flower-shaped MIMO antenna is proposed in paper [15], where the radiating structure consists of five circular patches, placed 45° apart, and connected to a circular hub through rectangular branches. The ground plane is common and trimmed to enhance gain in the operating band. The antenna achieves –17 dB of isolation and 7.8 dBi of gain.

In [16], a MIMO antenna having T-shaped elements arranged in a row with CPW feed is presented. The design commenced with a single element and a ground plane etched with split

ring slots. Two iterations are then carried out for the etching of split ring slots. A slot in the center with two slots at both of its ends was etched in the partial ground at an optimized distance determined in the first iteration. The second iteration has two even more truncated slots etched in opposite directions at an optimized distance. These slots play a crucial role in the generation of multiple resonant frequencies and provide the bandwidth of 25.1–37.5 GHz, return loss of less than –10 dB, and 10.6 dBi of peak gain at 36 GHz.

In [17], a circular shaped patch has been modified by etching two rectangular slots on either side of the patch to operate in the 28 GHz band, achieving an impedance bandwidth of 26–29.7 GHz. Such a radiating element was next extended to form a four-port MIMO in order to enhance channel capacity. Additionally, DGS was implemented to reduce mutual coupling. A two-element MIMO aerial with isolation of more than 36 dB is presented in [18]. The rectangular-shaped radiators are etched with two slits each and are aligned at 45° in a clockwise direction to achieve high isolation at 29 GHz.

In [19], a dual function DRA-MIMO design is proposed featuring circular polarization in the 25.5–27.79 GHz band. Orthogonal placement of identical ports provides –30 dB of mutual coupling with a gain of 5 dBi and the diversity parameters are within the acceptable limits for 5G communication. An antenna array shown in article [20] provides a low-profile mmWave metamaterial-based MIMO array with two ports, each containing 3×3 unit cells. The antenna achieves high isolation of >24 dB and offers acceptable diversity parameter values over a wide impedance bandwidth with a measured peak gain of 12.4 dBi at 28 GHz.

Authors of work [21] have investigated an antenna with an operational range of 26–29.5 GHz in a 2×2 MIMO configuration, utilizing polarization diversity to obtain a high degree of isolation and a low envelope correlation coefficient.

The MIMO antenna from article [22] has an impedance bandwidth of 9.23 GHz in the 22.43 to 31.66 GHz range and is capable of achieving significant isolation of at least 25 dB between the neighboring MIMO elements, even without the use of decoupling networks. A MIMO antenna consisting of two elements arranged in an array connected through a shared feeding network is proposed in [23]. A bow-tie shaped slot is etched in the middle of the patch with slit on the edges. DGS is deployed by etching vertical and horizontal slots to optimize performance. A zig-zag decoupling structure, as well as spatial and polarization diversity techniques help achieve isolation greater than –40 dB. This antenna operates in the 27.6–28.6 GHz range with its peak gain equaling 12.02 dBi.

In Fig. 2, isolation techniques used in the literature for 28 GHz designs are summarized, while Tab. 3 presents a comparison of decoupling techniques with the main parameters of antennas described in recent articles.

From Tab. 3, one may notice that for 2-port MIMO antennas spatial diversity technique brings maximum isolation within the compact size of the antenna, whereas the maximum gain is achieved with designs based on metamaterial or metasurfaces. For 4-port MIMO antennas, the maximum isolation comes from slots, defected ground structures and decouplers, while

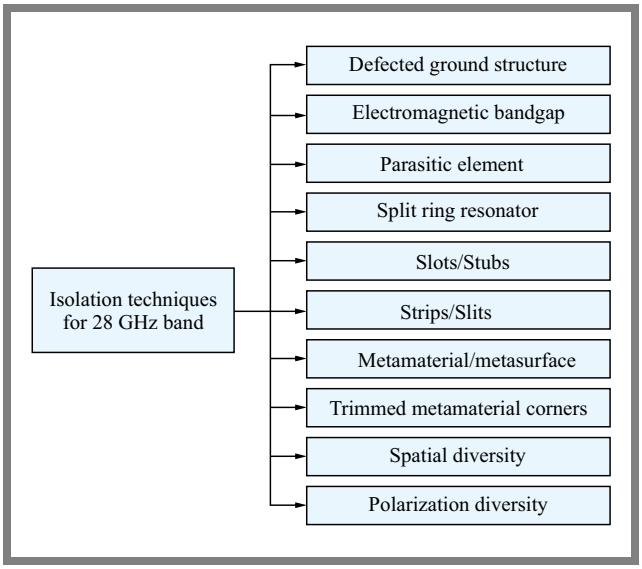

Fig. 2. Review of isolation techniques for MIMO antennas operating in 28.0 GHz band.

gain enhancement is achieved by electromagnetic bandgaps and spatial diversity as well.

3.3. Review of Dual Band Antennas for 28/38 GHz

28 GHz and 38 GHz frequencies are commonly used in 5G dual band antennas due to their low absorption rate.

In article [24], the authors investigated a two-port MIMO antenna with two arrays facing opposite directions. An EBG reflector was installed to minimize backward radiation while simultaneously enhancing the front-to-back ratio. It also helped in achieving a wideband impedance bandwidth with a realized gain of up to 11.5 and 10.9 dBi at 28 and 38 GHz frequencies, respectively.

Article [25] introduces a diamond-shaped pierced patch etched on a defected ground structure for dual-band operation. Orthogonal and spatial diversity techniques are used in four-port MIMO structures to achieve isolation of –50 dB with gain equaling 4.7/3.75 for 28/38 GHz, respectively.

In [26], a four-port MIMO with crescent-shaped radiators is proposed. Partial ground, spatial diversity, and polarization diversity techniques are used to enhance the emitter's isolation from the MIMO structure. The emitter which is 20×20 mm^2 in size exhibits more than –23 dB of isolation and 9.5/11.7 of gain at 28/38 GHz, respectively.

Paper [3] presents the parasitic element approach to two-port MIMO configurations. The isolation measured for the adjacent orientation of elements as well as for face-to-face placement of elements is less than –20 dB over the entire frequency range with a peak gain of 6.6/5.86 dBi at 28/38 GHz, respectively.

The four-port MIMO antenna demonstrated in paper [4] utilizes a parasitic element approach to enhance isolation which improved by 25 dB, with the maximum isolation value equaling –50 dB. The antenna offers a gain of 9.5/11.5 dB for 28/38 GHz bands, respectively. In [27], the authors reported a two port MIMO antenna having rectangular and triangular

Tab. 3. Summary of parameters achieved by decoupling techniques for 28 GHz MIMO antennas.

Ref.	Size [mm]	No. of elements	Decoupling technique	Isolation [dB]	Gain [dBi]	ECC	CCL
[13]	30×30×0.787	4	Pattern diversity	< –29	5.5	< 0.16	–
[14]	30×35×0.254	4	Decoupler and DGS	< –40	–	< 0.01	< 0.4
[15]	25×15×0.787	4	Pattern diversity	< –17	7.8	< 0.0001	–
[16]	12.7×50.8×0.800	4	Spatial diversity	< –22	5	< 0.01	–
[17]	30×30×1.575	4	Slot	< –30	–	< 0.005	< 0.15
[19]	15×30×0.254	2	Spatial diversity	< –35.8	–	< 0.005	< 0.1
[20]	54×23×0.790	2	Metamaterial	< –24	13.4	< 0.0013	< 0.42
[21]	20×20×7.608	4	Pattern diversity	< –25	14.1	< 0.008	–
[22]	24×24×0.254	4	Pattern diversity	< –25	5.6	< 0.0008	–
[23]	30 ×35×0.760	4	Slot	< –40	12	< 0.0003	< 0.4
[29]	110×75×0.760	4	Spatial diversity and DGS	< –22	9.3	< 0.05	< 0.1
[30]	42×85×0.508	2	Metasurface corrugations	< –37.1	18	–	–
[31]	15×25×0.203	2	Pattern diversity	< –30	5.9	< 0.005	< 0.12

stubs added to the radiator along with a partial ground to achieve resonance at 28 and 38 GHz bands. Isolation between MIMO elements is achieved by placing them orthogonally. The reported maximum isolation is less than –20 dB with a peak gain of 5.2/5.3 dBi and isolation of 30/22 for 28/38 GHz, respectively.

A slotted square radiator with a four-port MIMO configuration presented in [28] is characterized by left-hand circular polarization for 28 and 38 GHz bands. For such a design, isolation is below –36 dB and peak gain is 7.03/7.368 dB at 28/38 GHz, respectively.

A slotted rectangular two-port and four-port MIMO is described in article [32], where mutual coupling is reduced to –28.32 dB for the 28 GHz four-port MIMO and –26.27 dB for the 38 GHz four-port MIMO. Gain equals 7.95/8.27 dBi for 28/38 GHz, respectively.

The authors of [33] presented a design of a pentagon-shaped two-port MIMO antenna with a metamaterial array placed between the two symmetrically positioned radiators in order to isolate the electromagnetic fields. The achieved isolation is –39 dB at 28 GHz and –39 dB at 38 GHz with gain equaling 5.2 and 5.5 at 28 and 38 GHz, respectively.

Table 4 summarizes the aforementioned design features and provides the values of specific parameters, while Fig. 3 illustrates isolation techniques used in the literature for dual band antennas.

From Tab. 4, one may observe that metamaterials provide maximum isolation, while hybrid diversity (spatial and polarization) provides maximum gain for 2-port dual-band MIMO antennas, with hybrid diversity ensuring maximum isolation and high gain in 4-port dual-band MIMO antennas.

4. Conclusions

As the number of elements on the same patch increases, the problem of mutual coupling arises and it is a difficult challenge to establish optimum isolation to limit coupling. Numerous techniques have been analyzed which are deployed specifically to reduce mutual coupling or to improve diversity-related parameters. However, increasing gain while maintaining high diversity performance and decreasing the size of MIMO antennas are the few challenges that must be addressed.

This study proves that a substrate with low permittivity and a low loss tangent is necessary for enhancing gain and return loss. In addition, approaches relying on slots, dielectric lenses, and corrugation are required to eliminate the effect of cross-polarization. The use of dielectric lenses and para-

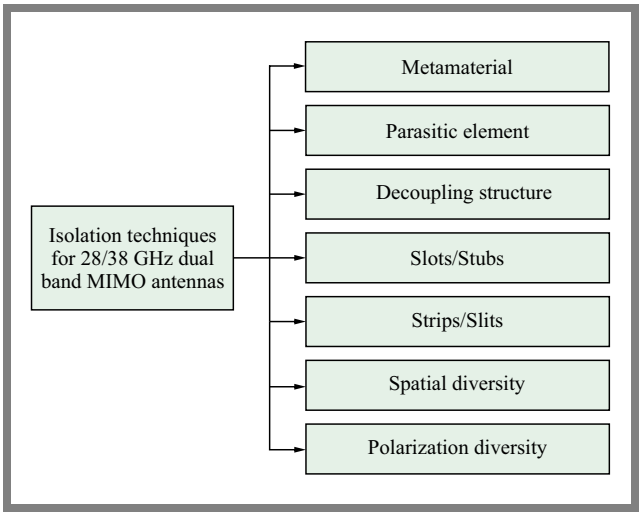

Fig. 3. Isolation techniques for 28/38 GHz band.

Tab. 4. Comparison of parameters of decoupling techniques for 28/38 GHz dual band MIMO antennas.

Ref.	Size [mm]	No. of elements	Decoupling technique	Isolation [dB]	Gain [dBi]	ECC	CCL
[4]	28×28×0.79	4	Parasitic element	–50	9.5/11.5	< 0.001	< 0.01
[24]	20×53×0.203	6	Electromagnetic band gap	< –20	15	< 0.12	–
[25]	150×75×3	4	Polarization and space diversity	–50	4.7/3.75	0	NC
[26]	20×20×0.25	4	Polarization and space diversity	–24/–25	9.5/11.7	< 0.0001	NC
[27]	27.65×12×0.273	2	Polarization and space diversity	–30/–22	5.2/5.3	< 0.0001/ < 0.0002	< 0.4
[28]	100×75×0.508	4	Polarization and space diversity	–36	7.6/8.12	< 0.005	< 0.4
[32]	55×110×0.508	4	Spatial diversity	–28/–26	8.27	< 0.005	NC
[33]	26×14.5×0.508	2	Metamaterial	–39/–38	5.2/5.5	< 0.0001	< 0.05
[34]	27.5×13.5×0.635	2	Spatial diversity	>20/20	NC	< 0.005/ < 0.0002	NC
[35]	30×15×0.203	2	Polarization and space diversity	–32.3/ –36.37	5.7/6.9	$< 10^{-4}$	< 0.3

sitic elements maximizes the radiation in the frontal direction for gain improvement. Metamaterials and DGS improve antenna gain, simultaneously reducing mutual coupling and antenna sizes, while techniques such as dielectric loading, interconnected ground strip, lumped components, orthogonal elements, meta-surfaces, FSS and polarization diversity contribute mostly to isolation improvement.

The major finding of the project is that spatial and polarization diversity isolation techniques (followed by metamaterials and electromagnetic band gap (EBG) techniques) offer higher isolation and improve gain in all the discussed frequency bands for two- and four-port MIMO antennas.

References

[1] K. Du, Y. Wang, and Y. Hu, "Design and Analysis on Decoupling Techniques for MIMO Wireless Systems in 5G Applications", *Applied Sciences*, vol. 12, no. 8, art. no. 3816, 2022 (https://doi.org/10.3390/app12083816).

[2] P. Sharma *et al.*, "MIMO Antennas: Design Approaches, Techniques and Applications", *Sensors*, vol. 22, no. 20, art. no. 7813, 2022 (https://doi.org/10.3390/s22207813).

[3] A.E. Farahat and K.F.A. Hussein, "Dual-band (28/38 GHz) Wideband MIMO Antenna for 5G Mobile Applications", *IEEE Access*, vol. 10, pp. 32213–32223, 2022 (https://doi.org/10.1109/ACCESS.2022.3160724).

[4] M. Hussain *et al.*, "Isolation Improvement of Parasitic Element-loaded Dual-band MIMO Antenna for mmWave Applications", *Micromachines*, vol. 13, no. 11, art. no. 1918, 2022 (https://doi.org/10.3390/mi13111918).

[5] T. Taga, "Analysis for Mean Effective Gain of Mobile Antennas in Land Mobile Radio Environments", *IEEE Transactions on Vehicular Technology*, vol. 39, no. 2, pp. 117–131, 1990 (https://doi.org/10.1109/25.54228).

[6] A. Yacoub, M. Khalifa, and D.N. Aloi, "Compact 2×2 Automotive MIMO Antenna Systems for Sub-6 GHz 5G and V2X Communications", *Progress In Electromagnetics Research B*, vol. 93, pp. 23–46, 2021 (https://doi.org/10.2528/PIERB21031606).

[7] D. Dileepan, S. Natarajan, and R. Rajkumar, "A High Isolation Multiband MIMO Antenna without Decoupling Structure for WLAN/WiMAX/5G Applications", *Progress in Electromagnetics Research C*, vol. 112, pp. 207–219, 2021 (https://doi.org/10.2528/PIERC21032605).

[8] T.M. Guan and S.K.A. Rahim, "Compact Monopole MIMO Antenna for 5G Application", *Microwave and Optical Technology Letters*, vol. 59, no. 5, pp. 1074–1077, 2017 (https://doi.org/10.1002/mop.30475).

[9] J. Huang *et al.*, "A Quad-port Dual-band MIMO Antenna Array for 5G Smartphone Applications", *Electronics*, vol. 10, no. 5, art. no. 542, 2021 (https://doi.org/10.3390/electronics10050542).

[10] J. Huang *et al.*, "Dual-band MIMO Antenna for 5G/WLAN Mobile Terminals", *Micromachines*, vol. 12, no. 5, art. no. 489, 2021 (https://doi.org/10.3390/mi12050489).

[11] Z.F. Al-Azzawi *et al.*, "Designing Eight-port Antenna Array for Multi-band MIMO Applications in 5G Smartphones", *Journal of Telecommunications and Information Technology*, no. 4, 2023 (https://doi.org/10.26636/jtit.2023.4.1297).

[12] S. Saxena *et al.*, "MIMO Antenna with Built-in Circular Shaped Isolator for sub-6 GHz 5G Applications", *Electronics Letters*, vol. 54, no. 8, pp. 478–480, 2018 (https://doi.org/10.1049/el.2017.4514).

[13] M.M. Kamal *et al.*, "Infinity Shell Shaped MIMO Antenna Array for mmWave 5G Applications", *Electronics*, vol. 10, no. 2, art. no. 165, 2021 (https://doi.org/10.3390/electronics10020165).

[14] K.R. Mahmoud and A.M. Montaser, "Optimized 4×4 Millimeter-wave Antenna Array with DGS Using Hybrid ECFO-NM Algorithm for 5G Mobile Networks", *IET Microwaves, Antennas & Propagation*, vol. 11, no. 11, pp. 1516–1523, 2017 (https://doi.org/10.1049/iet-map.2016.0959).

[15] S. Rahman *et al.*, "Nature Inspired MIMO Antenna System for Future mmWave Technologies", *Micromachines*, vol. 11, no. 12, art. no. 1083, 2020 (https://doi.org/10.3390/mi11121083).

[16] S.F. Jilani and A. Alomainy, "Millimeter-wave T-shaped MIMO Antenna with Defected Ground Structures for 5G Cellular Networks", *IET Microwaves, Antennas & Propagation*, vol. 12, no. 5, pp. 672–677, 2018 (https://doi.org/10.1049/iet-map.2017.0467).

[17] M. Hussain *et al.*, "Design and Characterization of Compact Broadband Antenna and its MIMO Configuration for 28 GHz 5G Applications", *Electronics*, vol. 11, no. 4, art. no. 523, 2022 (https://doi.org/10.3390/electronics11040523).

[18] A. Ahmad, D.Y. Choi, and S. Ullah, "A Compact Two Elements MIMO Antenna for 5G Communication", *Scientific Reports*, vol. 12, art. no. 3608, 2022 (https://doi.org/10.1038/s41598-022-07579-5).

[19] A. Kumar *et al.*, "Circularly Polarized Dielectric Resonator Based Two Port Filtenna for Millimeter-wave 5G Communication System", *IETE Technical Review*, vol. 39, no. 6, pp. 1501–1511, 2022 (https://doi.org/10.1080/02564602.2022.2028588).

[20] S.S. Al-Bawri *et al.*, "Hexagonal Shaped Near Zero Index (NZI) Metamaterial Based MIMO Antenna for Millimeter-wave Application", *IEEE Access*, vol. 8, pp. 181003–181013, 2020 (https://doi.org/10.1109/ACCESS.2020.3028377).

[21] N. Hussain, M. Jeong, J. Park, and N. Kim, "A Broadband Circularly Polarized Fabry-Perot Resonant Antenna Using a Single-layered PRS for 5G MIMO Applications", *IEEE Access*, vol. 7, pp. 42897–42907, 2019 (https://doi.org/10.1109/ACCESS.2019.2908441).

[22] M.I. Khan *et al.*, "A Compact mmWave MIMO Antenna for Future Wireless Networks", *Electronics*, vol. 11, no. 15, art. no. 2450, 2022 (https://doi.org/10.3390/electronics11152450).

[23] M. Bilal *et al.*, "High-isolation MIMO Antenna for 5G Millimeter-wave Communication Systems", *Electronics*, vol. 11, no. 6, article no. 962, 2022 (https://doi.org/10.3390/electronics11060962).

[24] A.A.R. Saad and H.A. Mohamed, "Printed Millimeter-wave MIMO-based Slot Antenna Arrays for 5G Networks", *AEU-International Journal of Electronics and Communications*, vol. 99, pp. 59–69, 2019 (https://doi.org/10.1016/j.aeue.2018.11.029).

[25] M.A. El-Hassan, K.F.A. Hussein, and A.E. Farahat, "Compact Dual-band (28/38 GHz) Patch for MIMO Antenna System of Polarization Diversity", *The Applied Computational Electromagnetics Society Journal*, vol. 37, no. 6, pp. 716–725, 2022 (https://doi.org/10.13052/2022.ACES.J.370607).

[26] N. Sghaier *et al.*, "Millimeter-wave Dual-band MIMO Antennas for 5G Wireless Applications", *Journal of Infrared, Millimeter, and Terahertz Waves*, vol. 44, pp. 297–312, 2023 (https://doi.org/10.1007/s10762-023-00914-5).

[27] A.R. Sabek, W.A.E. Ali, and A.A. Ibrahim, "Minimally Coupled Two-element MIMO Antenna with Dual Band (28/38 GHz) for 5G Wireless Communications", *Journal of Infrared, Millimeter, and Terahertz Waves*, vol. 43, pp. 335–348, 2022 (https://doi.org/10.1007/s10762-022-00857-3).

[28] F. Alnemr, M.F. Ahmed, and A.A. Shaalan, "A Compact 28/38 GHz MIMO Circularly Polarized Antenna for 5G Applications", *Journal of Infrared, Millimeter, and Terahertz Waves*, vol. 42, pp. 338–355, 2021 (https://doi.org/10.1007/s10762-021-00770-1).

[29] S.I. Naqvi *et al.*, "Integrated LTE and Millimeter-wave 5G MIMO Antenna System for 4G/5G Wireless Terminals", *Sensors*, vol. 20, no. 14, art. no. 3926, 2020 (https://doi.org/10.3390/s20143926).

[30] S. Gupta, Z. Briqech, A.R. Sebak, and T. Denidni, "Mutual-coupling Reduction Using Metasurface Corrugations for 28 GHz MIMO Applications", *IEEE Antennas and Wireless Propagation Letters*, vol. 16, pp. 2763–2766, 2017 (https://doi.org/10.1109/LAWP.2017.2745050).

[31] H. Zahra *et al.*, "A 28 GHz Broadband Helical Inspired End-fire Antenna and its MIMO Configuration for 5G Pattern Diversity Applications", *Electronics*, vol. 10, no. 4, art. no. 405, 2021 (https://doi.org/10.3390/electronics10040405).

[32] H.M. Marzouk, M.I. Ahmed, and A.A. Shaalan, "Novel Dual-band 28/38 GHz MIMO Antennas for 5G Mobile Applications", *Progress in Electromagnetics Research C*, vol. 93, pp. 103–117, 2019 (https://doi.org/10.2528/PIERC19032303).

[33] B.A. Esmail and S. Koziel, "High Isolation Metamaterial-based Dual-band MIMO Antenna for 5G Millimeter-wave Applications", *AEU-International Journal of Electronics and Communications*, vol. 158, art. no. 154470, 2023 (https://doi.org/10.1016/j.aeue.2022.154470).

[34] A. Omar, M. Hussein, I.J. Rajmohan, and K. Bathich, "Dual-band MIMO Coplanar Waveguide-fed-slot Antenna for 5G Communications", *Heliyon*, vol. 7, no. 4, art. no. 06779, 2021 (https://doi.org/10.1016/j.heliyon.2021.e06779).

[35] W.A.E. Ali, A.A. Ibrahim, and A.E. Ahmed, "Dual-band Millimeter Wave 2×2 MIMO Slot Antenna with Low Mutual Coupling for 5G Networks", *Wireless Personal Communications*, vol. 129, pp. 2959–2976, 2023 (https://doi.org/10.1007/s11277-023-10267-w).

Parminder Kaur, Ph.D.
Chitkara University Institute of Engineering and Technology
https://orcid.org/0000-0001-8161-1505
E-mail: parmindersandal@yahoo.com
Chitkara University, Punjab, India
https://www.chitkara.edu.in

Shivani Malhotra, Ph.D.
Chitkara University Institute of Engineering and Technology
https://orcid.org/0000-0001-8644-4689
E-mail: shivani.malhotra@chitkara.edu.in
Chitkara University, Punjab, India
https://www.chitkara.edu.in

Manish Sharma, Ph.D.
Chitkara University Institute of Engineering and Technology
https://orcid.org/0000-0002-1539-2938
E-mail: manishengineer1978@gmail.com
Chitkara University, Punjab, India
https://www.chitkara.edu.in

High-isolation Quad-port MIMO Antenna for 5G Applications

Noor Haider Saeed[1], Malik Jasim Farhan[1], and Qasim Hadi Kareem[2]

[1]*Mustansiriyah University, Baghdad, Iraq,*
[2]*Al-Iraqia University, Baghdad, Iraq*

https://doi.org/10.26636/jtit.2024.3.1582

Abstract — This paper presents a compact, high-isolation MIMO antenna with physical dimensions of 68 × 68 × 0.8 mm, designed for use in 5G applications. The antenna's bandwidth ranges from 3.25 GHz to 4.34 GHz and it offers a gain of 4.3 dBi, making it suitable for applications relying on 5G technology. Several improvements have been introduced to improve its overall efficiency, such as adjustments to the ground plane and integrating apertures in the radiating patch. The alterations referred to above were optimized using the sweep parameter method to ensure that their best configurations are achieved. Furthermore, much attention has been paid to enhance isolation by ensuring all terminals are positioned precisely at 90° angles. The CST Studio Suite was utilized to design and thoroughly simulate the proposed MIMO antenna.

Keywords — 5G, compact antenna, high-isolation, MIMO antenna, sub-6 GHz

1. Introduction

The inherent advantages of 5G wireless communication systems, including higher transmission rates, low latency, and improved spectral efficiency, have attracted increasing academia and industry interest [1], [2]. The frequency ranges employed in 5G communication can be categorized into two main groups: sub-6 GHz, which includes the spectrum below 6 GHz, and the 5G millimeter wave spectrum, above 24 GHz [3]. Although the 5G millimeter wave spectrum offers higher data transfer rates than sub-6 GHz, its widespread adoption could be improved due to its limited coverage range. The inclination towards sub-6 GHz frequencies has increased demand in many countries with well-developed wireless communication industries. Therefore, the frequency range of 3.4 to 3.8 GHz inside the sub-6 GHz spectrum has attracted significant interest recently [4].

To enhance channel capacity and spectral efficiency without using extra bandwidth or transmission power, 5G mobile devices use MIMO antennas [5] to achieve spatial diversity, i.e. transmit data across many antennas and facilitate propagation along spatial pathways that are not coupled. This approach enhances the dependability and robustness of systems in multipath fading scenarios.

The spatial multiplexing technique used in MIMO involves dividing the data stream into segments, each delivered along a distinct propagation path. This improves data transfer rates without requiring more bandwidth or transmission power [6]. The choice between these two methods depends on the prevailing wireless fading conditions. A MIMO system with spatial multiplexing is best suited for multipath wireless channels with a high signal-to-noise ratio (SNR), while spatial diversity is better for multipath wireless channels with a low SNR [7], [8].

However, increasing the number of antennas in MIMO might lead to problems with isolation. Consequently, this phenomenon can degrade antenna performance [8]. Hence, it is of crucial importance to attain superior performance of MIMO antenna systems, defined by the achievement of low envelope correlation coefficients (ECCs) and high isolation, while simultaneously maintaining compact antenna sizes [9]–[14].

Many research projects have been conducted to improve the isolation of MIMO antennas systems and boost their performance. A study described in [15] introduced a decoupling structure for MIMO antennas, aiming to enhance isolation by leveraging magnetic and electric coupling. This approach yielded isolation values below 20 dB in the 2.3–2.9 GHz frequency range. The authors of [22] present a dual-band, four-element MIMO system designed specifically for 5G mobile terminals. This system contains antenna pairs that are self-decoupled, resulting in enhanced isolation. The achieved isolation values surpass –17.5 dB and –20 dB in the 3.5 and 4.9 GHz frequency bands, respectively. In [17], a novel eight-port MIMO antenna array was developed specifically for 5G smartphones and a frequency of 2.6 GHz. This array has orthogonal polarization and strategically positioned elements, improving isolation between the antenna ports.

The study presented in [18] introduces an eight-element MIMO array developed specifically for 5G smartphones that operate inside the 3.45 GHz frequency band. This array utilizes U-shaped and L-shaped antenna arrays, relying on inverted-I ground slots and NL structure to achieve a 15 dB increased isolation. Moreover, in the study described in [19], an eight-port antenna array is proposed for 5G MIMO systems used in mobile devices. This antenna array operates in the 3.5 and 5 GHz frequency bands and includes a neutralized line to address the coupling issue.

The project described in [20] offers an additional contribution, as it discusses the integration of self-isolated antennas in a 5G MIMO system for mobile terminals. This integration achieves

This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) License.
For more information, see https://creativecommons.org/licenses/by/4.0/

isolation levels above 20 dB without additional decoupling devices.

In conclusion, the study described in [21] presents a compact MIMO antenna array design for 5G mobile devices. This design comprises eight elements and achieves a remarkable isolation performance of over 12.2 dB without additional elements or decoupling techniques. The MIMO antennas' mutual coupling is another important problem, since antenna components are located close to each other [22].

The main objective of this article is to propose and evaluate a compact MIMO antenna array designed specifically for 5G applications. The design has been optimized to ensure efficiency at a frequency of 3.6 GHz, employing a cost-effective FR-4 substrate material. The proposed antenna exhibits decent performance, confirmed by S-parameters, radiation patterns, gain characteristics, and envelope correlation coefficients. The tiny dimensions make it well-suited for easy integration into modern communication equipment.

The proposed design is unique due to the smaller dimensions of the antennas (measuring only 12 × 24 mm^2) compared to the solutions proposed in previous studies focusing on this particular area. Furthermore, it offers dual-band functionality, increased bandwidth, high isolation, and better gain.

The paper is organized as follows. Section 2 defines the geometry of the antenna system. Results of simulations concerning the proposed antenna (return loss, performance, and radiation pattern) are presented in Section 3. Section 4 compares the proposal with other research projects, with conclusions presented in Section 5.

2. Antenna Geometry

The process of designing an antenna may be divided into two distinct operational phases. First, an individual antenna element is developed. The second phase consists in integrating four individual aerials into a MIMO configuration. The geometry of the intended MIMO antennas was selected and CST software was used for simulation purposes. Through CST optimization, the values of specific parameters were fine-tuned to obtain the best possible results meeting the required specifications.

2.1. Single Antenna Structure

The primary element of the proposed antenna is the patch (radiating component, i.e. emitter). The remaining components include the ground plane, the dielectric substrate, and the input port. The patch has the form of a narrow strip of copper foil mounted on one side of the insulating substrate. The ground plane, made of the same metal as the patch, is placed on the opposite side of the substrate. In order to create some space between the patch and the ground plane, a dielectric substrate is used. Figure 1 shows the front and back of the proposed antenna, illustrating its components.

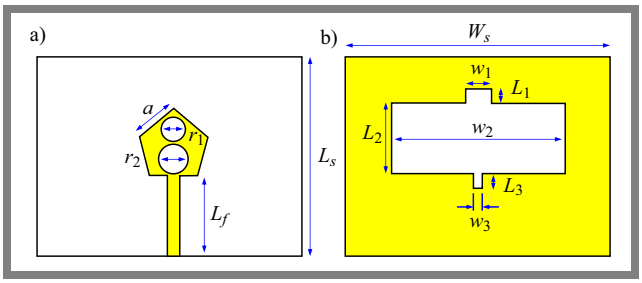

Fig. 1. Front a) and back b) view of the designed antenna.

2.2. MIMO Antenna Design

Figure 2 shows the proposed MIMO antenna configuration and a fabricated prototype comprising four ports. Each of these ports is positioned at a distance of less than one-third of the wavelength, on a 68 × 68 mm^2 FR-4 substrate with a dielectric constant of 4.3 and thickness of 0.8 mm. The detailed dimensions are provided in Tab. 1. The design incorporates symmetrical radiating elements, with optimized dimensions and suitable placement of the feed points ensuring that the most favorable impedance matching may be achieved.

Various techniques exist for feeding microstrip patch antennas, including microstrip aperture coupling, line feed, coaxial probe feed, electromagnetic coupling, and coplanar waveguide. In this case, a microstrip patch feed line with a characteristic impedance of 50 Ω is employed. The dimensions of the feed line, i.e. width W_f and length L_f, are fixed to ensure a 50 Ω impedance match. Figure 2a presents the configuration of such a feed line.

3. Results and Discussions

3.1. Single Antenna

The main parameter of the antenna is the maximum transfer of power, achieved by matching the feed line with input impedance. This relationship is typically represented by S_{11}

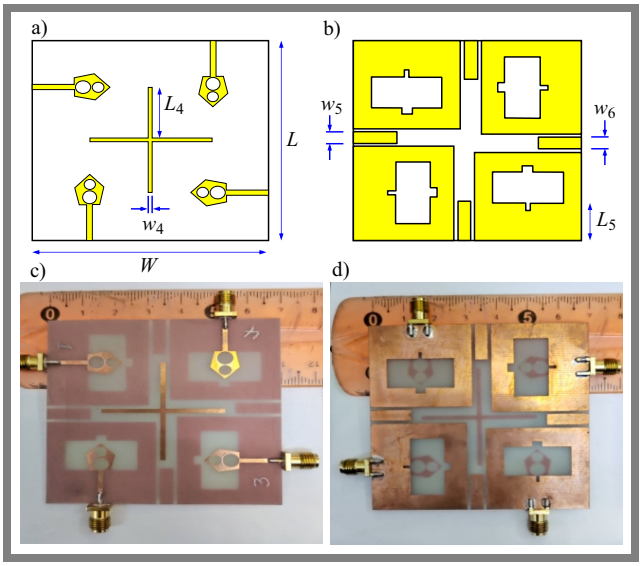

Fig. 2. Configuration of the proposed MIMO antenna: a) front view, b) back view, and c-d) both sides of the fabricated prototype.

Tab. 1. MIMO antenna dimensions.

Parameter	Value [mm]	Parameter	Value [mm]
W_s	32	w_1	3
L_s	30	L_1	2
W_f	1.5	w_2	20.8
L_f	12	L_2	10.65
a	6	w_3	1
L	12.3	L_3	2.35
r_1	2	L_4	17.25
r_2	2.4	w_4	1.5
h	0.8	w_5	6
L_5	13	w_6	4
L	68	W	68

quantifying the ratio between the amount of energy sent and reflected. Figure 3 shows S_{11} of an antenna as a function of frequency, known as its input reflection coefficient curve.

The antenna S-parameters are influenced by altering the radius of the circular slots on the patch. Various values specified by the sweep parameters are used. The return loss value must be minimized to less than –10 dB to ensure optimal antenna performance. By comparing different radius values, it was observed that with $r_1 = 2$ mm and $r_2 = 2.4$ mm, the antenna achieved a minimum return loss of –30 dB at a frequency of 3.6 GHz.

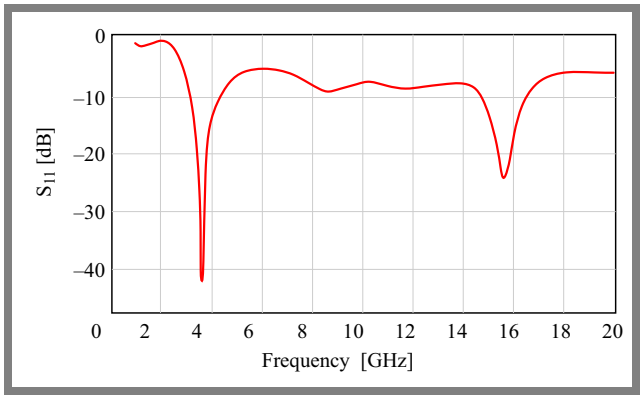

Fig. 3. Reflection coefficient of a single antenna.

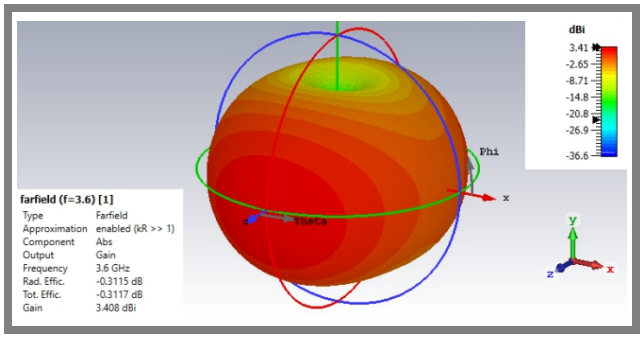

Fig. 4. Gain of single antenna.

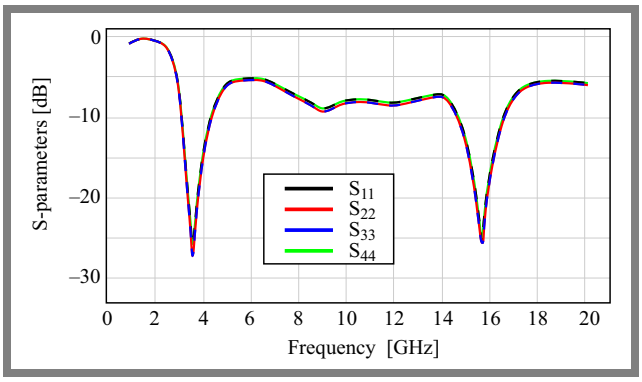

Fig. 5. Reflection coefficients of a MIMO antenna system.

The gain of an antenna refers to its capacity to focus radiated power in a specific direction, or to efficiently absorb incoming power from that direction. For a single patch antenna, this gain is approximately 3.4 dBi.

The simulated results demonstrate that the radiation patterns at far-field distance of the designed antenna are characterized by omnidirectional properties within the full bandwidth. Figure 4 illustrates gain characteristics of a single antenna.

3.2. MIMO Antenna

As the array technique enhances the behavior of a microstrip patch antenna, four array elements are used to show how the extra aerials improve the overall performance. The layout of the four components must be compact due to requirements prevailing in 5G mobile connectivity applications. To keep the proposed MIMO antenna compact and in to mitigate interference and fading caused by the larger number of antennas needed to achieve higher gains and wider bandwidths, an aerial spacing of 25 mm is recommended.

The reflection coefficient serves as a critical parameter for identifying the antenna's operating bands. Figure 5 shows the return loss simulation results (S_{11}–S_{44}). One may notice that the proposed MIMO antenna system achieves a bandwidth spanning from 3.25 to 4.34) GHz, as illustrated in the figure.

S_{12}, S_{13}, and S_{14} parameters play a significant role in characterizing mutual coupling within a MIMO system. As shown in Fig. 6, across the operational frequency spectrum, all these parameters remain below 25 dB. Due to the symmetrical structure of the MIMO design, S_{12} and S_{14} exhibit close similarity.

3.3. MIMO Antenna System Performance

The graphical representation of the simulated performance of the MIMO system is shown in Fig. 7.

One may notice that the gain and efficiency of the four antenna elements behave consistently throughout the operational frequency range. Specifically, gain remains constant at 4.3 dBi, while efficiency hovers at approximately 90%.

The two-dimensional electric field radiation patterns depicted in Fig. 8 provide a visual insight into the radiation characteristics of each antenna within the MIMO system. These patterns vividly illustrate the emission directions of the an-

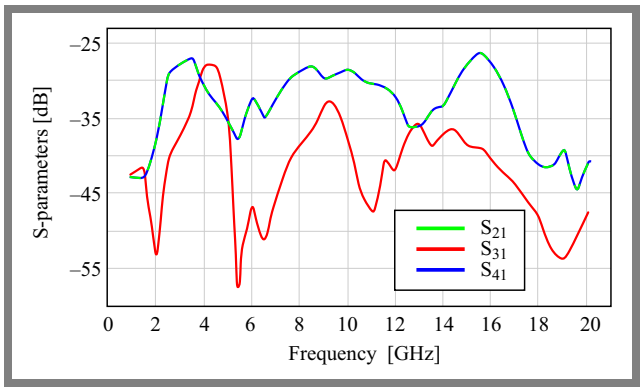

Fig. 6. Transmission coefficients of a MIMO system.

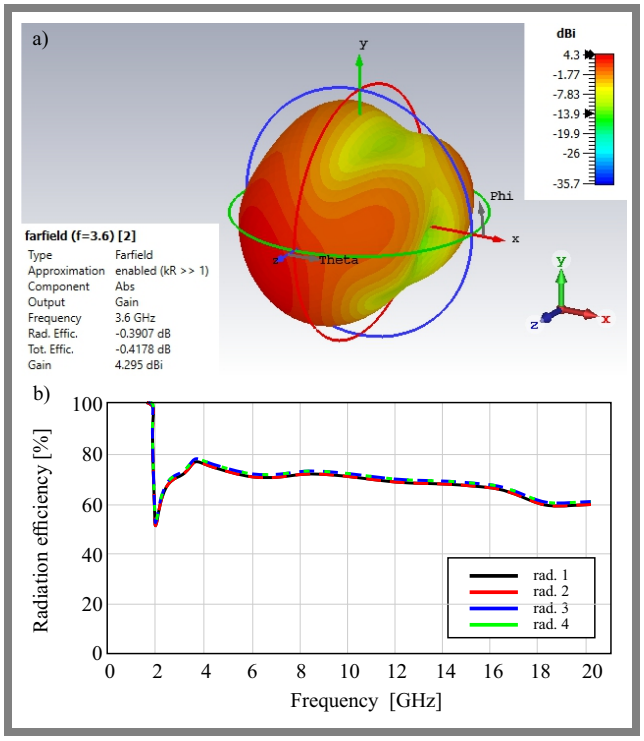

Fig. 7. Simulated MIMO antenna: a) gain and b) radiation efficiency.

tennas at the frequency of 3.6 GHz. These observations show that the proposed MIMO system effectively radiates in various directions, underscoring its capability to transmit signals across a broad angular spectrum.

The envelope correlation coefficient (ECC) is the most important metric used for assessing the correlation between different antenna elements within a MIMO system. ECC is utilized to evaluate the diversity performance of MIMO system configurations, where individual radiating elements can simultaneously and independently transmit data streams. As such, the radiators of antenna elements must exhibit low ECC values, ideally below 0.5, to ensure robust isolation between MIMO components. Correlation ϱ_e between two antenna elements i and j in an N port MIMO system can be calculated using the relationship provided in [23], [24].

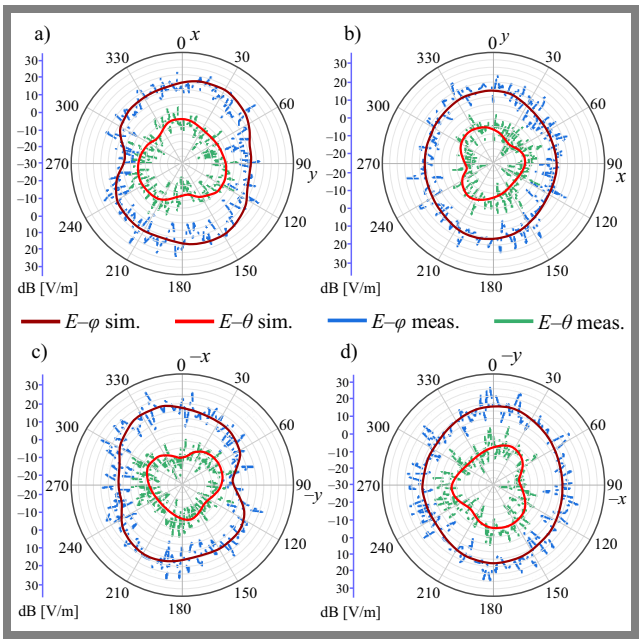

Fig. 8. Radiation pattern at 3.6 GHz for a) antenna 1, b) antenna 2, c) antenna 3, and d) antenna 4.

$$|\varrho_{(ij)}|^2 = \varrho_{(eij)} = \left| \frac{|S_{ii}^* S_{ij} + S_{ji}^* S_{jj}|}{\left|(1-|S_{ii}|^2-|S_{ji}|^2)(1-|S_{jj}|^2-|S_{ij}|^2)\eta_{rad\,i}\,\eta_{rad\,j}\right|^{\frac{1}{2}}} \right|^2 \quad (1)$$

Figure 9 illustrates the ECC for four elements, consistently remaining below 0.05 across the entire frequency band. This signifies strong isolation between the emitters, indicating that the designed MIMO system is characterized by good diversity performance.

The diversity gain (DG) of a MIMO antenna system is an important parameter in the context of various strategies of enhancing the signal-to-interference ratio [25]. It can be expressed as:

$$DG = 10\sqrt{1-|\varrho_{(eij)}|}\,. \quad (2)$$

Correlation coefficient ϱ_e predominantly influences diversity gain. The highest attainable diversity gain value of 10 is achieved when ϱ_e equals zero, signifying an optimal diversity scenario. Higher diversity gain values are achieved as ECC values decrease.

As shown in Fig. 10, DG approaches nearly 10 dB for the antenna's operating band. This underscores the considerable enhancement in diversity gain achieved within these frequency ranges.

4. Comparison with Previous Works

The quad-port MIMO antenna system introduced in this research is compared with other recent studies. As illustrated in Tab. 2, the proposed MIMO system is characterized a relatively small physical dimensions, despite consisting of four radiating elements. In contrast to the cited literature, the ar-

Tab. 2. Comparison with recent works dealing with 4-element MIMO antennas.

Reference	PCB size [mm^2]	Bandwidth [GHz]	Isolation [dB]	Efficiency [%]	ECC
[9]	130 × 74	3.3–3.6, 4.8–5 (–6 dB)	>12	>50	<0.1
[10]	120 × 65	3.3–4.2, 4.4–5	>18.8	—	<0.018
[11]	150 × 75	3.4–3.6	>16.3	>64	
[12]	115 × 65	3.5–5.9	>20	—	<0.005
[13]	150 × 75	3.21–3.81	>10	>70	<0.02
[14]	150 × 82	3.3–6.6	>23	—	<0.005
Proposed	68 × 68	3.25–4.32	>25	>88	≈0

chitecture described in this study features a wide operational bandwidth.

Furthermore, the proposed compact design demonstrates decent isolation and high gain due to the orthogonal alignment of the radiating antennas. Such an alignment facilitates the introduction of diverse patterns. It achieves isolation levels of more than 23 dB between radiators, without the need to use complex decoupling components.

5. Conclusion

This paper presents an efficient and compact quad-element MIMO antenna system designed for 5G applications, specifically focusing on the 3.25 to 4.32 GHz frequency range. The design incorporates four patch antennas carefully placed in the corners of FR-4 substrate, with the overall dimensions of the substrate equaling 68 × 68 × 0.8 mm^3. The shape of each antenna includes a polygonal patch element connected to a microstrip feed line. This specific layout improves isolation by guaranteeing that all antenna ports are oriented orthogonally at an angle of 90°. With isolation levels above 25 dB, radiation efficiency above 85%, and an ECC approximation close to zero, the proposed design is well-suited for use in 5G mobile terminals.

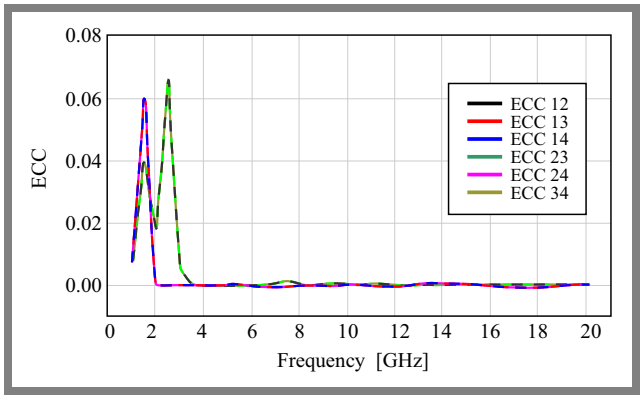

Fig. 9. ECC of the proposed MIMO system.

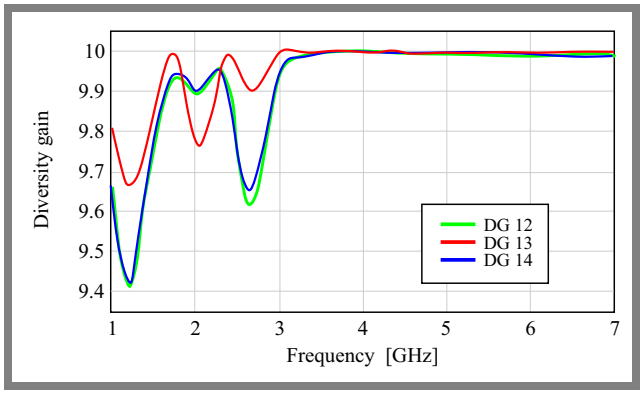

Fig. 10. Diversity gain of the proposed MIMO antenna system.

Acknowledgments

The authors wish to thank Mustansiriyah University in Baghdad, Iraq for assistance and support provided in this research.

References

[1] C.Z. Han, L. Xiao, Z. Chen, and T. Yuan, "Co-located Self-neutralized Handset Antenna Pairs with Complementary Radiation Patterns for 5G MIMO Applications", *IEEE Access*, vol. 8, pp. 73151–73163, 2020 (https://doi.org/10.1109/ACCESS.2020.2988072).

[2] Q.H.K. Al-Gertany, M.J. Farhan, and A.K. Jasim, "Design and Analysis a Frequency Reconfigurable Octagonal Ring-shaped Quad-port Dual-band Antenna Based on a Varactor Diode", *Progress in Electromagnetics Research C*, vol. 116, pp. 235–248, 2021 (https://doi.org/10.2528/PIERC21091705).

[3] A. Ghaffar *et al.*, "Design and Realization of a Frequency Reconfigurable Multimode Antenna for ISM, 5G-sub-6-GHz, and S-band Applications", *Applied Sciences*, vol. 11, no. 4, art. no. 1635, 2021 (https://doi.org/10.3390/app11041635).

[4] Q.H. Kareem and M.J. Farhan, "Compact Dual-polarized Eight-element Antenna with High Isolation for 5G Mobile Terminal Applications", *International Journal of Intelligent Engineering & Systems*, vol. 14, no. 6, pp. 187–197, 2021 (https://doi.org/10.22266/ijies2021.1231.18).

[5] A. Zhao and Z. Ren, "Size Reduction of Self-isolated MIMO Antenna System for 5G Mobile Phone Applications", *IEEE Antennas and Wireless Propagation Letters*, vol. 18, no. 1, pp. 152–156, 2018 (https://doi.org/10.1109/LAWP.2018.2883428).

[6] J. Sui and K.-L. Wu, "Self-curing Decoupling Technique for Two Inverted-F Antennas with Capacitive Loads", *IEEE Transactions on Antennas and Propagation*, vol. 66, no. 3, pp. 1093–1101, 2018 (https://doi.org/10.1109/TAP.2018.2790041).

[7] Y. Wu *et al.*, "A Survey on MIMO Transmission with Finite Input Signals: Technical Challenges, Advances, and Future Trends", *Proceedings of the IEEE*, vol. 106, no. 10, pp. 1779–1833, 2018 (https://doi.org/10.1109/JPROC.2018.2848363).

[8] Y. Li, Y. Luo, C.Y.D. Sim, and G. Yang, "High-isolation 3.5 GHz Eight-antenna MIMO Array Using Balanced Open-slot Antenna

Element for 5G Smartphones", *IEEE Transactions on Antennas and Propagation*, vol. 67, no. 6, pp. 3820–3830, 2019 (https://doi.org/10.1109/TAP.2019.2902751).

[9] M.K. Sharma, M. Kumar, J.P. Saini, and S.P. Singh, "Computationally Optimized MIMO Antenna with Improved Isolation and Extended Bandwidth for UWB Applications", *Arabian Journal for Science and Engineering*, vol. 45, pp. 1333–1343, 2020 (https://doi.org/10.1007/s13369-019-03888-6).

[10] W. Zhang, Z. Weng, and L. Wang, "Design of a Dual-band MIMO Antenna for 5G Smartphone Application", *2018 International Workshop on Antenna Technology (iWAT)*, Nanjing, China, 2018 (https://doi.org/10.1109/IWAT.2018.8379211).

[11] A. Biswas and V.R. Gupta, "Design and Development of Low Profile MIMO Antenna for 5G New Radio Smartphone Applications", *Wireless Personal Communications*, vol. 111, pp. 1695–1706, 2020 (https://doi.org/10.1007/s11277-019-06949-z).

[12] M.Y. Muhsin, J.K. Ali, and A.J. Salim, "A Compact High Isolation Four Elements MIMO Antenna System for 5G Mobile Devices", *Engineering and Technology Journal*, vol. 40, no. 8, pp. 1055–1061, 2022 (https://doi.org/10.30684/etj.2021.131103.1004).

[13] A.S. Abdullah, S.A. Hashem, W.S. Al-Dayyeni, and M.F. Hassoun, "Four-port Wideband Circular Polarized MIMO Antenna for Sub-6 GHz Band", *Proceedings of International Conference on Emerging Technologies and Intelligent Systems*, vol. 2, pp. 909–920, 2022 (https://doi.org/10.1007/978-3-030-85990-9_72).

[14] U. Rafique, S. Khan, S.M. Abbas, and P. Dalal, "Uni-Planar MIMO Antenna for Sub-6 GHz 5G Mobile Phone Applications", *2022 IEEE Wireless Antenna and Microwave Symposium (WAMS)*, Rourkela, India, 2022 (https://doi.org/10.1109/WAMS54719.2022.9848366).

[15] H.S. Najim, M.F. Mosleh, and R.A. Abd-Alhameed, "Design a MIMO Printed Dipole Antenna for 5G Sub-band Applications", *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 27, no. 3, pp. 1649–1660, 2022 (https://doi.org/10.11591/ijeecs.v27.i3.pp1649-1660).

[16] C.-D. Xue *et al.*, "MIMO Antenna Using Hybrid Electric and Magnetic Coupling for Isolation Enhancement", *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 10, pp. 5162–5170, 2017 (https://doi.org/10.1109/TAP.2017.2738033).

[17] S. Saxena *et al.*, "MIMO Antenna with Built-in Circular Shaped Isolator for Sub-6 GHz 5G Applications", *Electronics Letters*, vol. 54, no. 8, pp. 478–480, 2018 (https://doi.org/10.1049/el.2017.4514).

[18] M.-Y. Li *et al.*, "Eight-port Orthogonally Dual-polarized MIMO Antennas Using Loop Structures for 5G Smartphone", *2016 IEEE 5th Asia-Pacific Conference on Antennas and Propagation (APCAP)*, Kaohsiung, Taiwan, 2016 (https://doi.org/10.1109/APCAP.2016.7843165).

[19] W. Jiang, B. Liu, Y. Cui, and W. Hu, "High-isolation Eight-element MIMO Array for 5G Smartphone Applications", *IEEE Access*, vol. 7, pp. 34104–34112, 2019 (https://doi.org/10.1109/ACCESS.2019.2904647).

[20] J. Guo, L. Cui, C. Li, and B. Sun, "Side-edge Frame Printed Eight-port Dual-band Antenna Array for 5G Smartphone Applications", *IEEE Transactions on Antennas and Propagation*, vol. 66, no. 12, pp. 7412–7417, 2018 (https://doi.org/10.1109/TAP.2018.2872130).

[21] A. Zhao and Z. Ren, "Multiple-input and Multiple-output Antenna System with Self-isolated Antenna Element for Fifth-generation Mobile Terminals", *Microwave and Optical Technology Letters*, vol. 61, no. 1, pp. 20–27, 2019 (https://doi.org/10.1002/mop.31515).

[22] Q.H.K. Al-Gertany, R.A. Shihab, and H.H. Kareem, "Compact Dual-polarized Reconfigurable MIMO Antenna Based on a Varactor Diode for 5G Mobile Terminal Applications", *Progress in Electromagnetics Research C*, vol. 137, pp. 185–198, 2023 (https://doi.org/10.2528/PIERC23072204).

[23] M.Y. Muhsin, A.J. Salim, and J.K. Ali, "An Eight-element MIMO Antenna System for 5G Mobile Handsets", *2021 International Symposium on Networks, Computers and Communications (ISNCC)*, Dubai, United Arab Emirates, 2021 (https://doi.org/10.1109/ISNCC52172.2021.9615663).

[24] M.H. Reddy and D. Sheela, "MIMO Antenna Design and Optimization with Enhanced Bandwidth for Wireless Applications", *Journal of Telecommunications and Information Technology*, vol. 4, pp. 22–26, 2020 (https://doi.org/10.26636/jtit.2020.145520).

[25] S. Blanch, J. Romeu, and I. Corbella, "Exact Presentation of Antenna System Diversity Performance from Input Parameter Description", *Electronics Letters*, vol. 39, no. 9, pp. 705–707, 2003 (https://doi.org/10.1049/el:20030495).

Noor Haider Saeed, M.Sc.
Electrical Engineering Department
https://orcid.org/0009-0006-1954-1546
E-mail: Noorhaider98@uomustansiriyah.edu.iq
Mustansiriyah University, Baghdad, Iraq
https://www.uomustansiriyah.edu.iq

Malik Jasim Farhan, Ph.D.
Electrical Engineering Department
https://orcid.org/0000-0002-8952-6134
E-mail: malik.jasim@uomustansiriyah.edu.iq
Mustansiriyah University, Baghdad, Iraq
https://www.uomustansiriyah.edu.iq

Qasim Hadi Kareem, Ph.D.
Electrical Engineering Department
https://orcid.org/0000-0002-0586-4436
E-mail: qasim.hadi2017@gmail.com
Al-Iraqia University, Baghdad, Iraq
https://en.aliraqia.edu.iq

A Deep Learning-based Approach for Channel Estimation in Multi-access Multi-antenna Systems

Mahmoud M. Qasaymeh[1], Ali Alqatawneh[1], Mahmoud A. Khodeir[2], and Ahmad F. Aljaafreh[1,3]

[1]*Tafila Technical University, Tafila, Jordan,*
[2]*Jordan University of Science and Technology, Irbid, Jordan,*
[3]*University of Detroit Mercy, Detroit, USA*

https://doi.org/10.26636/jtit.2024.3.1701

Abstract — This paper studies estimating the channel state information at the end of receiver (CSIR) for multiple transmitters communicating with only one receiver so that the latter can decode the incoming signal more efficiently. The transmitters and the receiver are all equipped with multi-antennas and using orthogonal space-time block codes (OSTBC). An algorithm is developed based on deep learning for estimating multi-user multiple-input multiple-output (MU-MIMO) channels. The algorithm could estimate the CSIR using a single pilot block. The proposed convolutional neural network (CNN) architecture designed for this task begins with an input layer that accepts grayscale images, followed by six convolutional blocks for feature extraction and processing. The network concludes with a fully connected layer to output the estimated channel information. It is trained using a regression loss function to map input images to accurate channel information accurately. The performance of the proposed method is compared with classical methods like least square and subspace-based methods, including Capon and rank revealing QR (RRQR) methods. CNN achieved better performance in comparison with the reference. Computer simulations are included to validate the proposed method.

Keywords — channel estimation, CNN, CSI, CSIR, least square, MU-MIMO, OSTBC

1. Introduction

Multi-user multiple-input multiple-output (MU-MIMO) and orthogonal space-time block code (OSTBC) are advanced technologies that enhance data rate and reliability in wireless communication systems, while combined, they provide dependable performance for networks with many users and antennas. By employing MU-MIMO, the spatial dimension can be achieved to simultaneously serve several users via the same time-frequency resources using space division multiple access (SDMA), beamforming, and interference management techniques like zero-forcing (ZF) and minimum mean square error (MMSE).

The OSTBC codes are employed to encode the data across various antennas and time slots, and this enhances spatial diversity and ensures reliable transmission despite the channel's variability. Initially developed by Alamouti for systems with two transmit antennas, the OSTBC codes have been integrated into 5G systems to exploit their diversity and their robustness in high mobility scenarios [1], [2].

Consider a mobile base station equipped with several antennas to serve multiple users, each with several aerials. Using MU-MIMO techniques, the base station shares its spatial resources among individuals, allowing for concurrent communication. Those resources allocated to each user make it possible for them to freely communicate even though they may experience changes in their transmission paths [3]–[11].

The STBC codes can be classified based on the availability of a channel state information (CSI), distinguishing between codes that require no CSI, CSI at the transmit side (CSIT), or CSI at the receiving side (CSIR). Papers [12] and [13] were expanded to encompass a multi-user multi-antenna (MUMA) at the transmitter communicating with a single receiver under the perfect CSIR condition. Understanding this information can help the recipient better interpret the received signal and identify any potential corruption by the channel, thereby significantly improving the reliability and quality of the received signal [14], [15]. However, it is essential to note that ideal channel state information (ICSI) is unavailable in real-world environments. Therefore, wireless communication systems should utilize channel estimation schemes. These schemes can be blind, semi-blind channels or leverage pilot symbols already known to the receiver.

Minimizing the number of transmitted pilot symbols is desirable to enhance spectral efficiency in bandwidth-constrained environments, making blind or semi-blind channel estimation techniques preferable [16]–[21]. Well-known least squares and subspace estimation algorithms, such as rank revealing QR factorization (RRQR), propagator method (PM) and Capon method, were employed to determine the null space or signal subspace for estimating the CSI [22], [23].

The telecommunications industry is shifting towards the sixth generation (6G) to accommodate the increasing volume of information exchange and the need for rapid system responses. This transition is supported by advanced multi-input multi-output (MIMO) configurations and machine learning (ML)-

This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) License.
For more information, see https://creativecommons.org/licenses/by/4.0/

based channel estimation techniques in conjunction with traditional methods like OSTBC [24], [25]. Over the past few years, artificial intelligence has emerged as a powerful tool for enhancing the performance of wireless communications. In wireless communication systems, ML can be used either to enhance the performance of a specific functional part of the communication system, such as signal detection, noise cancelation, channel estimation, or as an end-to-end, where the transmitter and the receiver can be replaced by a neural network (NN) [26]–[28].

Classical signal processing has demonstrated effectiveness in channel estimation for wireless communication systems. However, these methods exhibit limitations, particularly in sophisticated wireless environments like those encountered in 5G and beyond. These limitations include the computational complexity mainly for massive MIMO systems, the necessity for prior channel characteristics knowledge in signal processing blind-based channel estimation and the significant overhead introduced by pilot-aided methods. Fortunately, ML can efficiently handle these challenges [29], [30].

Chun *et al.* in [31] proposed a deep learning-based approach for channel estimation in massive MIMO systems. The pilot and the channel coefficients estimator were designed using a two-layer neural network (TNN) and a deep neural network (DNN). Based on datasets collected from a practical wireless environment, a DL-based channel estimation scheme was proposed in [32] for mmWave massive MIMO systems. The generated channel matrix was utilized as the input of a DL neural network with REsUNet to enhance channel feature extraction.

Meenalakshmi *et al.* [33] applied a DL method to 5G MIMO-OFDM systems, highlighting the method's potential in high-dimensional signal processing. In a MIMO-OFDM system, the overhead on pilot symbols used for channel estimation increases with the number of antennas at both ends. In order to increase performance, authors in [34] integrated DL with bidirectional long short-term memory (bi-LSTM) networks for joint channel estimation and symbol detection in MIMO-OFDM systems.

Motivated by the opportunities and trends for incorporating artificial intelligence into wireless communications, as well as the potential advantages of machine learning over traditional signal processing methods in addressing problems in wireless communication systems, we proposed a DL-based approach for estimating channel parameters in a MUMA system using a convolutional neural network (CNN). The CNN architecture is tailored to extract and learn features from input signal data images, which enables precise channel estimation. It comprises multiple convolutional, batch normalization, rectified linear unit (ReLU), pooling layers, a fully connected layer, and a regression layer, all optimized using stochastic gradient descent with momentum (SGDM).

The structure of this paper is as follows. Section 2 presents the MU-MIMO environment and gives the mathematical model for the problem formulation. Section 3 develops the proposed method which includes the CNN architecture and the training process. Section 4 illustrates and compares the performance of the proposed CNN across a wide range of signal-to-noise ratios (SNR) with previous work such as: least squares (LS), Capon and rank revealing QR (RRQR) methods. Finally, Section 5 presents some concluding remarks, summarizing the key findings.

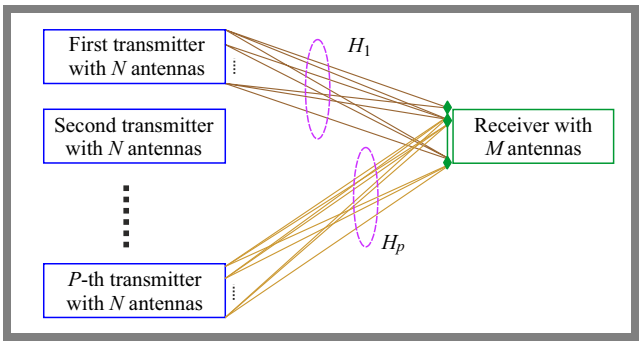

Fig. 1. MU-MIMO environment of P transmitters and a single receiver.

2. Problem Formulation

Let us consider an uplink reception scenario, where P users transmit their signals encoded with OSTBC in a MU-MIMO environment as shown in Fig. 1.

The base station applies multiuser detection algorithms to separate the users' signals and decode the OSTBC-encoded data. Assuming that all transmitters are equipped with N antennas and the receiver is equipped with m antennas. To simplify further, we assume that all transmitters employ the same OSTBC code of block length T. Moreover, we consider a flat block fading channel [22]. The p-th user CSI is $\mathbf{H}_p$ of size $N \times M$, where its elements are assumed to be independent complex Gaussian random variables. Thus, the received signal matrix of size $T \times M$ is given by:

$$\mathbf{Y} = \sum_{p=1}^{P} \mathbf{X}(\boldsymbol{s}_p)\, \mathbf{H}_p + \mathbf{Z}\,, \tag{1}$$

where $\mathbf{X}(\boldsymbol{s}_p)$ is the transmitted signal matrix of size $T \times N$, $\boldsymbol{s}_p = [s_1, s_2, \ldots, s_K]$ is the transmitted frame vector of size $K \times 1$, which consists of the quadrature phase shift keying (QPSK) modulated data symbols and $\mathbf{Z}$ is the corresponding noise matrix. The elements of OSTBC matrix $\mathbf{X}(\boldsymbol{s}_p)$ are function of $\boldsymbol{s}_p$ and its complex conjugates [22].

The matrix $\mathbf{X}(\boldsymbol{s}_p)$ can be rewritten by separating real and imaginary parts as:

$$\mathbf{X}(\boldsymbol{s}_p) = \sum_{k=1}^{K} \left(\mathbf{C}_k\, \mathrm{Re}\{s_k\} + \mathbf{D}_k\, \mathrm{Im}\{s_k\}\right). \tag{2}$$

Here, matrices $\mathbf{C}_k := \mathbf{X}(\boldsymbol{e}_k)$ and $\mathbf{D}_k := \mathbf{X}(i\boldsymbol{e}_k)$. Also, $\boldsymbol{e}_k$ is the $K \times 1$ identity matrix vector corresponding to k-th column and i is an imaginary unit $\sqrt{-1}$. Using Eq. (2), we can express the reshaped real received vector of size $2MT \times 1$ for single transmitted frame per user as [23]:

$$\boldsymbol{y} = \overline{\boldsymbol{Y}} = \sum_{p=1}^{P} \mathbf{A}(\boldsymbol{h}_p)\, \overline{\boldsymbol{s_p}} + \overline{\boldsymbol{Z}}. \tag{3}$$

Define $\overline{\boldsymbol{B}}$ as a real vector after reshaping the complex matrix $\mathbf{B}$ using the vectorization operator $\mathrm{vec}\{.\}$ by stacking all columns of a matrix on top of each other as:

$$\overline{\boldsymbol{B}} := \begin{bmatrix} \mathrm{vec}\{\mathrm{Re}(\mathbf{B})\} \\ \mathrm{vec}\{\mathrm{Im}(\mathbf{B})\} \end{bmatrix}. \tag{4}$$

Define also, $\boldsymbol{h}_p := \overline{\boldsymbol{H}}_p$ of size $2MN \times 1$. An extended vector of all channels $\boldsymbol{g}$ of size $2MNP$ is formed as:

$$\boldsymbol{g} = \begin{bmatrix} \boldsymbol{h_1} \\ \boldsymbol{h_2} \\ \vdots \\ \boldsymbol{h_p} \end{bmatrix}. \tag{5}$$

The vectorized form of the matrix $\mathbf{A}(\boldsymbol{h}_p)$ is given by:

$$\begin{aligned} \mathbf{A}(\boldsymbol{h}_p) &:= \left[\overline{\boldsymbol{C}_1\boldsymbol{H}_p}, \ldots, \overline{\boldsymbol{C}_K\boldsymbol{H}_p}, \overline{\boldsymbol{D}_1\boldsymbol{H}_p}, \ldots, \overline{\boldsymbol{D}_K\boldsymbol{H}_p}\right] \\ &= \left[\boldsymbol{a}_1(\boldsymbol{h}_p)\boldsymbol{a}_2(\boldsymbol{h}_p) \ldots\ldots \boldsymbol{a}_{2K}(\boldsymbol{h}_p)\right]. \end{aligned} \tag{6}$$

The matrix in Eq. (6) is an orthogonal matrix with column norm of $\|h_p\|^2$. As $\mathbf{A}(\boldsymbol{h}_p)$ is linear in $\boldsymbol{h}_p$ there exists $2K$ real matrices $\boldsymbol{\Phi}_k, k = 1, 2, \ldots, 2K$ with dimensions $2MT \times 2MN$ such that:

$$\boldsymbol{a}_k(\boldsymbol{h}_p) := \boldsymbol{\Phi}_k \boldsymbol{h}_p \quad \text{for} \quad k = 1, 2, \ldots, 2K\,. \tag{7}$$

Here, $\boldsymbol{\Phi}_k$ for $k = 1, 2, \ldots, 2K$ is OSTBC specific and known. Based on Eq. (7) we can rewrite Eq. (6) as:

$$\mathbf{A}(\boldsymbol{h}_p) := \left[\boldsymbol{\Phi}_1\boldsymbol{h}_p\ \boldsymbol{\Phi}_1\boldsymbol{h}_p \ldots\ldots \boldsymbol{\Phi}_{2K}\boldsymbol{h}_p\right], \tag{8}$$

and

$$\mathrm{vec}\left\{\mathbf{A}(\boldsymbol{h}_p)\right\} = \boldsymbol{\Phi}\boldsymbol{h}_p\,, \tag{9}$$

where:

$$\boldsymbol{\Phi} := \left[\boldsymbol{\Phi}_1^T, \boldsymbol{\Phi}_2^T, \ldots, \boldsymbol{\Phi}_{2K}^T\right]^T. \tag{10}$$

The $2MT \times 2MT$ covariance matrix based in one training block formulated from the received data as:

$$\mathbf{R}_{cov} := \mathbf{E}\left\{\overline{\boldsymbol{Y}}\,\overline{\boldsymbol{Y}}^T\right\}. \tag{11}$$

For the case of multiple training frames per user J, the reshaped real received vector of size $2JMT \times 1$ is formed by concatenating the corresponding of the reshaped real received vector of size $2MT \times 1$ as:

$$\boldsymbol{y} = \begin{bmatrix} \boldsymbol{y_1} \\ \boldsymbol{y_2} \\ \vdots \\ \boldsymbol{y_J} \end{bmatrix}. \tag{12}$$

The $2JMT \times 2JMT$ covariance matrix based in J training block is formulated from the received data as:

$$\mathbf{R}_{cov} := \mathbf{E}\left\{\boldsymbol{y}\boldsymbol{y}^T\right\}. \tag{13}$$

The problem is to estimate the P channel matrices $\mathbf{H}_p$, each of size $N \times M$ or simply $\boldsymbol{g}$ in Eq. (5) based on the covariance matrix in Eq. (13) and the J transmitted pilot data frames each of size $K \times 1$ symbols $\mathbf{s}_p$ per user, to form an overall pilot vector of size $PJK \times 1$. This is a classical estimation problem which was studied earlier in [22]. The novelty of this article is to utilize DL methods to build a CNN that can extract the CSI.

The input data for the neural network constitutes a carefully structured 4D array of single-channel images. Each image of size $2JMT \times (2JMT+1)$ is derived from the concatenation of the covariance matrix with the overall pilot vector after ensuring that the vector is reshaped appropriately by padding zeros to the pilot vector to match the matrix dimensions.

In large number of users scenario, were the pilot vector couldn't be accommodated within $2JMT$, the covariance matrix is extended by padding a complete rows of zeros and the image size will be given by $2PJK$, so in general the image size is given by $\max(2JMT, 2PJK) \times (1 + \max(2JMT, 2PJK)$.

The neural network's output data is a 3D array of vectors representing the accurate CSI. Each vector operates as the ground truth for one instance of the training set, with the entire array structured to match the number of Monte Carlo simulations. Such data is essential for training the neural network to estimate the channel characteristics accurately based on the input signal data.

This format allows CNN to process and learn from the signal data, ultimately aiming to perform accurate channel estimation based on the provided training set. The CNN architecture is designed to learn and extract features from the input signal data images, facilitating accurate channel estimation. It consists of multiple convolutional, batch normalization, ReLU and pooling layers, followed by a fully connected layer and a regression layer optimized using SGDM.

3. Development of The Proposed Methods

Recently, there has been a significant and growing shift towards exploring machine learning models as alternatives to conventional signal processing-based channel estimation methods. These models are designed to learn the channel response from large sets of measurements more accurately than traditional statistical-based methods, reflecting the current trends in our field.

Adopting machine learning channel estimation models presents multiple potential advantages. First, these models may be more precise than their traditional statistical-based counterparts, particularly in complex or fast-changing channels. Second, they can also be more flexible and learn to adjust themselves over time in ways that traditional statistical-based models cannot. Finally, machine learning models might estimate a wider variety of channel properties than can be

Tab. 1. MU-MIMO parameters.

Variable	Description	Value
P	Number of users/transmitters	2
p	p-th user$= 1, 2, \ldots, P$	1, 2
N	Number of antennas per user	4
M	Number of antennas at the receiver	4
$\boldsymbol{s}_p$	p-th user information/pilot data vector	3×1
$\mathbf{K}$	Length of $\boldsymbol{s}_p$	3
$\mathbf{T}$	The block length of OSTBC code	4
$\mathbf{J}_t$	Number of pilot blocks	2, 5, 10
$2JMT \times (2JMT+1)$	Input image size	

done using classical and statistics-based channel estimation models.

However, using ML models for channel estimation comes with its own set of challenges. First, these models can be resource-intensive when being built and in operation. Additionally, they require proper training sets for good results because poor input could result in poor outputs. Also, the performance of machine learning models is influenced by the similarity of the training signals. Finally, such systems are complex to interpret, making it difficult to understand how they operate. The most straightforward DL architecture is the multi-layer perceptron (MLP), which consists of a succession of fully connected layers separated by activation functions. Despite their simplicity, MLP remains an essential tool when the dimension of the signal to be processed is manageable.

3.1. Datasets Generation

The research's first aim is to generate a synthetic dataset which imitates the actual CSI complexity. CSI of various SNRs having different numbers will be utilized. We choose ¾-rate OSTBC ($K = 3, T = 4$). In this case study, we shall have two users ($P = 2$) participating in a multi-access channel where each sender employs four transmit antennas ($N = 4$) to communicate with another partner via four receive antennas each ($M = 4$). There is an assumption that one of the users has a SNR that is stronger than the others by 2.5 dB. The estimation of CSI requires the receiver to know only two to ten blocks of information assuming $J_t = 2, 5$, or 10. Each case was simulated for 1000 independent channel realizations. All used parameters are given in Tab. 1.

3.2. Dataset Preprocessing

The dataset of the received signal went through preprocessing procedures to form an image of size $2JMT \times (2JMT + 1)$, which is derived from the concatenation of the covariance matrix with the overall training vector after ensuring that the vector is reshaped appropriately by padding zeros to match

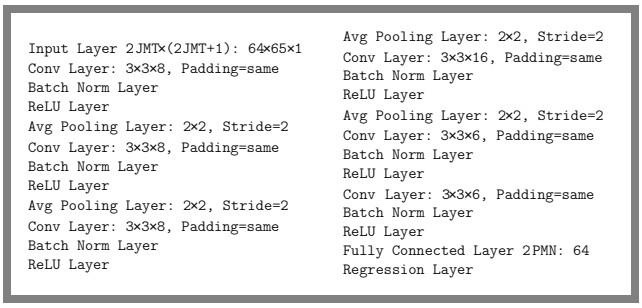

Fig. 2. Deep neural network architecture (the diagram was automatically generated by Matlab).

the matrix dimensions to guarantee consistency and make model training easier. After that, the dataset was split into training (80%) and testing (20%) groups [31]. The output layer utilizes a linear activation function to accommodate this variability, allowing for the flexibility required to capture such experiment-specific nuances. The size of the learning set is considered 1000.

3.3. Neural Network Architecture

The convolutional neural network (CNN) architecture based on two pilot frames, as shown in Fig. 2, is designed explicitly for channel estimation in MU-MIMO systems. The CNN architecture designed for this task begins with an input layer that accepts grayscale images. Six convolutional blocks follow this, each performing specific operations to extract, and process features from the input data. The CNN architecture ends with an out layer of a fully connected layer to output the estimated channel information. A regression loss function is employed to train the neural network to accurately learn the mapping of the input images to the accurate channel information.

3.4. Input Layer

The CNN input layer is an image input layer with dimensions $[64, 65, 1]$ representing the reshaped data designed to handle grayscale images derived from the received signal matrices and training vectors. It acts as the first point of data entry, identifying the shape and configuration of the input data for the network to process.

3.5. Hidden Layers

The hidden layers consist of the following layers.

First convolutional block. Three convolutional layers are stacked one after the other. Each layer applies convolutional filters to extract spatial features from the input data. The convolutional filters are 3×3 in size, and the number of filters gradually increases from 8 to 32. Followed by batch normalization, which is applied after each convolutional layer. The previous layer's activations are normalized, which helps to stabilize and accelerate the training process. Every batch normalization layer is immediately succeeded by ReLU activating functions. ReLU introduces non-linearity to the network by setting negative values to zero, allowing the model to learn complex patterns in the data. Four average pooling

layers are added to down sample the spatial dimensions of the feature maps. Each pooling layer reduces the spatial resolution by a factor of 2, helping to reduce computation and control overfitting.

Second convolutional block. Like the first convolutional layer, this layer further refines the features extracted by the previous layers. Followed by batch normalization, ReLU activation and average pooling layers as described above.

Third Convolutional block. It continues the process of feature extraction with additional filters. Followed by batch normalization, ReLU activation and average pooling layers.

Fourth convolutional block. Increases the number of filters to 16, maintaining the same filter size and padding. This block also includes batch normalization, ReLU activation, and average pooling layers. The fifth and sixth convolutional blocks reduce the filter count to 6, while maintaining the same architecture with convolution, batch normalization and ReLU activation layers.

Fifth convolutional layer. It further increases the number of filters to 32. It is followed by batch normalization and ReLU activation layers.

Sixth convolutional layer. It maintains the number of filters at 32 for deeper feature extraction. Four average pooling layers are added to down sample the spatial dimensions of the feature maps. Each pooling layer reduces the spatial resolution by a factor of 2, helping to reduce computation and control overfitting.

3.6. Output Layer

The output layer consists of two layers:

Fully connected layer. After the convolutional blocks, the network includes a fully connected layer with 64 output neurons, which connects all input neurons to each output neuron, facilitating the combination of extracted features into final predictions.

Regression layer. It is used to compute the loss for regression tasks, making this architecture suitable for continuous output prediction, such as estimating channel information. This structured approach allows the CNN to learn hierarchical feature representations effectively, enhancing its predictive accuracy from input images to true channel information.

3.7. Neural Network Training Process

The training process for the neural network is designed to estimate the CSI. The critical parameters set for training include the mini-batch size, maximum epochs, and the initial learning rate. Specifically, we choose the size of the mini batch to be 128 in order to settle the trade-off between learning speed and stability.

The network is trained over a maximum of 30 epochs, ensuring sufficient iterations for convergence. A piecewise learning rate schedule starts with an initial learning rate of 0.004. This schedule decreases the learning rate by a factor of 0.1 every 20 epochs. This approach helps fine-tune the model as it approaches convergence, allowing for more precise

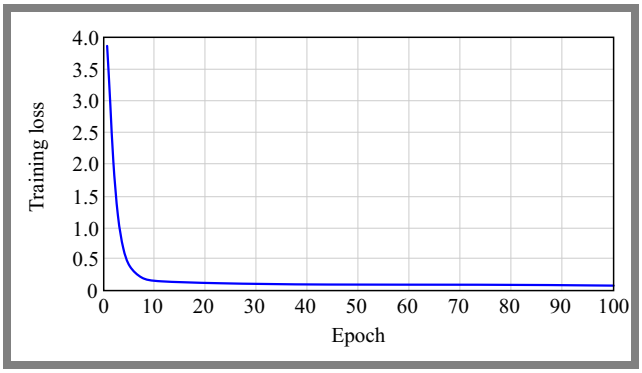

Fig. 3. Training loss function.

adjustments to the weights in the later stages of training. The data is shuffled at the beginning of each epoch to ensure robust learning. This shuffling prevents the network from learning any spurious patterns related to the order of the training data.

We use validation data, evaluated every 100 iterations, to regularly assess the model's performance on unseen data. This typical validation makes it possible to assess the ability of the model to generalize and note any signs of early overfitting during the training process. The SGDM algorithm is used to optimize the network. We chose this optimizer because it accelerates convergence and improves training stability. SGDM minimizes the loss function by iteratively adjusting the network's weights. It incorporates momentum to avoid oscillations and helps navigate the optimization landscape more efficiently.

The loss function used in this context is the regression loss, which measures the difference between the predicted channel information and the actual values. By minimizing this loss, the network learns to make more accurate predictions, as shown in Fig. 3. Upon completing the training process, the neural network can make precise predictions on new, unseen data. This capability results from the network's ability to effectively estimate the channel response, demonstrating the model's learned expertise and robustness in dealing with varying channel conditions. Combining carefully chosen training parameters, validation techniques, and an effective optimization algorithm ensures that the neural network performs reliably in practical applications.

4. Simulation Results

We conducted extensive simulations for the purpose of validating the proposed method. OSTBC with a rate of $K = \frac{3}{4}$ ($K = 3, T = 4$) was under consideration. The multi-access scenario in question headed for two users ($P = 2$), each having $N = 4$ transmitting antennas and communicating through a single receiver with $M = 4$ receiving antennas.

We assumed that one of the two users had a transmit power that was 2.5 dB higher than the other and carried out each of these scenarios using $M_C = 1000$ independent channel realizations. The normalized root-mean-square error (NRMSE) for p-th user was:

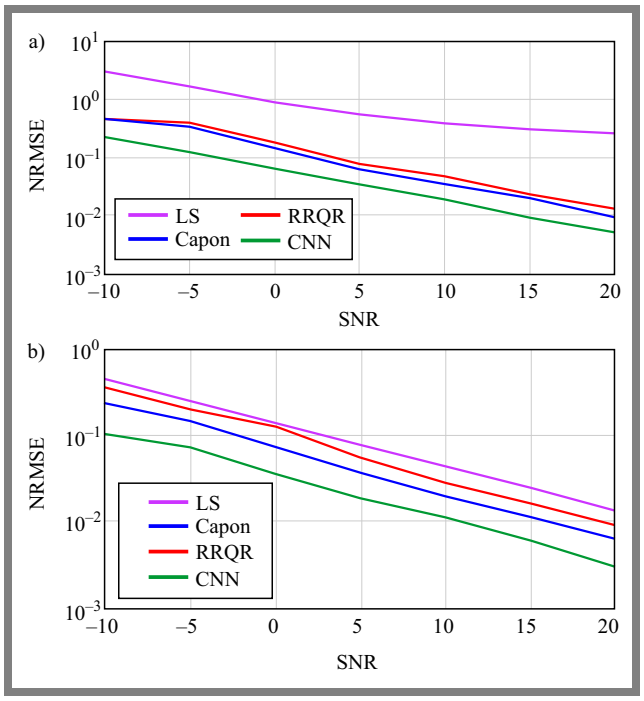

Fig. 4. Performance comparison of different methods for stronger transmitter using: a) two pilot frames and b) ten pilot frames.

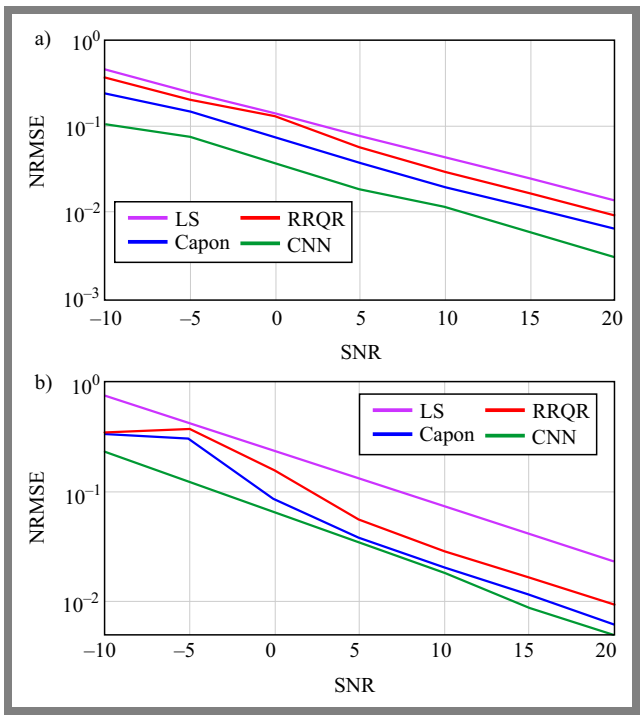

Fig. 5. Performance comparison of different methods for weaker transmitter using: a) two pilot frames and b) ten pilot frames.

$$NRMSE^p = \frac{1}{M_c}\sum_{n=1}^{M_c}\left(\sqrt{\frac{1}{2MN}\sum_{i=1}^{2MN}\left(\boldsymbol{h}_p(i)-\hat{\boldsymbol{h}}_p(i)\right)^2}\right). \tag{14}$$

A comparison of NRMSE on a more robust transmitter using two pilot frames is presented between the CNN estimator, LS, RRQR, and Capon methods at various SNR levels, as shown in Fig. 4a. The CNN method demonstrates a significant improvement over all conventional estimators. For instance, the CNN achieves an NRMSE of 0.001 at SNR of 11 dB, whereas the Capon estimator requires an SNR of 20 dB to reach the same NRMSE. Additionally, at an SNR of 5 dB, the NRMSE values for the CNN, Capon, RRQR, and LS methods are 0.0182, 0.0634, 0.0763, and 0.306, respectively. Similarly, Fig. 4b. indicates that using ten pilot frames provides a 5 dB advantage for the CNN over the Capon method in achieving the same NRMSE.

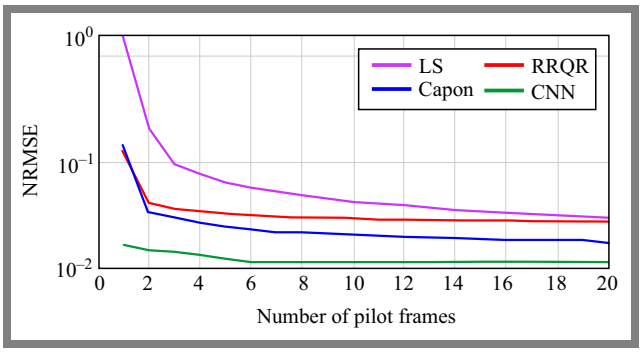

Fig. 6. Performance comparison of different methods of estimating for stronger transmitter at 10 dB SNR.

The NRMSE versus SNR are presented in Fig. 5a. and Fig. 5b for a system with a weaker transmitter's CNN estimator using two and ten pilot frames, respectively. Also, the performance of the CNN estimator is compared with the performance of the LS, RRQR, and Capon methods. It is observed that with two pilot frames, the performance of the CNN-based channel estimator significantly outperforms the performance of equivalent signal processing-based channel estimation models such as LS, Capon, and RRQR. The NRMSE versus the number of pilot frames at 10 dB is illustrated in Fig. 6. Smaller NRMSE is consistently shown by CNN-based estimator in comparison with other signal processing-based estimators as regards the number of employed pilot frames, thus indicating higher reliability of CNN-based estimator than other signal processing-based estimators. Indeed, the CNN method is better, even with a single pilot frame. Moreover, the number of pilot frames will influence the performance of the LS, Capon, and QR methods. Finally, the graph shows that as the number of pilot frames increases, the gradual drop-off occurs for the estimator's performance based upon CNN, which tends to maintain its stability.

5. Conclusion

This paper presents a novel deep-learning algorithm to estimate the channel state at the receiver end (CSIR) in an MU-MIMO system using OSTBC. The proposed CNN model, featuring six convolutional blocks for extracting features and one fully connected layer for outputting channel estimates, demonstrated better performance than other classical subspace-based estimators. We used just one pilot block and validated proposed approach by extensive computer simulations, which indicated significant improvements over classical

techniques such as least squares, Capon, and RRQR. Then, the usage of pilot blocks was extended up to 20 frames. These findings proved that this novel method based on deep learning can be relied upon for accurate CSIR estimation, thereby improving the signal decoding efficiency of multi-antenna communication systems.

References

[1] J. Kaur *et al.*, "Machine Learning Techniques for 5G and Beyond", *IEEE Access*, vol. 9, pp. 23472–23488, 2021 (https://doi.org/10.1109/ACCESS.2021.3051557).

[2] C. Xu *et al.*, "Two Decades of MIMO Design Tradeoffs and Reduced-complexity MIMO Detection in Near-capacity Systems", *IEEE Access*, vol. 5, pp. 18564–18632, 2017 (https://doi.org/10.1109/ACCESS.2017.2707182).

[3] Z. An *et al.*, "Blind High-order Modulation Recognition for Beyond 5G OSTBC-OFDM Systems via Projected Constellation Vector Learning Network", *IEEE Communications Letters*, vol. 26, no. 1, pp. 84–88, 2022 (https://doi.org/10.1109/LCOMM.2021.3124244).

[4] R. Chataut and R. Akl, "Massive MIMO Systems for 5G and Beyond Networks – Overview, Recent Trends, Challenges, and Future Research Direction", *Sensors*, vol. 20, no. 10, art. no. 2753, 2020 (https://doi.org/10.3390/s20102753).

[5] C.M. Lau, "Performance of MIMO Systems Using Space Time Block Codes (STBC)", *Open Journal of Applied Sciences*, vol. 11, pp. 273–286, 2021 (https://doi.org/10.4236/ojapps.2021.113020).

[6] Y. Huo *et al.*, "Technology Trends for Massive MIMO Towards 6G", *Sensors*, vol. 23, no. 13, art. no. 6062, 2023 (https://doi.org/10.3390/s23136062).

[7] H. Sadia, H. Iqbal, and R.A. Ahmed, "MIMO-NOMA with OSTBC for B5G Cellular Networks with Enhanced Quality of Service", *2023 10th International Conference on Wireless Networks and Mobile Communications (WINCOM)*, Istanbul, Turkiye, 2023 (https://doi.org/10.1109/WINCOM59760.2023.10322880).

[8] M.G. Gaitán, G. Javanmardi, and R. Sámano-Robles, "Orthogonal Space-time Block Coding for Double Scattering V2V Links with LOS and Ground Reflections", *Sensors*, vol. 23, no. 23, art. no. 9594, 2023 (https://doi.org/10.3390/s23239594).

[9] R.-Y. Wei, K.-H. Lin, and J.-R. Jhang, "High-diversity Bandwidth-efficient Space-time Block Coded Differential Spatial Modulation", *IEEE Open J. of the Communications Society*, vol. 5, pp. 3331–3339, 2024 (https://doi.org/10.1109/OJCOMS.2024.3404427).

[10] J. Park, J. Lee, I. Ha, and S.-K. Han, "Mitigation of Dispersion-induced Power Fading in Broadband Intermediate-frequency-over-Fiber Transmission Using Space-time Block Coding", *2024 Optical Fiber Communications Conference and Exhibition (OFC)*, San Diego, USA, 2024 (https://doi.org/10.1364/OFC.2024.M3F.4).

[11] J. Zhang *et al.*, "Automatic Identification of Space-time Block Coding for MIMO-OFDM Systems in the Presence of Impulsive Interference", *IEEE Transactions on Communications*, 2024 (https://doi.org/10.1109/TCOMM.2024.3382326).

[12] V. Singh *et al.*, "Diversity Combining Scheme for Time-varying STBC NGSO Multi-satellite Systems", *IEEE Communications Letters*, vol. 28, no. 4, pp. 882–886, 2024 (https://doi.org/10.1109/LCOMM.2024.3359329).

[13] M. Li, F. El Bouanani, S. Muhaidat, and M. Dianati, "Secure STBC-Aided NOMA in Cognitive IIoT Networks", *IEEE Internet of Things Journal*, vol. 11, no. 1, pp. 1256–1271, 2024 (https://doi.org/10.1109/JIOT.2023.3288452).

[14] H. Jafarkhani and V. Tarokh, "Multiple Transmit Antenna Differential Detection from Generalized Orthogonal Designs", *IEEE Transactions on Information Theory*, vol. 47, no. 6, pp. 2626–2631, 2001 (https://doi.org/10.1109/18.945280).

[15] H. Li, X. Lu, and G.B. Giannakis, "Capon Multiuser Receiver for CDMA Systems with Space-time Coding", *IEEE Transactions on Signal Processing*, vol. 50, no. 5, pp. 1193–1204, 2002 (https://doi.org/10.1109/78.995086).

[16] D. Reynolds, X. Wang, and H.V. Poor, "Blind Adaptive Space-time Multiuser Detection with Multiple Transmitter and Receiver Antennas", *IEEE Transactions on Signal Processing*, vol. 50, no. 6, pp. 1261–1276, 2002 (https://doi.org/10.1109/TSP.2002.1003052).

[17] S. Zhou, B. Muquet, and G.B. Giannakis, "Subspace-based (semi-) Blind Channel Estimation for Block Precoded Space-time OFDM", *IEEE Transactions on Signal Processing*, vol. 50, no. 5, pp. 1215–1228, 2002 (https://doi.org/10.1109/78.995088).

[18] P. Stoica and G. Ganesan, "Space-time Block Codes: Trained, Blind, and Semi-blind Detection", *Digital Signal Processing*, vol. 13, no. 1, pp. 93–105, 2003 (https://doi.org/10.1016/S1051-2004(02)00009-X).

[19] S. Shahbazpanahi, A.B. Gershman and G.B. Giannakis, "Semi-blind Multi-user MIMO Channel Estimation Based on Capon and MUSIC Techniques", *IEEE Int. Conference on Acoustics, Speech, and Signal Processing (ICASSP '05)*, Philadelphia, USA, 2005 (https://doi.org/10.1109/ICASSP.2005.1416123).

[20] S. Shahbazpanahi *et al.*, "Minimum Variance Linear Receivers for Multi-access MIMO Wireless Systems with Space-time Block Coding", *IEEE Transactions on Signal Processing*, vol. 52, no. 12, pp. 3306–3313, 2004 (https://doi.org/10.1109/TSP.2004.837442).

[21] S. Shahbazpanahi, A.B. Gershman, and J.H. Manton, "Closed-form Blind MIMO Channel Estimation for Orthogonal Space-time Block Codes", *IEEE Transactions on Signal Processing*, vol. 53, no. 12, pp. 4506–4517, 2005 (https://doi.org/10.1109/TSP.2005.859331).

[22] G. Hiren *et al.*, "Semiblind Multiuser MIMO Channel Estimators Using PM and RRQR Methods", *2009 Seventh Annual Communication Networks and Services Research Conference*, Moncton, Canada, 2009 (https://doi.org/10.1109/CNSR.2009.11).

[23] M. Qasaymeh, "Semi Blind Estimation for Multiuser MIMO Channel Using Orthogonal Space Time Block Coding", *Journal of Theoretical and Applied Information Technology*, pp. 37–42, 2010.

[24] H. Tataria *et al.*, "6G Wireless Systems: Vision, Requirements, Challenges, Insights, and Opportunities", *Proceedings of the IEEE*, vol. 109, no. 7, pp. 1166–1199, 2021 (https://doi.org/10.1109/JPROC.2021.3061701).

[25] N. Kato *et al.*, "Ten Challenges in Advancing Machine Learning Technologies Toward 6G", *IEEE Wireless Communications*, vol. 27, no. 3, pp. 96–103, 2020 (https://doi.org/10.1109/MWC.001.1900476).

[26] S. Liu, T. Wang, and S. Wang, "Toward Intelligent Wireless Communications: Deep Learning-based Physical Layer Technologies", *Digital Communications and Networks*, vol. 7, no. 4, pp. 589–597, 2021 (https://doi.org/10.1016/j.dcan.2021.09.014).

[27] M.M. Qasaymeh, "A Novel Machine Learning Approach for Blind Carrier Offset Estimation in OFDM Systems", *International Journal of Electrical and Electronic Engineering & Telecommunications*, vol. 13, no. 4, pp. 286–292, 2024 (https://doi.org/10.18178/ijeetc.13.4.286-292).

[28] M.M. Qasaymeh and A.F. Aljaafreh, "Joint Time Delay and Frequency Estimation Based on Deep Learning", *Journal of Communications*, vol. 19, no. 1, pp. 1–6, 2024 (https://doi.org/10.12720/jcm.19.1.1-6).

[29] P. Dong *et al.*, "Deep CNN-based Channel Estimation for mmWave Massive MIMO Systems", *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 5, pp. 989–1000, 2019 (https://doi.org/10.1109/JSTSP.2019.2925975).

[30] M.M. Qasaymeh, A.A. Alqatawneh, and A.F. Aljaafreh, "Channel Estimation Methods for Frequency Hopping System Based on Machine Learning", *Journal of Communications*, vol. 19, no. 3, pp. 143–151, 2024 (https://doi.org/10.12720/jcm.19.3.143-151).

[31] C.-J. Chun, J.-M. Kang, and I.-M. Kim, "Deep Learning-based Channel Estimation for Massive MIMO Systems", *IEEE Wireless Communications Letters*, vol. 8, no. 4, pp. 1228–1231, 2019 (https://doi.org/10.1109/LWC.2019.2912378).

[32] Z. Miyuan and C. Xibiao, "Channel Estimation for mmWave Massive MIMO Systems Based on Deep Learning", *IRO Journal on Sustainable Wireless Systems*, vol. 3, no. 4, pp. 226–241, 2022 (https://doi.org/10.36548/jsws.2021.4.003).

[33] M. Meenalakshmi, S. Chaturvedi, and V.K. Dwivedi, "Deep Learning-based Channel Estimation in 5G MIMO-OFDM Systems", *2022 8th International Conference on Signal Processing and Communication (ICSC)*, Noida, India, 2022 (https://doi.org/10.1109/ICSC56524.2022.10009461).

[34] A.K. Nair and V. Menon, "Joint Channel Estimation and Symbol Detection in MIMO-OFDM Systems: A Deep Learning Approach Using Bi-LSTM", *2022 14th International Conference on Communication Systems & Networks (COMSNETS)*, Bangalore, India, 2022 (https://doi.org/10.1109/COMSNETS53615.2022.9668456).

Mahmoud M. Qasaymeh, Ph.D.
Computer Engineering and Communication Department
https://orcid.org/0009-0006-0836-6113
E-mail: mqasaymeh@gmail.com
Tafila Technical University, Tafila, Jordan
https://www.ttu.edu.jo/en/

Ali Alqatawneh, Ph.D.
Computer Engineering and Communication Department
https://orcid.org/0000-0002-3084-6774
E-mail: ali.qatawneh@ttu.edu.jo
Tafila Technical University, Tafila, Jordan
https://www.ttu.edu.jo/en/

Mahmoud A. Khodeir, Ph.D.
Department of Electrical Engineering
https://orcid.org/0000-0002-3487-0237
E-mail: makhodeir@just.edu.jo
Jordan University of Science and Technology, Irbid, Jordan
https://www.just.edu.jo

Ahmad F. Aljaafreh, Ph.D.
Computer Engineering and Communication Department, Dept. of Computer Science and Software Engineering
https://orcid.org/0000-0002-8329-8804
E-mail: aljaafah@udmercy.edu
Tafila Technical University, Tafila, Jordan
https://www.ttu.edu.jo/en/
University of Detroit Mercy, Detroit, USA
https://www.udmercy.edu

TS-based Mobility-aware Multi-hierarchical Caching Model with Vehicle Clustering and Content Popularity Prediction

Radouane Baghiani[1], Lyamine Guezouli[2], and Ahmed Korichi[1]

[1]*LINATI Laboratory, Kasdi Merbah University, Ouargla, Algeria,*
[2]*LEREESI Laboratory, HNS-RE2SD, Batna, Algeria*

https://doi.org/10.26636/jtit.2024.3.1616

Abstract — This research is concerned with the fusion of artificial intelligence (AI) and machine learning within multi-hierarchical caching systems, specifically targeting vehicular and edge caching domains. This study introduces an innovative architecture harmonizing Thompson sampling learning-based caching policies with advanced vehicle clustering and content-popularity prediction methods (TS-MMCM). Simulations show substantial performance improvements and a big impact of the proposed approach on system efficiency in dynamic network environments. The proposal demonstrates a notable gain in cache hit rates and decreased latency levels, highlighting the potential of AI to improve caching techniques in dynamic network environments.

Keywords — clustering, Internet of Vehicles, machine learning, mobile edge computing, multi-hierarchical caching, Thompson sampling learning

1. Introduction

Caching is an essential computing approach that improves the performance and efficiency of a system by keeping frequently accessed data in fast access storage. It improves the rate at which data can be processed and greatly reduces the time needed to retrieve data for applications that require high performance and real-time capabilities. The cache handles recurring data requests, thereby minimizing the workload of the backend system, simultaneously enabling efficient scalability of services during high traffic periods.

Caching also enhances user experience by speeding up data retrieval, making it especially advantageous in web construction tasks, as it allows accelerated website loading. From an economic standpoint, data catching reduces the expenses associated with data transport and computations and facilitates offline data access, hence decreasing bandwidth utilization and alleviating network congestion. Furthermore, caching enhances scalability of the system by alleviating data access bottlenecks. Through the use of advanced techniques, caching also guarantees data consistency in dispersed systems, ensuring that the cached data remain accurate and valid.

Vehicular and edge caching are important developments in modern networking that aim to satisfy mobile users' growing need for data and services. Vehicular caching utilizes the storage capabilities of modern vehicles and the characteristics of vehicular networks to enhance both accessibility and availability of data. This approach is particularly beneficial in urban and mobile environments, as it minimizes access latency and improves user experience standards in applications such as video streaming, gaming, and augmented reality [1]. It enables real-time analysis and decision-making for IoT devices and services, allows for effective network expansion by distributing the data workload, and promotes sustainable networking by reducing the amount of energy used for transmitting data.

Vehicular and edge caching work together to achieve a decentralized and efficient networking model. This approach addresses such issues as high bandwidth requirements, minimal latency, and the necessity of ensuring uninterrupted connectivity in mobile and distributed environments.

Artificial intelligence (AI) and machine learning (ML) are transforming caching techniques in computing by improving efficiency and performance levels through predictive caching, adapting to changing data access patterns, as well as optimizing data storage. AI and ML enable caching systems to predict upcoming data queries by analyzing user behaviors and access patterns, thereby resulting in decreased latency. Furthermore, these technologies dynamically adapt caching algorithms in real time to optimize data delivery and resource allocation, which is particularly advantageous in content distribution networks (CDNs) and for handling the influx of data from IoT devices and edge computing. AI and ML enhance caching systems by identifying abnormalities that signal potential security risks or malfunctions, thereby enabling remedial measures to be taken in a timely manner. Tailored caching in web and mobile applications customizes content delivery based on individual user preferences, thus resulting in improved user experience.

In this study, a long short-term memory (LSTM) model is utilized to forecast content request numbers in a time series

This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) License.
For more information, see https://creativecommons.org/licenses/by/4.0/

tailored for Internet of Vehicles (IoV) scenarios. It focuses on caching decisions and proposes a reinforcement learning-based algorithm called the Thompson-based mobility-aware multi-hierarchical caching model with vehicle clustering and content popularity prediction methods (TS-MMCM). This model aims to enhance cache hit ratios by predicting content requests, applying the k-means model to cluster cars depending on their speed and position in order to stabilize communication, forecasting request numbers via the LSTM model, estimating content popularity with the Zipf model, and using reinforcement learning to optimize caching decisions.

The combination of Thompson sampling (TS) with vehicle clustering and content-popularity prediction methods enhances caching rules in dynamic situations, such as CDNs, edge computing, and vehicular networks. This method integrates the adaptive learning of TS, a Bayesian approach renowned for effectively managing the trade-off between exploration and exploitation, with clustering to categorize comparable content or requests, and predicting content popularity. The objective is to improve the effectiveness of caching by utilizing real-time predictions and user behaviors in order to determine which information to cache.

The main benefits of this technique include the ability to make detailed caching decisions based on the popularity of content within certain groups, the flexibility to adapt caching policies in real time according to changes in content demand, and the emphasis on caching content that is expected to be required in the future. This method not only enhances resource utilization, but also amplifies user experience by customizing cached information based on the interests of specific groups or locations, resulting in expedited content delivery and reduced network congestion.

The remainder of this paper is organized as follows. In Section 2, we discuss the related literature. Section 3 outlines the proposed system model, while Section 4 presents the formulation of the problem. The proposed caching strategy is detailed in Section 5. Section 6 discusses the simulation results. Finally, Section 7 concludes the paper and suggests future research.

2. Related Works

Several studies have employed a reinforcement learning approach to enhance caching algorithms. Reference [2] addressed the device-to-device (D2D) caching issue as a multi-agent multi-armed bandit (MAB) learning problem. It employed Q-learning to determine the optimal coordination of caching decisions among multiple agents to maximize caching benefits. Paper [3] utilized a recurrent neural network (RNN) to predict user behavior and content popularity. A Q-learning strategy that employs learning automata was proposed for cooperative caching. This considers the popularity of content and the positions of mobile users to select the best options in a random and unchanging environment.

In [4], the authors investigated user terminal (UT) edge caching in D2D-enabled cellular networks, emphasizing content popularity and UT positioning. The problem of stochastic games was modelled and solved using the proposed multi-agent cooperative alternating Q-learning (CAQL) algorithm, allowing UTs to update caching placement policies for better performance. Article [5] explored the tidal effect in a mobile edge computing (MEC) network with multi-user and multicast data. It formulated the problem as an infinite-horizon average cost Markov decision process, aiming to optimize bandwidth usage and reduce data transmission. The problem was restated in reinforcement learning to learn file popularity and user requests. The design of Q-learning and a deep Q-network (DQN) was proposed, providing insights into the network caching design.

2.1. Caching in Vehicles

Caching-related issues arising from vehicle mobility are different from those found in cellular networks. The authors of [6] investigated the idea of automobiles acting as moving cache nodes and creating an automobile cloud to distribute the requested content to consumers using 6G. To prepare edge smart technologies for an upcoming 6G vehicle network, work [7] suggested utilizing parked vehicles as additional edge nodes. Such a solution will provide abundant resources in conjunction with existing ground infrastructure.

The architecture proposed in paper [8] consists of three layers: an airship, UAVs, and vehicles, all of which are used for vehicular caching. The authors employed the DQN method to obtain the most advantageous caching approach. In study [9], a sophisticated machine learning method was applied to information-centric networking-based vehicular networks. This technique expects future user requests by predicting their future evaluations of videos.

In [10], deep reinforcement learning (DRL) and federated learning (FL)-based content caching methods were proposed for efficient transmission tasks meeting the applicable latency requirements. These methods optimize latency by considering regional preferences and latency constraints. Multi-agent reinforcement learning (MARL) and FL were used to forecast content popularity and determine cached content in each area.

Another study in [11] proposed a deep learning system for content caching in autonomous vehicles that makes use of MEC servers to cache high-probability content and relies on a multi-layer perceptron (MLP) to predict content demands in certain regions. A convolutional neural network (CNN) was employed for the prediction of passenger profiles, while the self-driving car utilized binary classifications and the k-means approach to choose pertinent infotainment content for downloading and caching. In [12], researchers focused on issues pertaining to automated vehicle control and proactive caching in roadside units. This study employed deep reinforcement learning to enhance efficiency and quality of experience (QoE) by implementing proactive caching actions.

The authors of [13] proposed caching for infotainment content in self-driving cars based on MEC servers and utilizing CNN-

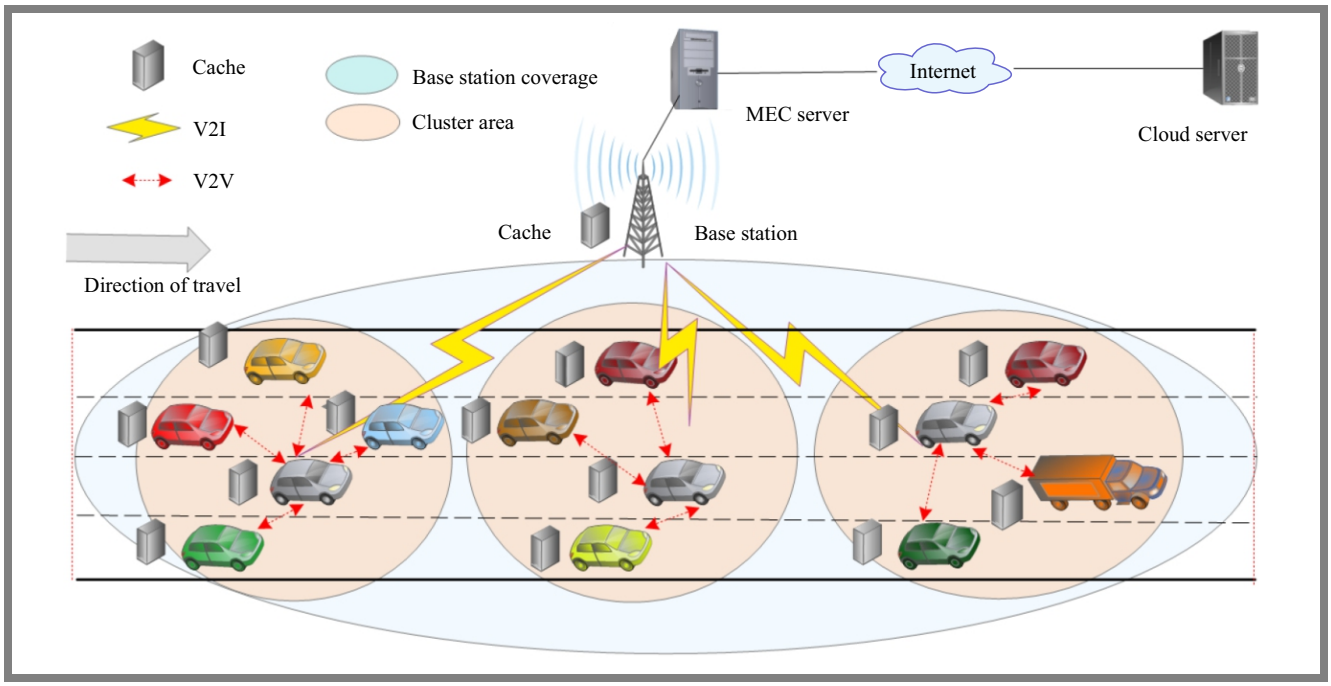

Fig. 1. Proposed architecture.

learned passenger features for high expected probabilistic values in their areas.

2.2. Caching in Edge

Recently, numerous research projects have been focusing on edge caching. In [14], a learning-based cooperative content caching policy was considered for the MEC architecture in situations where user preferences are unknown and only the demands for historical content are visible. This study proposes the MARL-based approach to address the cooperative content caching problem by modeling it as a multi-agent multi-armed bandit problem.

In [15], the authors proposed a novel edge-assisted intelligent caching framework that improved the cache hit rate. This framework can independently develop a caching technique in real time by analyzing the request sequence without requiring preliminary data processing or engineering features. Another study described in [16] investigated the prediction of spatial content preferences using distributed learning approaches and mobility predictions.

In contrast to the majority of other studies, paper [17] applied deep reinforcement learning to collaborative caching and set varying content sizes, while in [18], a collaborative caching approach for mobile edge computing servers incorporated multi-agent reinforcement learning.

2.3. Combining Vehicle and Edge Caching

Recent studies have integrated local vehicle and edge device caches, leveraging user and regional preferences to understand and optimize vehicle usage. The architecture proposed in [19] suggested a collaborative caching approach for the IoV by utilizing content request prediction (CCCRP) to minimize latency in obtaining content. The authors employed k-means clustering to group vehicles, LSTM networks to predict content requests, and reinforcement learning to improve caching decisions, hence enhancing the quality of service for vehicle requests.

The study described in [20] presented a cooperative caching approach for downloading content in reinforcement learning, applying the k-means algorithm for vehicle clustering and long short-term memory for content prediction, as well as the DRL algorithm and DQN to determine the most effective caching technique.

3. Proposed Architecture

In this section, we present an architecture for multilevel data caching, as depicted in Fig. 1, which includes cluster member users, cluster heads, base stations (BSs), and edge servers to cache various types of content. Each edge server described in this proposal can handle the BSs located within its coverage area. Several users were present in each base station's dominant region. This is because the emphasis is placed on the cache capacity of the vertical architecture.

Vehicles covered by the same BS insurance were grouped into clusters according to their mobility traits. Cluster heads are assigned to each cluster, and the cluster head selection remains constant in each time frame. Vehicles within the range of a single BS can form multiple non-overlapping clusters, with each cluster uniquely assigned to one BS. The cluster head for each vehicle corresponds to the cluster it is part of and is marked as $H_{i,1}$, where i represents the index of clusters within the coverage area of the BS.

Further, vehicles within the cluster may also be denoted by $H_{i,J}$, where j ($j \geqslant 2$) represents the index of the vehicle within cluster i. For example, the cluster head $H_{i,1}$ serves cluster members. All notations are summarized in Tab. 1. In this cache structure, when the vehicle user initiates a request, the local cache is first checked to locate the requested content. When the content is available in the local cache, it becomes directly accessible.

Otherwise, the user seeks content from $H_{i,1}$. If the content is cached in $H_{i,1}$, the user retrieves it directly. If it is not found in $H_{i,1}$, then the user sends a request to the BS covering their location. The user can instantly access the content from the covering BS if it is cached. If cache retrieval is unsuccessful, the request is forwarded to the edge server and, if not found, the request is forwarded to the cloud server.

The reliance on mobility traits to group vehicles into clusters enhances the reliability of communication lines between vehicles connected to the same base station. A cluster head is assigned to each cluster. It is important to remember that the cluster head is always chosen in the same manner for consecutive time slots. Normalized characteristics, such as the position and speed of each vehicle, have been derived. This process aims to group N_v–generated vehicles into M_c vehicle clusters.

Because every vehicle in the cluster, including the cluster head, has the ability to cache content, the cluster heads can establish one-hop communication with other vehicles in the same cluster. However, no communication can be established between any two vehicles that are part of different clusters.

3.1. Caching Model

Set $L = \{1, 2, \ldots, l\}$ represents the content repository, where l is the total amount of content. The cloud server stores the complete content repository L, whereas the edge server collaborates with the base stations, cluster heads, and vehicle users to cache the content. Each content has a distinct size, represented by $Z = \{z_1, z_2, \ldots, z_l\}$, where l is an element of L.

We grouped these content types into n categories, denoted by $cat = \{\text{category } 1, \text{category } 2, \ldots, \text{category } m\}$. The edge server has a local cache with capacity C_{edge}. The base station offers a cache capacity denoted as C_{BS} and the cluster head has a local cache with a capacity of $C_{Hi,1}$, where M_c is the number of clusters. $V = \{v_1, v_2, \ldots, V_u\}$ represents a group of users, assuming that certain content is stored in the user's local cache with a cache capacity of $C_{Hi,j}$. When $S(l, edge) = 1$, content l is stored at the edge, otherwise $S(l, edge) = 0$. For $S(l, BS) = 1$, content l is stored in the BS. $S(l, BS) = 0$ signifies that the content is not stored in the BS.

For $S(l, H_{i,1}) = 1$, the content l is stored in the cluster head, otherwise $S(l, H_{i,1}) = 0$ denotes the content l is not stored in $H_{i,1}$. Similarly, $S(l, H_{i,j}) = 1$ implies that content l is cached in the local cache of users in cluster i, whereas $S(l, H_{i,j}) = 0$ indicates otherwise. It is crucial to ensure that the cache size of any device does not exceed its total cache capacity at any given time t:

$$\begin{cases} \sum\limits_{e=1}^{l} z_e \cdot S(e, edge) \leqslant C_{edge}, & \forall\, e \in L \\ \sum\limits_{r=1}^{l} z_r \cdot S(r, BS) \leqslant C_{BS}, & \forall\, r \in L \\ \sum\limits_{b=1}^{l} z_b \cdot S(b, H_{i,1}) \leqslant C_{H_{i,1}}, & \forall\, b \in L \\ \sum\limits_{d=1}^{l} z_d \cdot S(d, H_{i,j}) \leqslant C_{H_{i,j}}, & \forall\, d \in L \end{cases},$$

where z_e, z_r, z_b, and z_d represent the size of the cached content at each piece of equipment or location indexed by e, r, b, and d, respectively.

Tab. 1. Summary of notations used.

Symbol	Description
BS	Base station
$H_{i,1}$	Cluster head
$H_{i,j}$	Cluster member
Nv	Total number of vehicles in the system
Mc	Number of clusters grouping the vehicles
L	Content repository
Z	Size of each file repository
cat	Category of content
V	Set of vehicles
C_{edge}	Cache capacity of edge cache
C_{BS}	Cache capacity of base station cache
$C_{H_{i,1}}$	Cache capacity of cluster head cache
$C_{H_{i,j}}$	Cache capacity of local cache
$S(l, edge)$	Cache state of the file l in the edge
$S(l, BS)$	Cache state of the file l in the BS
$S(l, H_{i,1})$	Cache state of the file l in the cluster head
$S(l, H_{i,j})$	Cache state of the file l in the cluster member
$HitR_t^{edge}$	Cache hit rate of the edge
$HitR_t^{BS}$	Cache hit rate of the BS
$HitR_t^{H_{i,1}}$	Cache hit rate of the cluster head
$HitR_t^{H_{i,j}}$	Cache hit rate of the cluster member
Θ_{edge}	Caching edge policy
$\Theta_{H_{i,1}}$	Caching cluster head policy
$\Theta_{H_{i,j}}$	Caching cluster member policy
P_l	Probability of each file
$L_{i,l}(t)$	Location of vehicle
$S_{i,l}(t)$	Speed of vehicle
CL	Connectivity lifetime

$S(e, edge), S(r, BS), S(b, H_{i,1})$, and $S(d, H_{i,j})$ denote the cache state of the content for the edge, BS, cluster head, and cluster members, respectively, with the respective indices. $C_{edge}, C_{BS}, C_{H_{i,1}}$, and $C_{H_{i,j}}$ represent the maximum cache capacities for the edge, BS, cluster head, and cluster members, respectively, owing to the constraint of the restricted caching capacity of the edge, BS, $H_{i,1}$, and $H_{i,j}$.

We propose a cache-placement technique that relies on the popularity of content. The chance of a given piece of content being requested is impacted by its popularity ranking, with more popular content having a higher level of probability of being requested. Zipf distribution is commonly used to represent the probability of a content request. The likelihood of content being requested can be ascertained by employing the Zipf model which considers the number of requests for that content. We define R_l to represent content l that is ranked at the R-th level of popularity. The Zipf model is described as:

$$P(R, s, N) = \frac{1/R^s}{\sum_{q=1}^{N} \frac{1}{q^s}}, \quad q \in \{1, 2, \ldots, l\} , \tag{1}$$

where N represents the total amount of content and s is the exponent parameter that characterizes the Zipf distribution. A greater value of s indicates that the content queries are more focused on the highest ranking. Evaluation of caching strategies involves examining the cache hit rate to understand its benefits and drawbacks. The cache hit rate of the edge server, cluster head, and cluster members is:

$$HitR_t^{edge} = \sum_{e=1}^{l} P_e \cdot S(e, edge), \quad \forall\, e \in L , \tag{2}$$

$$HitR_t^{BS} = \sum_{r=1}^{l} P_r \cdot S(r, BS), \quad \forall\, r \in L , \tag{3}$$

$$HitR_t^{H_{i,1}} = \sum_{b=1}^{l} P_b \cdot S(b, H_{i,1}), \quad \forall\, b \in L , \tag{4}$$

$$HitR_t^{H_{i,j}} = \sum_{d=1}^{l} P_d \cdot S(d, H_{i,j}), \quad \forall\, d \in L . \tag{5}$$

The probability vector $P_l = [p_1, p_2, \ldots, p_l]$ indicates the likelihood of each content being required by vehicle users in the upcoming time period.

3.2. Content Request Model

All the components of the proposed system, including vehicle users, cluster heads, base stations, and edge servers, are capable of caching content. If a vehicle within the cluster requires specific content, there are five potential locations where it can be obtained:

- Local cache. The vehicle user's local cache is first checked to locate the requested content. If the required content is present in the local cache, it is directly accessible.
- Cluster head. If the content is not available in the local cache, the user seeks it from the cluster head. The requested content is directly retrieved from the cluster head if it is cached using vehicle-to-vehicle (V2V) connectivity.
- Base station. If the content is not located in the cluster head, the user sends the request to the BS that covers it. The user will instantly access the content from the BS if it is cached using vehicle-to-infrastructure (V2I) connectivity.
- Edge server. For content not found in the BS, the user transmits the query to the edge server through the BS. If cache retrieval remains unsuccessful, the request is forwarded to the cloud server. However, this results in greater latency.

4. Problem Formulation

The cache hit rate is an important performance metric used in caching techniques, as it indicates the extent to which the content in the cache aligns with user requests. Our objective was to create an online caching technique that maximizes the cache hit ratio. Owing to the large volume of files, we categorized the individual pieces of content into various types and assigned varying probabilities to each of the selection types. Once the file type is identified based on probability, one file of that type is selected for the cache. The objective of the caching method for each vehicle user, cluster, and edge is to maximize the cache hit rate as follows:

$$\max_{\Theta_{H_{i,j}}} HitR_t^{H_{i,j}}(\Theta_{H_{i,j}}) , \tag{6}$$

$$\max_{\Theta_{H_{i,1}}} HitR_t^{H_{i,1}}(\Theta_{H_{i,1}}) , \tag{7}$$

$$\max_{\Theta_{edge}} HitR_t^{edge}(\Theta_{edge}) . \tag{8}$$

Our technique aims to improve the cache hit rates at both the edge server and the car user layer. Considering the proposed architecture model, limitations regarding the vehicle users, cluster heads, and edge servers' ability to cache data, as well as the mobility constraint imposed by the vehicle, the proposed issue formulation is as follows:

$$\begin{aligned} &\max \{\theta_{H_{i,j}}, \theta_{H_{i,1}}, \theta_{edge}\} \quad \text{for } HitR_t^{H_{i,j}} , \\ &\max \{\theta_{H_{i,j}}, \theta_{H_{i,1}}, \theta_{edge}\} \quad \text{for } HitR_t^{H_{i,1}} , \\ &\max \{\theta_{H_{i,j}}, \theta_{H_{i,1}}, \theta_{edge}\} \quad \text{for } HitR_t^{edge} , \\ &\max \{\theta_{H_{i,j}}, \theta_{H_{i,1}}, \theta_{edge}\} \quad \text{for } HitR_t^{total} . \end{aligned} \tag{9}$$

Given the constraints set C_1, we have:

$$\begin{cases} S(l, edge) \\ S(l, BS) \\ S(l, H_{i,1}) \\ S(l, H_{i,j}) \end{cases} \in \{0, 1\}, \quad \forall\, l \in L .$$

Given the constraints set C_2, we have:

$$\begin{cases} \sum_{e=1}^{l} z_e \cdot S(e, edge) \leqslant C_{edge}, & \forall\, e \in L \\ \sum_{r=1}^{l} z_r \cdot S(r, BS) \leqslant C_{BS}, & \forall\, r \in L \\ \sum_{b=1}^{l} z_b \cdot S(b, H_{i,1}) \leqslant C_{H_{i,1}}, & \forall\, b \in L \\ \sum_{d=1}^{l} z_d \cdot S(d, H_{i,j}) \leqslant C_{H_{i,j}}, & \forall\, d \in L \end{cases},$$

where the initial constraint indicates that each content, represented by 1 or 0, can either be cached or not cached. Hence, values from the set of $[0, 1]$ must be assigned to $S(l, edge), S(l, BS), S(l, H_{i,1})$, and $S(l, H_{i,j})$. Furthermore, the last constraint guarantees that the cumulative amount of the content cached in the edge, BS, $H_{i,1}$, and $H_{i,j}$ does not exceed their respective caching capacities. Reinforcement learning (RL) can be a prospective approach for solving the caching optimization problem described above. We propose to use an effective RL approach to address this multi-agent decision-making issue.

5. Proposed Solution

The proposed approach consists of three components: clustering vehicles using k-means [21], predicting content requests using LSTM, and implementing the TS algorithm for content caching decisions.

5.1. Vehicle Clustering

Our method involves clustering vehicles according to their mobility to reduce the signaling load from V2V broadcasting and improve vehicular communication connections. In each cluster, the cluster head can directly communicate with other members located within the same cluster using single-hop V2V communication. This setup ensures efficient and direct communication pathways within the cluster, optimizing the network's overall functionality and reducing the need for complex, multi-hop communication strategies. Vehicles are classified according to their mobility characteristics, into distinct clusters within a single BS coverage area.

We created M_c clusters from a set of N_v generated vehicles by using normalized attributes, such as position and speed,

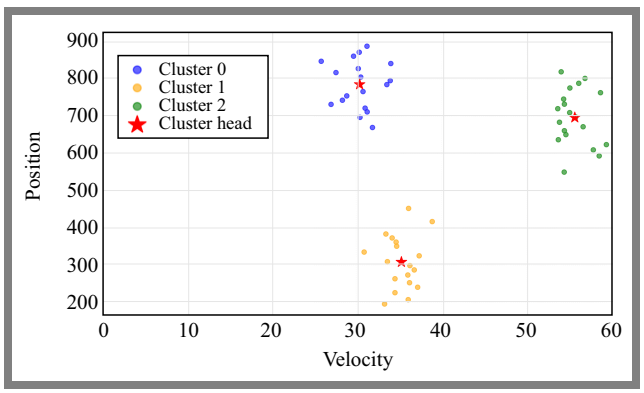

Fig. 2. Vehicle clustering with cluster heads.

obtained from each vehicle (Fig. 2). In vehicle clustering, the primary challenges include determining the clustering technique and choosing the cluster head. This study uses the k-means algorithm for clustering. As shown in Algorithm 1, four main steps are part of the process: choosing the M_c cluster centers, grouping the rest of the vehicles, updating the cluster centers, and repeating the process while the cluster centers remain constant.

To speed up clustering results, the algorithm first gives priority to choosing M_c vehicles that are widely separated. It is important to remember that the k-means algorithm depends on calculations of Euclidean distance. It is possible that the k-means cluster centers do not perfectly line up with the individual vehicles. Choosing cluster heads that match these clusters is a further refining step.

The cluster head provides the content cache source and manages access, requiring a reliable connection with other cars. Therefore, we utilized a distinct parameter called connectivity lifetime (CL) to choose cluster heads based on the durability of vehicle connections [22].

Let us represent the l-th car in cluster i as $N_{i,l}$. The vehicle's location and speed at time t are represented as $L_{i,l}(t)$ and $S_{i,l}(t)$, respectively. Vehicles occasionally send a "hello" beacon packet to other vehicles within the coverage area of the same BS. The hello beacon packet includes the vehicle's mobility information, such as its location $L_{i,l}(t)$ and speed $S_{i,l}(t)$. Therefore, every car can calculate its associations with other vehicles using this information. CL is the time it takes for a communication link to occur between two vehicles. Therefore, when the CL expires, the link between the two vehicles is expected to end, signifying that their relative distance is projected to reach the maximum broadcasting range, represented as δ, during this time frame.

The requirement for preserving a link between cars $N_{i,l}$ and $N'_{i,l}$ (where $l' \neq l$) follows this principle:

$$\left(L_{i,l}(t + \mathrm{CL}_{i,l}^{i,l'}) - L_{i,l'}(t + \mathrm{CL}_{i,l}^{i,l'})\right)^2 = \delta^2 \,. \tag{10}$$

If the vehicle maintains the same speed for the following time interval τ, the equation for the location of the car is:

$$L(t + \tau) = L(t) + \tau \cdot S(t) \,. \tag{11}$$

Therefore, Eq. (10) can be written as:

$$\left(L_{i,l}(t) - L_{i,l'}(t) + \mathrm{CL}_{i,l}^{i,l'}\left[S_{i,l}(t) - S_{i,l'}(t)\right]\right)^2 = \delta^2 \,, \tag{12}$$

where $CL_{i,l}^{i,l'}$ can be calculated by solving Eq. (12) in which the speed and location are 2D coordinates as follows:

$$\mathrm{CL}_{i,l}^{i,l'} = \frac{\delta^2 - |\mathbf{L}_{i,l}(t) - \mathbf{L}_{i,l'}(t)|^2}{|\mathbf{S}_{i,l}(t) - \mathbf{S}_{i,l'}(t)|} \,. \tag{13}$$

Once the clusters are determined using the k-means algorithm, the average CL for each vehicle $N_{i,l}$ in the corresponding cluster i can be computed as:

$$\overline{\mathrm{CL}}_{i,l} = \frac{1}{n_i - 1} \sum_{l'=1}^{n_i} \mathrm{CL}_{i,l}^{i,l'}, \quad l' \neq l \ . \tag{14}$$

Consequently, the vehicle with the highest average CL within this cluster may be designated as the cluster head. This choice ensures that the cluster head and every other vehicle within the same cluster enjoy optimal link stability.

Algorithm 1 Vehicle clustering based on the k-means approach.

Input: Vehicle set $V = \{V_1, V_2, \ldots, V_{nv}\}$, number of vehicle clusters M_c
Output: Vehicle cluster set $\{c_1, c_2, \ldots, c_{mc}\}$
Select initial cluster centers $\{t_1, t_2, \ldots, t_{mc}\}$
repeat
  for $l = 1$ to N_v **do**
    Determine the distance between N_l and each center t_i, denoted as $d_{i,l} = \|N_l - t_i\|^2$
    Assign N_l to the cluster whose cluster center is the closest
  end for
  for $i = 1$ to M_c **do**
    Get the fresh cluster centre $t_i^* = \frac{1}{|C_i|} \sum_{N_l \in C_i} N_l$
    if $t_i^* \neq t_i$ **then**
      Refresh the existing cluster center $t_i = t_i^*$
    else
      Maintain the current cluster center constant
    end if
  end for
until cluster centers become fixed

5.2. Content Popularity Prediction

The proposed design architecture emphasizes the importance of predicting content popularity, wherein the caching technique relies on content popularity. Many previous studies assume that the popularity of the content is pre-established or use the frequency of requests as a criterion for assessing content popularity. In practice, content popularity requires an element of freshness. In fact, vehicle content requests already showed a steady trend over time [23]. By analyzing past request-related data, it is possible to uncover the fundamental trends governing the frequency of content requests, which can then be used to forecast the number of requests in upcoming time intervals.

Here, this forecast request volume is defined as popularity of content. The use of a machine learning technique like RNN has enabled us to forecast the popularity of newly generated content. RNNs is capable of uncovering underlying temporal patterns by analyzing the historical request times for each content. However, conventional RNNs may encounter issues such as vanishing or exploding gradients with prolonged propagation. As a remedy, LSTM is favored, representing an optimized version of RNN [24].

The structure of an LSTM unit is shown in Fig. 3. Let μ_t and h_t represent the input and output data at time slot t, respectively. At each time slot t, the cell receives a new input μ. Three gates, named input gate i, forget gate f, and output gate o, are positioned within the cell. The value of these gates can be computed as follows:

$$f_t = \sigma\big(W_f \cdot [h_{t-1}, \mu_t] + b_f\big) \ , \tag{15}$$

$$i_t = \sigma\big(W_i \cdot [h_{t-1}, \mu_t] + b_i\big) \ , \tag{16}$$

$$o_t = \sigma\big(W_o \cdot [h_{t-1}, \mu_t] + b_o\big) \ , \tag{17}$$

where W_f, W_i, and W_o indicate the weight matrices, whereas b_f , b_i, and b_o represent the variable biases for the three gates. $\sigma(\cdot)$ is the non-linear activation sigmoid function.

The procedure for updating information in LSTM is as follows [25]. The forget gate f_t determines the portion of C_{t-1} to discard, such as:

$$f_t \cdot C_{t-1} \ . \tag{18}$$

The current state information is updated:

$$i_t \cdot \tilde{C}_t \ , \tag{19}$$

where

$$\tilde{C}_t = \mathrm{tgh}\left(W_c \cdot [h_{t-1}, \mu_t] + b_c\right) . \tag{20}$$

Here, b_c and w_c denote the variable bias and weight matrix of the memory cell, and $\mathrm{tgh}(\cdot)$ is the hyperbolic tangent function. The current unit is updated in the following manner:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \ . \tag{21}$$

Consequently, the output data is calculated as:

$$h_t = o_t \cdot \mathrm{tgh}(C_t) \ . \tag{22}$$

The LSTM approach is capable of predicting the volume of content requests in the future, allowing to optimize the coaching of content for car users. By using time series context, the LSTM model is provided with the count of content requests from the preceding time step $t-1$, as its input, denoted by:

$$\mu(t-1) = \big\{\mu_1(t-1), \mu_2(t-1), \ldots, \mu_Q(t-1)\big\} \ . \tag{23}$$

The LSTM output predicts the count of content requests for the forthcoming time period, as a set of values represented by:

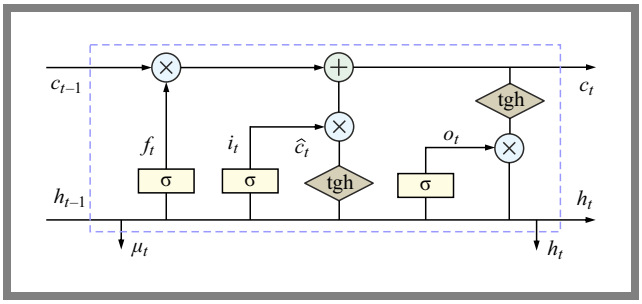

Fig. 3. LSTM cell.

$$\theta(t) = \left\{\theta_1(t), \theta_2(t), \ldots, \theta_Q(t)\right\} . \tag{24}$$

Items within function $\theta(t)$ are sorted to create a sorted index vector such as:

$$\lambda = \left\{\lambda_1, \lambda_2, \ldots, \lambda_Q\right\} . \tag{25}$$

This sorted index vector is then input into the Zipf model to determine the popularity of content in the repository, as described in the caching model, producing a popularity distribution:

$$\pi = \left\{\pi_1, \pi_2, \ldots, \pi_Q\right\} . \tag{26}$$

Our LSTM-based prediction model for predicting content request volumes is structured as follows:

- Input layer provides historical data of content request numbers, processed and normalized.
- First LSTM layer:
  - consists of 128 units (number of memory cells in the layer) with `return_sequences=true` to maintain temporal sequence processing for deeper layers.
  - uses the rectified linear unit (ReLU) activation function for intermediate layers to introduce non-linearity.
  - includes a dropout rate of 0.2 to mitigate overfitting by randomly omitting a subset of features during training.
- Second LSTM layer:
  - comprises 64 units configured with `return_sequences=true` and takes the sequence output from the previous LSTM layer and processes it further, outputting a sequence of 64-dimensional hidden states.
  - ReLU activation function.
  - this layer also includes a dropout of 0.2.
- Third LSTM layer:
  - 32 units.
  - ReLU activation function.
  - false return sequences (to return the final hidden state for the next layer).
  - 0.2 dropout.

  This layer takes the sequence output from the previous LSTM layer and processes it to output a 32-dimensional hidden state representing the entire sequence.
- Dense layer (output layer) features a single neuron with a linear activation function to predict the count of future content requests, reflecting the anticipated popularity.

The model compilation includes:

- Optimizer – the Adam optimizer is utilized for its efficient computation and adaptive learning rate of 0.001, enhancing the convergence of training.
- Loss function – mean squared error (MSE) is employed to quantify the accuracy of predictions, providing a clear measure of prediction error in the context of regression.

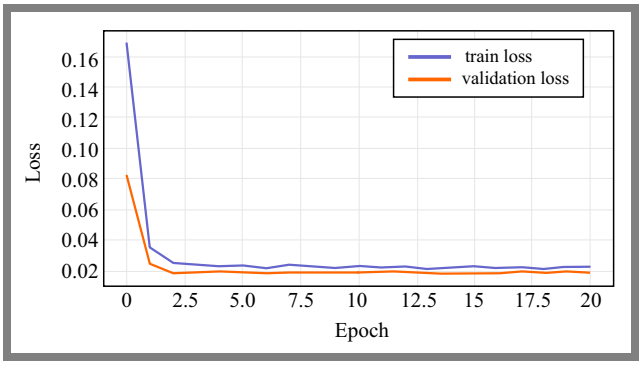

Fig. 4. Model loss over an epoch.

- Metrics – mean absolute error (MAE) provides a measure of the mean absolute difference between observed values and predictions, offering a more direct interpretation than MSE.

The training configuration is as follows:

- Epochs and batch size – the model is trained over 100 epochs with a batch size of 32, balancing the model's exposure to the training data with computational efficiency.
- Early stopping is implemented with a patience of 10 epochs; training is stopped early if the validation loss does not improve for 10 consecutive epochs, preventing overfitting.

For model training, 20% of the data is used for validating performance of the model during training, which facilitates tuning and prevents overfitting. Early stopping is used to halt the training process when the model's performance on the validation set stops improving.

5.3. Evaluation of Proposed Model

The dataset used for predicting content requests originates from the Big Data Challenge [26]. It includes data for ten types of content, collected between November 1, 2013 and January 1, 2014, with each data point representing a 10-minute interval. This analysis is based on a dataset of 1,000 such data points.

The first 800 data points are utilized as the training set, and the model's performance is subsequently tested on the last 200 data points. Once trained, the model's performance is evaluated on the validation set. Figure 4 illustrates the training and validation losses over epochs and reveals the following insights:

- Sharp decline in initial epochs. There is a significant drop in both training and validation losses within the first few epochs, indicating that the model quickly captures dominant patterns in the data.
- Convergence. After the initial sharp decline, both losses level off, showing only slight further reductions. This trend suggests that continuing training beyond these epochs might not yield substantial improvements, pointing towards potential diminishing returns.
- Closeness of losses. The training and validation losses remain close throughout the process, suggesting that the model generalizes well and is not overfitting. This closeness

is indicative of a model that performs well on both seen and unseen data.

This comprehensive evaluation highlights the LSTM-based content prediction model's ability to learn and predict content request volumes effectively, showcasing its utility in practical scenarios based on the dataset from the Big Data Challenge. The analysis of the loss plot further confirms the model's capability to generalize, making it a reliable tool for forecasting content requests in similar settings.

By integrating LSTM predictions into our caching strategy, we can dynamically adjust the cached content based on real-time popularity forecasts. This approach not only enhances the efficiency of the caching system but also ensures that the most relevant and demanded content is available to users promptly, thereby reducing latency and improving user experience.

This sophisticated LSTM framework forms the backbone of the proposed predictive model, enabling robust and accurate forecasting of content popularity, which is crucial for the effective management of cache resources in vehicular networks.

5.4. Cooperative Caching-based TS Algorithm

Reinforcement learning is an algorithmic approach centered on mapping behaviors based on the environmental state [27]. In this framework, the agent within the reinforcement learning system chooses and performs actions from a set based on the system's state, subsequently receiving a reward corresponding to the action taken. This reward serves as a metric for assessing the effectiveness of the chosen action. After the training phase, the agent can swiftly identify and perform actions that are linked to higher rewards, depending on the system's state.

In the context of MAB problems, each arm corresponds to a distinct action, yielding a reward upon selection. However, the likelihood of receiving a reward varies across arms, presenting the challenge of efficiently choosing arms within a limited number of trials to optimize rewards. Balancing exploration and exploitation is a widely recognized challenge in this scenario. To maximize rewards, decision-makers must strike a balance between exploring to leverage arms more effectively and exploiting the knowledge gained from prior exploration. The MAB problem, utilizing strategies like TS, aims to maximize the expected reward, ensuring the best possible outcome [28].

TS is more effective for caching issues with high uncertainty, as its exploration of varied strategies can enhance the overall performance. TS is a stochastic algorithm used for decision making in situations of uncertainty, commonly applied to MAB problems. It assigns scores to each arm by assuming the reward probability for each arm conforms to a beta distribution [29].

The latter is a continuous distribution of probabilities inside the interval $[0, 1]$, distinguished by two positive shape parameters denoted by a and b. The average value of the beta distribution is determined by $\frac{a}{a+b}$. A larger a results in a higher average value of $\mathrm{beta}(a, b)$, whereas a higher b decreases this average. The TS algorithm estimates the return, represented by ϕ, by sampling $\mathrm{beta}(a, b)$ [29].

Following the outcome of arm selections, updates are made to the selected arm. Receiving a reward of 1 increases the corresponding a value by 1. When the reward is 0, the b value is incremented by 1. The sampling's inherent randomness enables TS to naturally achieve a balance between exploration and exploitation (EE).

Additionally, adjustments to the EE balance can be made by altering the update mechanism for the TS parameters a and b. We designated the edge server, base station, cluster head, and each user with a local cache as agents. Each category is considered an arm. Each arm has a unique likelihood of being chosen. Once the file category is identified based on probability, the file is selected from that category to cache.

In this study, the TS technique is used to iteratively select actions (arms) based on their predicted reward probability. The algorithm tracks reward distribution estimations for each arm using beta distributions. The algorithm samples beta distributions at each iteration to evaluate the likelihood of each arm being the best selection. Next, it chooses the arm with the highest calculated likelihood and adjusts its settings according to the incentives it has received. In this context, rewards refer to cache hit rates.

The reward is an indicator of success linked to choosing a specific arm. The reward is calculated according to the hit rate attained by completing the required tasks. The hit rate indicates the ratio of successful task completions to the total number of requested tasks. To update the parameters of the beta distribution at time-slot $t+1$, the following formula was used:

$$\begin{aligned} a_{d,cat_i}^{t+1} &= a_{d,cat_i}^{t} + R_{d,cat_i}^{t} \\ b_{d,cat_i}^{t+1} &= b_{d,cat_i}^{t} + (1 - R_{d,cat_i}^{t}) \end{aligned}, \tag{27}$$

where R_{d,cat_i}^{t} is the hit rate (reward) of the selected arm of the file in the cache of each device d at time t.

Algorithm 2 illustrates the proposed TS-MMCM approach, while the TS algorithm process is outlined in Algorithm 3. The TS-based algorithm sorts content into numerous categories. Without categorizing content, each agent has an action space of around $(2^{|L|})$ when selecting from $|L|$ files. Therefore, the algorithm becomes ineffective. The TS-based approach is suggested to decrease the action space of Q-learning [30]. Initially, the algorithm initializes the beta distribution of each arm as follows:

$$\phi_{cat_i}^{0} \sim \mathrm{beta}(a_{cat_i}^{0}, b_{cat_i}^{0}) \leftarrow (1, 1), \text{ where } 1 \leqslant i \leqslant N\ , \tag{28}$$

where N is the number of arms.

Therefore, the arm with the highest sampled value is selected as the optimal arm. According to the stored probabilities of each arm, content will be cached based on its probability of being requested in the future. Once the items are stored in the caches of each device, the rewards are acquired. The posterior distribution of each selected arm is updated as follows:

$$\phi_{cat_i}^{t+1} \sim \mathrm{beta}(a_{cat_i}^{t+1}, b_{cat_i}^{t+1}), \quad \text{where } 1 \leqslant i \leqslant N\ . \tag{29}$$

Algorithm 2 TS-based mobility-aware multi-hierarchical caching model with vehicle clustering and content popularity prediction methods (TS-MMCM).

Phase 1. Initialization
Initialize parameters: number of vehicles, vehicle clusters, period T, content categories $|cat|$, repository, content sizes, parameter s, cache capacity
Initialize beta distribution parameters:
$\phi^0_{cat_i} \sim \text{beta}(a^0_{cat_i}, b^0_{cat_i}) \leftarrow (1,1), 1 \leqslant i \leqslant N$
Phase 2. Vehicle clustering
Perform Algorithm 1 to obtain vehicle clusters
Phase 3. Obtain cluster heads
Choose cluster heads based on CL selection criterion:
$\overline{\text{CL}}_{i,l} = \frac{1}{n_i - 1} \sum_{l'=1}^{n_i} \text{CL}^{i,l'}_{i,l}, \quad l' \neq l$
Phase 4. Find the best caching decision
repeat for each round
    Obtain the predicted number of content requests $\theta(t)$ using LSTM
    Sort $\theta(t)$ to obtain sorted index vector λ_Q
    Obtain content popularity vector π_Q using Zipf model: $P(R,s,N) = \frac{1/R^s}{\sum_{q=1}^{N} \frac{1}{q^s}}, q \in \{1,2,\ldots,l\}$
    for each cluster H_i **do**
        for each vehicle user $H_{i,j} \in H_i$ **do**
            Determine caching method in each local cache using a TS-based algorithm
        end for
        Determine the caching policy in each cluster head $H_{i,1}$ using a TS-based algorithm
    end for
    Determine caching policy in the edge server using a TS-based algorithm
    Update content in each cache
until convergence criteria are met or fixed number of rounds T

Algorithm 3 Thompson sampling.

Input: Initialize parameters
Output: Report the selected arms and their corresponding probabilities
Number of categories N
Initialize the posterior distributions of each arm as a function given by Eq. (28)
Initialize variable for reward
repeat
    for each round **do**
    Pull arms:
        Sample from the posterior distributions of each arm
        Select the arm with the highest sampled value
    Obtain reward:
        Execute the chosen action/arm in the environment
        Update the cache
        Observe the reward obtained
    Update parameters for the chosen arm:
        Update the posterior distribution of the selected arm based on observed reward according to Eq. (29)
        Update the reward for the chosen arm
    end for
Next round
until convergence or fixed number of rounds

6. Simulation Results

In this section, we validate the effectiveness and efficiency of proposed approach by conducting simulations using Python. The performance metrics we considered are the hit rate and latency. The hit rate refers to the proportion of data or resource requests that are successfully retrieved from the cache, while the latency is the time duration from when a request is sent by the vehicle user to when the last data packet is received. Table 2 lists the simulation's parameters and their values.

In the simulation scenario, a library containing 1000 items was used, each having a size ranging from 5 to 100 MB. The total size of all pieces of content was 51 GB. The cache capacity of the edge server was configured to be between 10% and 25% of the entire volume of content. The cache capacity allocated to each user and cluster head was adjusted to be between 10% and 25% of the cache size of the edge server, ensuring a fair and efficient distribution of caching resources across the network [30]. We modelled the arrival of requests from cars as a Poisson process in each time slot. To simulate user requests, we varied the shape parameter values of the Zipf distribution.

6.1. Exploring Caching Strategy Performance

The proposed strategy is compared with the following approaches: LRU, LFU, fuzzy logic, and ICSAD [30].

Figure 5 depicts the performance of the preceding algorithms in relation to Zipf parameters. Figure 5a shows the average hit rate as a function of the Zipf parameter for many caching or selection algorithms. The proposed algorithm (TS-MMCM) showed superior performance compared to all other algorithms, demonstrating continuous and significant improvement as the Zipf value grows from 1.0 to 1.6. This demonstrates that our algorithm efficiently gives priority to caching the most widely accessed content.

Its average hit rate grows from 0.3 to above 0.6, indicating a strong ability to adjust to the Zipf parameter-defined popularity distribution. This demonstrates that the proposed approach is more efficient in obtaining a greater hit rate when

Tab. 2. Simulation parameters.

Parameter	Value
Distance	2 [km]
Vehicle speed	20, 60 [km/h]
Request rate	20, 30 [requests/m]
Number of vehicle users	55
Number of file categories	5
Cache capacity rate	10, 15, 20, 25 [%]

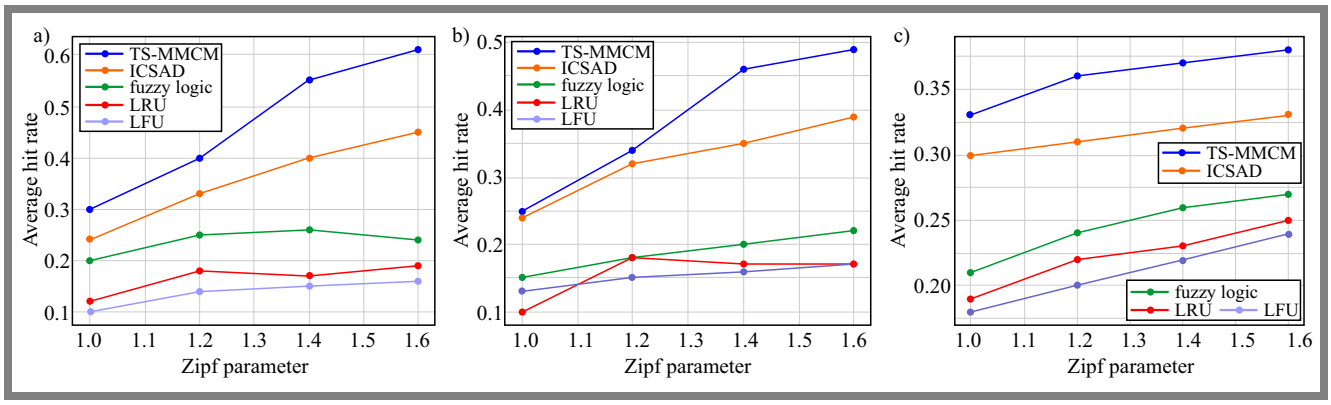

Fig. 5. Comparative analysis of caching strategies across different cache locations: a) for locale cache, b) for cluster head cache, and c) for edge cache.

the distribution of requests becomes more imbalanced (which is indicated by an increase in the Zipf parameter).

ICSAD is the algorithm that offers second best performance, showing a moderate rising trend. The average hit rate started at approximately 0.25 when the Zipf parameter was set to 1.0 and gradually increased to just below 0.5 when the Zipf parameter was set to 1.6. This hit rate is consistently lower than that of the TS-MMCM algorithm in all cases. Fuzzy logic showed a rather flat performance with a slight upward trend. It started at an average hit rate of approx. 0.2 and ended at approx. 0.25. Although performance improved as the Zipf parameter increased, the gains are not as great as those observed when using our algorithm or ICSAD.

LRU, the traditional caching strategy, followed a slight upward trend, starting just above a hit rate of 0.1 and approaching 0.2 as the Zipf parameter increased to 1.6. This suggests its modest sensitivity to popularity distribution.

The LFU algorithm exhibited the least amount of improvement, indicating it may be least suitable for adapting to changes in content popularity distributions. Figure 5b illustrates the effectiveness of different caching algorithms in relation to the Zipf parameter within the cluster head cache. Our approach routinely surpasses ICSAD, fuzzy logic, LRU, and LFU approaches in terms of performance, with a noticeable positive relationship between the average hit rate and the Zipf parameter. The initial hit rate of the proposed algorithm equaled approx. 0.25 and it gradually increased to nearly 0.5.

In contrast, the remaining algorithms demonstrated different levels of improvement as the Zipf parameter increased, but none of them surpassed a hit rate of 0.4. The figure indicates that TS-MMCM easily adjusts to various data request distributions, as evidenced by the Zipf parameter. Figure 5c shows a comparison of the performance of hit rate-based algorithms in the edge cache as the Zipf parameter increased from 1.0 to 1.6.

Performance of the proposed algorithm surpassed that of the others, demonstrating a consistent rise in the hit rate as the Zipf parameter increased. The ICSAD algorithm is ranked second, with the fuzzy logic, LRU, and LFU algorithms following in that order. The disparity in performance among algorithms indicates different degrees of efficiency, with the conventional LRU and LFU caching techniques falling behind more recent approaches. The figure likely illustrates the superiority of a new algorithm in the context of data caching or retrieval, driven by the distribution of data, as indicated by Zipf.

Figure 6 shows the relationships between the cache hit ratio of the aforementioned algorithms and the overall caching capacity rate. The algorithms are compared at capacity rates of 0.1, 0.15, 0.2, and 0.25. These capacity rates indicate the combined cache capacity rate of the local cache, cluster head, and edge server. The cache hit ratio is the ratio of requests successfully retrieved from the cache to the total number of requests made.

The proposed algorithm offers exceptional performance, since it consistently achieves the greatest hit rates across all capacity rates. The frequency of successful cache retrievals improved as the capacity rate increased, indicating that it efficiently utilized the available cache space to store frequently requested content.

The ICSAD algorithm demonstrates a correlation between the hit rate and the capacity rate, but it fell short of achieving the same degree of performance as TS-MMCM, albeit it ranked as the second most effective method. Performance of the fuzzy logic algorithm was inferior to that of the proposed algorithm and ICSAD at the 0.1 capacity rate, but increased at the 0.15 capacity rate. However, the improvement is not as great at

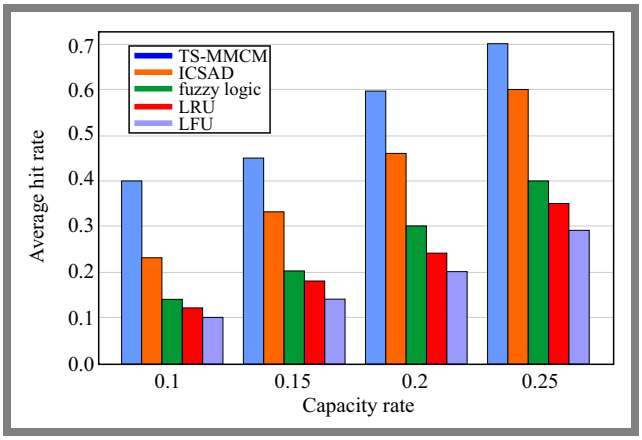

Fig. 6. Cache hit ratio versus total caching capacity rate.

higher capacity rates when compared with TS-MMCM and ICSAD.

The LRU method constantly exhibited worse performance when compared to TS-MMCM, ICSAD, and fuzzy logic. However, it surpassed the LFU algorithm. The hit rate increased when the capacity rate escalated, although it stayed below 0.4 even at the greatest capacity.

The LFU algorithm initially had the lowest capacity rate of 0.1, but it showed a progressive rise in performance as the capacity rate increased. However, it consistently remained the least-performing algorithm compared to other approaches at all capacity rates.

As the capacity rate increased, the average hit rate for each algorithm also increased. This is an expected behavior as greater capacity generally means that more data can be stored, which is likely to translate into higher success rates.

This is because more content may be cached with a bigger caching capacity rate, allowing the local cache, cluster head, and edge server to respond to more requests from the vehicle users. The TS-MMCM algorithm and ICSAD had the highest average hit rates at a capacity rate of 0.25, suggesting that they are more effective at making use of bigger capacity rates than fuzzy logic, LRU, and LFU.

In general, Fig. 6 shows that the TS-MMCM algorithm and ICSAD are the most effective solutions in terms of hit rate at the capacity rates tested, with the proposed algorithm outperforming ICSAD. Meanwhile, LFU seems to be the least effective across all capacity rates.

6.2. Cache Hit Rate and Latency Optimization via Advanced Clustering Technique

To improve the accuracy of content popularity determination and to increase the cache hit rate even further, this paper described a procedure in which the LSTM method is used to predict content requests. We evaluated the effectiveness of vehicle grouping according to speed and position. Vehicles were grouped into clusters of three, six, and nine, and their performance was analyzed based on the request rate, using a Zipf parameter of 1.0 and setting the total cache capacity rate at 0.15.

Figure 7 illustrates the correlation between the cache hit rate in a cooperative caching system and the rate at which requests or data requests are made. The figure consists of three lines, each indicating a distinct number of clusters. The *y*-axis displays the hit rate, which is the fraction of requests that the cache serves without having to retrieve data from the origin server. A higher hit rate signifies superior performance, as it implies a larger number of requests being promptly served from the cache. The *x*-axis represents the request rate, which is assumed to be a measure of how often requests are submitted to the cache. The rate is often quantified as the number of requests per unit of time.

Cluster 3 exhibited the lowest hit rate among all request rates, initially equaling 0.82 and gradually increasing to 0.92 as the request rate rose from 20 to 30. These findings indicate that cluster 3 is the least effective among the three in handling

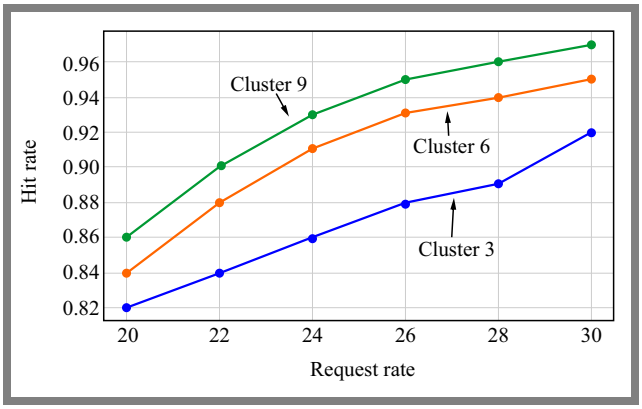

Fig. 7. Cache hit ratio vs. request rate.

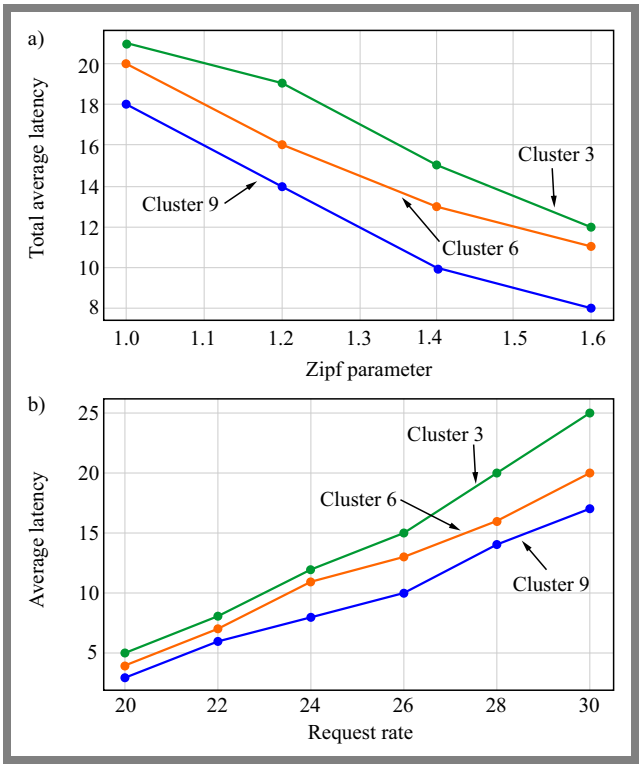

Fig. 8. Vehicle clustering performance: a) total average latency vs. Zipf parameter and b) average latency vs. request rate.

cache requests, although it demonstrated progress as the rate of requests increased. The hit rate of cluster 6 initially stood at 0.84 and gradually rose to just above 0.94 when the request rate escalated from 20 to 30. It surpassed cluster 3 in terms of performance at all request rate levels.

Cluster 9 exhibited the highest hit rate, starting at 0.86 and peaking at nearly 0.97 during the highest request rate shown. It offered superior performance compared to the two other clusters.

All clusters experienced an increase in their hit rates as the request rate rose. This indicates that the cooperative caching mechanism is efficient in using the higher volume of requests to accurately anticipate and cache popular content. Increased request rates can enhance the precision of predicting popular items, allowing for prefetching or caching and thus resulting in an improved hit rate.

Overall, Fig. 7 demonstrates that in a cooperative caching framework, an increase in the request rate leads to a corresponding rise in the hit rate across all clusters. Cluster 9 exhibited the highest performance, while cluster 3 showed the lowest performance.

Figure 8 shows the total average latency based on the Zipf parameter, as well as the average latency based on the request rate for three cluster sizes.

It follows from Figure 8a that the total average latency decreases as the number of clusters increases, indicating that vehicle clustering may efficiently reduce this time, while Fig. 8b demonstrates how the proposed clustering architecture might lead to a significant decrease in latency. Increasing the number of clusters can further reduce average latency.

Simulation results show that TS-MMCM algorithm significantly improves IoV caching system's performance compared to that achieved with the use of previous algorithms, enabling efficient adjustment to rapidly evolving network structures in time-sensitive IoV environments.

7. Conclusions

This study shows that intentional vehicle clustering combined with machine learning to forecast content popularity enhances both system performance and user experience by enabling better caching decisions and fostering more dependable communication links within vehicular networks. The goal of the proposed future study is to apply sophisticated machine learning models, energy efficiency considerations, and multidisciplinary research to ensure the model's practical feasibility in real-world scenarios while simultaneously enhancing the caching model's scalability, efficiency, and security.

References

[1] B. Radouane, G. Lyamine, K. Ahmed, and B. Kamel, "Scalable Mobile Computing: From Cloud Computing to Mobile Edge Computing", *2022 5th International Conference on Networking, Information Systems and Security: Envisage Intelligent Systems in 5G/6G-based Interconnected Digital Worlds (NISS)*, Bandung, Indonesia, 2022 (https://doi.org/10.1109/NISS55057.2022.10085600).

[2] W. Jiang *et al.*, "Multi-agent Reinforcement Learning for Efficient Content Caching in Mobile D2D Networks", *IEEE Transactions on Wireless Communications*, vol. 18, no. 3, pp. 1610–1622, 2019 (https://doi.org/10.1109/TWC.2019.2894403).

[3] Z. Yang, Y. Liu, Y. Chen, and L. Jiao, "Learning Automata Based Q-learning for Content Placement in Cooperative Caching", *IEEE Transactions on Communications*, vol. 68, no. 6, pp. 3667–3680, 2020 (https://doi.org/10.1109/TCOMM.2020.2982136).

[4] X. Fang, T. Zhang, Y. Liu, and Z. Zeng, "Multi-agent Cooperative Alternating Q-learning Caching in D2D-enabled Cellular Networks", *IEEE Global Communication Conference (GLOBECOM)*, Waikoloa, USA, 2019 (https://doi.org/10.1109/GLOBECOM38437.2019.9014053).

[5] Y. Qian *et al.*, "Reinforcement Learning-based Optimal Computing and Caching in Mobile Edge Network", *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 10, pp. 2343–2355, 2020 (https://doi.org/10.1109/JSAC.2020.3000396).

[6] P. Liu, Y. Zhang, T. Fu, and J. Hu, "Intelligent Mobile Edge Caching for Popular Contents in Vehicular Cloud Toward 6G", *IEEE Transactions on Vehicular Technology*, vol. 70, no. 6, pp. 5265–5274, 2021 (https://doi.org/10.1109/TVT.2021.3076304).

[7] W. Qi *et al.*, "Extensive Edge Intelligence for Future Vehicular Networks in 6G", *IEEE Wireless Communications*, vol. 28, no. 4, pp. 128–135, 2021 (https://doi.org/10.1109/MWC.001.2000393).

[8] J. Shi *et al.*, "A Novel Deep Q-learning-based Air-assisted Vehicular Caching Scheme for Safe Autonomous Driving", *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4348–4358, 2021 (https://doi.org/10.1109/TITS.2020.3018720).

[9] Z. Zhang, C.-H. Lung, M. St-Hilaire, and I. Lambadaris, "Smart Proactive Caching: Empower the Video Delivery for Autonomous Vehicles in ICN-based Networks", *IEEE Transactions on Vehicular Technology*, vol. 69, no. 7, pp. 7955–7965, 2020 (https://doi.org/10.1109/TVT.2020.2994181).

[10] Y. Liu and B. Mao, "On a Novel Content Edge Caching Approach Based on Multi-Agent Federated Reinforcement Learning in Internet of Vehicles", *2023 32nd Wireless and Optical Communications Conference (WOCC)*, Newark, USA, 2023 (https://doi.org/10.1109/WOCC58016.2023.10139417).

[11] A. Ndikumana *et al.*, "Deep Learning Based Caching for Self-driving Cars in Multi-access Edge Computing", *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 5, pp. 2862–2877, 2021 (https://doi.org/10.1109/TITS.2020.2976572).

[12] Z. Zhu *et al.*, "Proactive Caching in Auto Driving Scene via Deep Reinforcement Learning", *2019 11th International Conference on Wireless Communications and Signal Processing (WCSP)*, Xi'an, China, 2019 (https://doi.org/10.1109/WCSP.2019.8928131).

[13] A. Ndikumana and C.S. Hong, "Self-driving Car Meets Multi-access Edge Computing for Deep Learning-based Caching", *2019 International Conference on Information Networking (ICOIN)*, Kuala Lumpur, Malaysia, 2019 (https://doi.org/10.1109/ICOIN.2019.8718113).

[14] W. Jiang, G. Feng, S. Qin, and Y.-C. Liang, "Learning-based Cooperative Content Caching Policy for Mobile Edge Computing", *ICC 2019 – 2019 IEEE International Conference on Communications (ICC)*, Shanghai, China, 2019 (https://doi.org/10.1109/ICC.2019.8761121).

[15] C. Zhang *et al.*, "Toward Edge-assisted Video Content Intelligent Caching with Long Short-term Memory Learning", *IEEE Access*, vol. 7, pp. 152832–152846, 2019 (https://doi.org/10.1109/ACCESS.2019.2947067).

[16] Y. Ye, M. Xiao, and M. Skoglund, "Mobility-aware Content Preference Learning in Decentralized Caching Networks", *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 1, pp. 62–73, 2020 (https://doi.org/10.1109/TCCN.2019.2937519).

[17] Y. Zhang *et al.*, "Cooperative Edge Caching: A Multi-agent Deep Learning Based Approach", *IEEE Access*, vol. 8, pp. 133212–133224, 2020 (https://doi.org/10.1109/ACCESS.2020.3010329).

[18] X. Xu, M. Tao, and C. Shen, "Collaborative Multi-agent Multi-armed Bandit Learning for Small-cell Caching", *IEEE Transactions on Wireless Communications*, vol. 19, no. 4, pp. 2570–2585, 2020 (https://doi.org/10.1109/TWC.2020.2966599).

[19] R. Wang *et al.*, "Cooperative Caching Strategy with Content Request Prediction in Internet of Vehicles", *IEEE Internet of Things Journal*, vol. 8, no. 11, pp. 8964–8975, 2021 (https://doi.org/10.1109/JIOT.2021.3056084).

[20] X. Bi and L. Zhao, "Collaborative Caching Strategy for RL-based Content Downloading Algorithm in Clustered Vehicular Networks", *IEEE Internet of Things Journal*, vol. 10, no. 11, pp. 9585–9596, 2023 (https://doi.org/10.1109/JIOT.2023.3235661).

[21] T. Kanungo *et al.*, "An Efficient K-means Clustering Algorithm: Analysis and Implementation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 881–892, 2002 (https://doi.org/10.1109/TPAMI.2002.1017616).

[22] D. Zhang *et al.*, "New Multi-hop Clustering Algorithm for Vehicular Ad Hoc Networks", *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 4, pp. 1517–1530, 2019 (https://doi.org/10.1109/TITS.2018.2853165).

[23] C. Zhang *et al.*, "Deep Transfer Learning for Intelligent Cellular Traffic Prediction Based on Cross-domain Big Data", *IEEE Journal*

on Selected Areas in Communications, vol. 37, no. 6, pp. 1389–1401, 2019 (https://doi.org/10.1109/JSAC.2019.2904363).

[24] C. Olah, *Understanding LSTM Networks*, Github, 2015 [Online] Available: http://colah.github.io/posts/2015-08-Understanding-LSTMs/.

[25] C. Wang, Z. Zhao, Q. Sun, and H. Zhang, "Deep Learning-based Intelligent Dual Connectivity for Mobility Management in Dense Network", *2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)*, Chicago, USA, 2018 (https://doi.org/10.1109/VTC-Fall.2018.8690554).

[26] G. Barlacchi *et al.*, "Multi-source Dataset of Urban Life in the City of Milan and the Province of Trentino", *Scientific Data*, no. 2, art. no. 150055, 2015 (https://doi.org/10.1038/sdata.2015.55).

[27] Y. Cui, Y. Liang, and R. Wang, "Resource Allocation Algorithm with Multi-platform Intelligent Offloading in D2D-enabled Vehicular Networks", *IEEE Access*, vol. 7, pp. 21246–21253, 2019 (https://doi.org/10.1109/ACCESS.2018.2882000).

[28] S. Agrawal and N. Goyal, "Thompson Sampling for Contextual Bandits with Linear Payoffs", *ArXiv*, 2012 (https://doi.org/10.48550/arXiv.1209.3352).

[29] Y. Chen, J. Liu, Y. Shi, and M. Sheng, "Prefetch and Cache Replacement Based on Thompson Sampling for Satellite IoT Network", *ICC 2021 – IEEE International Conference on Communications*, Montreal, Canada, 2021 (https://doi.org/10.1109/ICC42927.2021.9500508).

[30] L. Zhao *et al.*, "Intelligent Content Caching Strategy in Autonomous Driving Toward 6G", *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 9786–9796, 2022 (https://doi.org/10.1109/TITS.2021.3114199).

Radouane Baghiani, Ph.D. student
Computer Sciences and Information Technologies Department
https://orcid.org/0009-0001-7694-7204
E-mail: baghiani.radouane@univ-ouargla.dz
LINATI Laboratory, Kasdi Merbah University, Ouargla, Algeria
https://www.univ-ouargla.dz

Lyamine Guezouli, Ph.D.
Renewable Energies and New Technologies Department
https://orcid.org/0000-0002-7406-7633
E-mail: lyamine.guezouli@hns-re2sd.dz
LEREESI Laboratory, HNS-RE2SD, Batna, Algeria
http://www.hns-re2sd.dz

Ahmed Korichi, Professor
Computer Sciences and Information Technologies Department
https://orcid.org/0000-0002-5741-4963
E-mail: ahmed.korichi@univ-ouargla.dz
LINATI Laboratory, Kasdi Merbah University, Ouargla, Algeria
https://www.univ-ouargla.dz

A Collaborative Approach to Detecting DDoS Attacks in SDN Using Entropy and Deep Learning

Narayan D.G., Heena W, and Amit K

KLE Technological University, Hubballi, Karnataka, India

https://doi.org/10.26636/jtit.2024.3.1609

Abstract — Software-defined networking (SDN) is an approach to network management allowing to enhance the performance of the network and making it more flexible. The centralized architecture of SDN makes it vulnerable to cyberattacks, especially distributed denial of service (DDoS) attacks. Existing research investigates the detection of DDoS attacks separately on the control plane and data plane. However, there is a need for efficient and accurate detection of these attacks using features obtained from both control and data planes. Therefore, we present a mechanism for identifying DDoS attacks using entropy, multiple feature selection mechanisms, and deep learning. Initially, we use entropy on the control plane to detect anomalous activity and identify suspicious switches. Next, we capture traffic on the suspicious switches to detect DDoS attacks. To detect these attacks, we utilize multi-layer perceptron (MLP) deep learning models, convolutional neural network (CNN), and the long short-term memory (LSTM) approach. An InSDN dataset is used to train the model and test data are generated using Mininet emulation and the Ryu controller. The results reveal that LSTM outperforms MLP and CNN, achieving an accuracy of 99.83%.

Keywords — DDoS, deep learning, LSTM, machine learning, random forest, SDN

1. Introduction

Software-defined networking (SDN) is a popular technology providing a centralized and efficient network infrastructure. In contrast to traditional distributed and composite network designs, SDN separates the network traffic forwarding process (infrastructure layer) from the control process (control plane). This separation offers numerous advantages, such as programmable configuration capabilities. By leveraging open APIs within software applications, SDN allows network behavior to be programmed centrally across the entire network and its constituent parts. However, SDN is also associated with some drawbacks stemming from its centralized architecture, i.e. a single point of failure. In the event of an intrusion, this can lead to disruptions throughout the entire network.

Distributed denial of service (DDoS) attacks pose a significant risk to SDN security. The primary objective of these attacks is to prevent authorized users from accessing resources. Additionally, these attacks consume the targeted source's capacity, leading to delays and disruptions in normal operations. While several studies have explored the use of classification algorithms to identify and prevent DDoS attacks, current research faces such challenges as achieving realistic performance rates for the detection system, addressing detection delays, and effectively handling high volumes of network traffic [1]. The absence of a comprehensive real-world solution to detect and prevent different types of attacks has motivated the development of a dedicated DDoS attack detection system.

Recently, AI-driven methodologies for identifying DDoS attacks have been utilized [2] and most of these research projects rely on machine/deep learning approaches to detect DDoS attacks [3]. While some studies in the literature have focused on addressing DDoS attacks by employing entropy or unsupervised learning, they have turned out to be less accurate [4]. Hence, there exists a necessity to integrate an entropy-based approach with machine/deep learning approaches to augment detection precision.

In this work, we have introduced a mechanism that employs an entropy-based approach to identify suspicious switches in the SDN control plane. We utilize a multi-feature selection mechanism with a deep learning-based method on the data plane to identify attacker machines. This deep learning module captures packets from potentially compromised switches over a predetermined time frame and employs a trained model to identify and classify potential attacks. Through the integration of both entropy and deep learning modules, the proposed approach is capable of effectively detecting unauthorized attacks and provides robust defense mechanisms against potential unknown threats targeting the application layer.

This study offers the following contributions:

- an entropy-based method is proposed that serves as a primary detection method on the control plane. This module identifies OpenFlow switches that are potentially under attack,
- machine and deep learning modules are developed that capture traffic on the susceptible switches to identify a DDoS attack. A multi-feature selection mechanism using SelectKBest, ANOVA F-value scores and feature importance scores from the random forest classifier is proposed,

This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) License.
For more information, see https://creativecommons.org/licenses/by/4.0/

- a comprehensive performance evaluation of the proposed method is performed, including an assessment of the accuracy and efficiency of the detection techniques.

The remaining part of the article is divided into four sections. Section 2 reviews related research, while Section 3 outlines the proposed model and algorithms. The results are presented in Section 4, and conclusions are drawn in Section 5.

2. Related Work

In [5], the authors developed a real-time DDoS attack detection mechanism using a deep neural network solution and the CICIDS 2017 dataset. The results show that the proposed model performs better than other models with different datasets, achieving an accuracy of 97.59% while consuming less time and resources. The authors of [6] initially set up a DDoS attack detection mechanism on the data plane. Then, to enhance the precision of detection, k-nearest neighbors and k-means machine learning algorithms were employed. These algorithms utilize a combination of five features to identify vulnerable flows in the network. The results indicate that the proposed method achieves a superior accuracy rate of 97.53%, outperforming both entropy and self-organizing map techniques.

Some of the work is based on statistical techniques. In [7], a joint entropy-based DDoS defense scheme is proposed. The authors use joint entropy measurements from multiple network features to detect anomalies. These features include, inter alia, packet rate, flow count, and byte count. By calculating the entropy of these features and analyzing their joint distributions, deviations indicative of DDoS attacks are identified. Paper [8] introduces an entropy-based detection method that analyzes various network traffic metrics to detect DDoS attacks. This approach aims to address challenges posed by the dynamic and potentially large-scale nature of DoS attacks in SDN networks. Through experimental validation, the paper demonstrates the effectiveness of such a framework in accurately identifying DDoS attacks.

In [9], the authors proposed a DDoS attack detection method using a combination of entropy and ensemble learning techniques. An inspection module deployed on the data plane collects real-time packets from the network through the edge switch. When vulnerable packets are detected, the controller is immediately notified. To achieve more accurate detection, a detection module is established on the control plane. This module incorporates five features and utilizes the random forest algorithm to classify network traffic. Once an attack packet is confirmed, the controller takes immediate action by dropping the packets from the edge switch, thereby blocking the attack. The simulation results, focusing on ICMP and SYN flood attacks, demonstrate a quick response and a more efficient detection rate.

In article [10], a cross-plane DDoS mitigation approach is introduced relying on two events to detect and react to DDoS attacks. Firstly, a data plane detection module is established, including a flow-monitoring algorithm. Secondly, a machine learning-based detection module is implemented on the control plane to ensure more precise attack detection. The results demonstrate that the system achieves a higher accuracy rate of 96%. In [11], a hybrid approach is proposed that combines features from both control and data planes. An entropy module is utilized as the primary detection feature, while a machine learning module is employed for more accurate attack detection. The results reveal that the integration of data and control plane features leads to better accuracy with reduced overhead.

Another framework, as proposed in [12], focuses on the control plane. The work emphasizes the security of both data and control planes. A deep neural network is used for packet classification. Preliminary detection occurs on the data plane, and attack traffic is mitigated. To detect attacks at an earlier stage, a detection module is set up on the control plane, providing better service to users using a priority queue. In article [13], the authors developed a safeguard mechanism to protect the control plane from DDoS attacks in software-defined networks. The safeguard framework is deployed across multiple control planes. The detection module works on the data plane.

In [14], the authors proposed a DDoS attack detection method using new features. On the control plane, flow messages are collected from the data plane, which is followed by extraction of important features. Next, the SVM model is trained using a dedicated dataset and is tested with real-time data. The results reveal that the system achieves a better accuracy rate compared to other existing solutions.

In paper [15], a method for detecting DDoS attacks by combining general entropy and the neural network particle swarm optimization with personal best (PSO-BP) is introduced. The entropy module is positioned on the OpenFlow switch to detect traffic, and the detection result is classified into normal and anomalous. To detect an attack, the controller uses the neural network PSO-BP. Results reveal that the DDoS attack detection rate is accurate and imposes a low CPU load on the controller. In [16], the authors proposed a mechanism for detecting and mitigating DDoS attacks using the data layer. The results show that the time to detection and mitigation is between 100 and 150 s.

The authors of [17] have developed a method to mitigate DDoS attacks originating from compromised OpenFlow switches on the controller. The research shows how a DDoS attack can be launched on an SDN controller with cooperating switches. It also directly mitigates such an attack within the second recurring request. In [18], the authors have proposed security mechanisms against attacks based on entropy in SDN-cloud using a POX controller. The developed mechanism has three advantages: high detection rate, low false-positive rate, and the ability to mitigate the attack. The proposed framework has the highest accuracy of 98.2%.

In [19], the authors introduced a framework to detect DDoS attacks with the cross-plane. The primary detection mechanism is set up at the data plane, and for precise DDoS attack detection the mechanism is set up at the control plane. The result reveals that the mentioned system is efficient and reduces

detection delay with less communication over the heading rate at the southbound.

Article [20] presents a mechanism on the control plane for a hybrid classification model, which utilizes SVM and SVM-SOM (self-organized map) machine learning algorithms for packet classification. The machine-learning module responsible for classification is implemented on the control plane. The results indicate that SVM-SOM achieves higher accuracy compared to SVM and SOM, with a lower false-positive rate than other algorithms. In paper [21], an efficient framework for an intrusion detection system on the data plane using data plane development kit (DPDK) is used to process packets at high speeds and monitor the data plane. The results demonstrate that the proposed DDoS attack detection framework effectively addresses high-speed networks in intrusion detection systems.

An automated system using a hybrid machine-learning algorithm for traffic packet classification is presented in [22]. The results show that a hybrid model combining SVM with random forest classification achieves the highest accuracy of 98.8%, with a low false-positive rate. The authors of [23] focus on the detection and mitigation of DDoS attacks using deep learning. CNN and recurrent neural network (RNN) long short-term memory models are employed for classification. RNN and LSTM perform well in detecting and mitigating DDoS attacks, achieving an accuracy of 89.63%. Performance of deep learning models is better when using a split ratio of 7:3 compared to 8:2 or 6:4. In [24], a big data framework for large-scale SDN that combines OpenStack with SDN using the ONOS controller, Apache Spark, and Apache Kafka is researched. The work overcomes the limitations of traditional data processing and validates the strength, scalability, and efficiency of the proposed framework.

The framework from [25] mitigates DDoS attacks on the data plane. The detection method analyzes header fields of the TCP connection and hashes them. Upon detecting an attack, the packets are dropped. In [26], the authors develop a mechanism using deep learning. They employ a standard dataset for traffic classification after preprocessing. The results show that the stacked auto-encoder multi-layer perceptron achieves an accuracy rate of 99.75%. The aim of the paper is to classify traffic using deep learning techniques.

The majority of the work mentioned above was carried using Internet datasets, such as CICIDS and NSL-KDD. Furthermore, most of the authors did not focus on feature selection techniques. We address these research gaps by employing a hybrid approach incorporating the inSDN dataset [27] and multi-feature selection methods.

3. Proposed Methodology

This section outlines the design of the proposed system, including the entropy module, the multi-feature selection mechanism, the InSDN dataset utilized for model training, and the deep learning algorithms employed for classifying network traffic in a test environment, using Mininet emulation.

3.1. System Model

As shown in Fig. 1, the process of detecting DDoS attacks in SDN environments involves a two-layered methodology combining entropy-based anomaly detection and the deep learning technique. The approach begins by leveraging entropy thresholds to identify anomalous network traffic behaviors. Entropy, in this context, measures the unpredictability or randomness of source IP addresses within a defined packet window (e.g. 50 packets). When entropy drops below a specified threshold during this window, it means that potentially suspicious activity has been detected. We monitor the drop in entropy for 10 consecutive packet windows (entropy drop counter), providing a cumulative measure of suspicious behavior.

Upon detecting sustained drops in entropy indicative of potential DDoS activity (e.g. when entropy drop counter reaches a predefined threshold, such as 10, i.e. an anomaly threshold), the next phase involves identifying the switches associated with the suspicious IP addresses using SDN flow rules. These switches are critical nodes within the network infrastructure where abnormal traffic patterns are observed.

To further validate and enhance the detection process, a deep learning module is employed. This module operates by capturing real-time network traffic data from identified switches for fixed time intervals (3 s) and then applies a deep learning algorithm to classify traffic patterns into those that are normal and malicious. The deep learning model is trained using the InSDN dataset to recognize the signature characteristics of DDoS attacks, enabling it to accurately distinguish between legitimate and malicious traffic.

The iterative, entropy-based anomaly detection process combined with deep learning ensures a proactive approach to detecting and mitigating DDoS attacks in SDN environments.

3.2. Entropy Computation

In an OpenFlow-enabled SDN network, when a TCP or UDP packet arrives at a switch, the switch checks its flow table for an existing rule. If no rule is found, the switch encapsulates the packet in a “packet_in” message and sends it to the SDN controller over a TCP connection. The controller analyzes the packet_in message, determines the optimal forwarding path, and sends back a “flow_mod” message with the new rule to the switch. The switch then forwards the packet according to the new rule and uses this rule to handle future packets.

We then use the packet_in to compute entropy and detect the suspicious host machines. In the SDN controller, entropy for packet_in messages using the source IP is calculated. This measures the randomness or unpredictability of the source IP addresses from a dataset of 50 packets. Entropy is a statistical measure that quantifies the level of uncertainty or diversity in a dataset. For packet_in messages, this can be done by first collecting the source IP addresses from 50 incoming packet_in messages.

To calculate entropy, the following steps are taken:

- determine the frequency of each unique source IP address in the 50 packets,

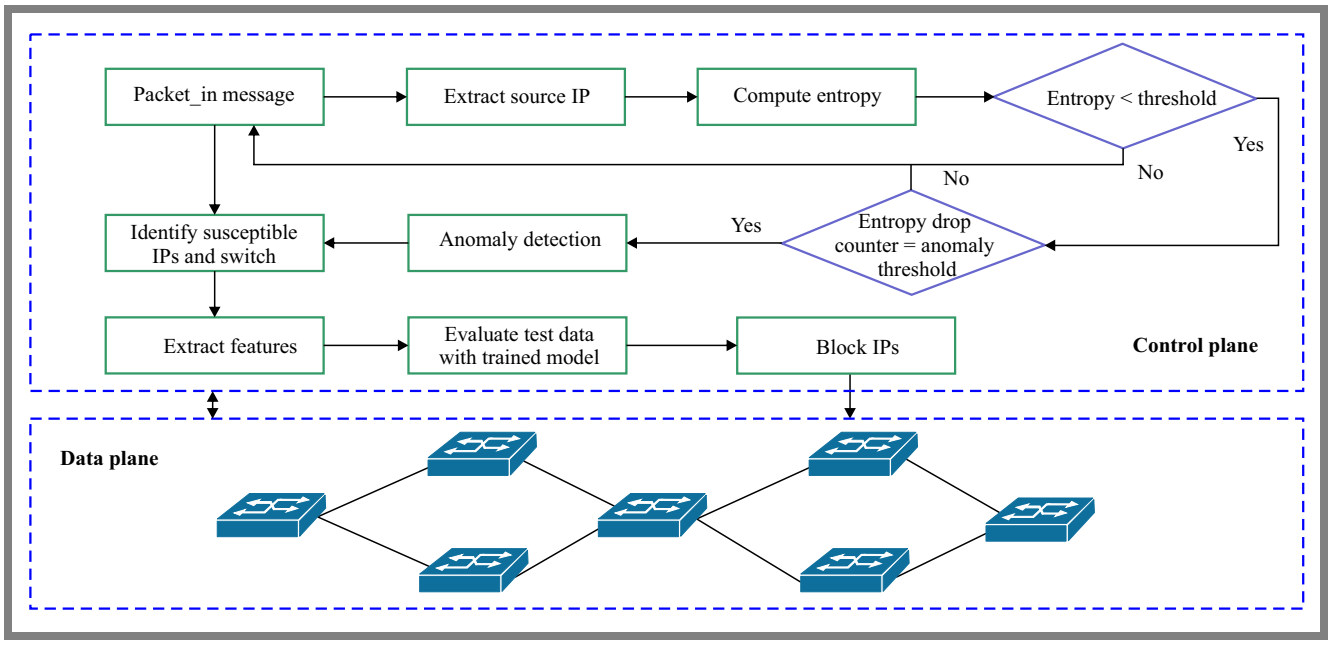

Fig. 1. Diagram of the proposed system.

- compute probability p_i of each source IP i occurring, which is the frequency of i divided by 50 (the total number of packet_in messages),
- use the entropy formula $H(X) = -\sum p_i \log_2(p_i)$, where X represents the set of source IPs.

The resulting entropy value $H(X)$ indicates the distribution of source IP addresses. A higher entropy value suggests a more diverse and unpredictable set of source IPs, while a lower entropy value indicates less diversity and more predictability. For instance, if each of the 50 packets has a unique source IP, the entropy will be higher compared to a scenario where all packets originate from a few IPs. The process of computing entropy is shown in Algorithm 1.

We utilize an entropy threshold of 0.5 to flag potentially suspicious IPs. The entropy drop counter serves as a metric to monitor such activity. Entropy is calculated over a 50–packet window. If entropy falls below the threshold during this window, indicating suspicious behavior, the counter is incremented. Detection of suspicious activity occurs when entropy remains below the threshold for 10 consecutive instances. Suspicious switches associated with these IPs are identified through flow rules. Table 1 presents the statistically derived entropy values from the different scenarios which were used to set the 0.5 threshold.

Once an anomaly is detected, a deep learning module is then employed to capture test traffic for 3 s on the identified switch to detect potential DDoS attacks. This iterative process continues until the end of the emulation process. The algorithm for malicious activity detection is presented as Algorithm 2.

3.3. Data Preprocessing

InSDN [27] addresses the lack of publicly available datasets for evaluating IDSs in SDN environments. The dataset is created using a virtualized network testbed with various VMs to simulate both legitimate and malicious traffic.

Algorithm 1 Calculate_Entropy subroutine

Input: IP_List, packet_count
Output: $Entropy$

```
Entropy ← 0.0
for each (IP, value) in IP_List.items() do
    p ← value / packet_count
    if p > 0 then
        Entropy ← Entropy − (p · log2(p))
    end if
end for
return Entropy
```

The dataset includes attacks such as DDoS, scanning, spoofing, and other common network attacks with 81 features. This dataset was used for training and the test data were collected from Mininet emulation with Ryu as the controller. Preprocessing of the InSDN dataset involves several key steps to prepare the data for effective use in training and evaluating machine learning models for DDoS attack detection in SDN environments. The preprocessing phase includes cleaning the dataset to remove noise and inconsistencies, handling missing values, and normalizing numerical features to ensure uniformity and enhance model performance. Feature engi-

Tab. 1. Entropy value with different attack rates.

Scenario	Attack packet rate [%]	Entropy value
1	80	0.39
2	60	0.54
3	40	0.66

Algorithm 2 Detect_Malicious_Activity procedure

Input: Network traffic
Output: Hosts with malicious activity

Initialize $packet\_count \leftarrow 0$
Initialize $Entropy\_Drop\_Counter \leftarrow 0$
Initialize $Threshold \leftarrow 0.3$
Initialize IP_List as empty dictionary
Initialize $Flag\_DDoS\_Detected \leftarrow 0$
while receiving packets **do**
    $src\_ip \leftarrow$ get_source_ip(packet)
    $packet\_count \leftarrow packet\_count + 1$
    if src_ip exists in IP_List **then**
        $IP\_List[src\_ip] \leftarrow IP\_List[src\_ip] + 1$
    else
        $IP\_List[src\_ip] \leftarrow 1$
    end if
    if $packet\_count = 50$ **then**
        $Entropy \leftarrow$ Calculate_Entropy
        if $Entropy \leqslant Threshold$ **then**
            Increment $Entropy\_Drop\_Counter$
        else
            $Entropy\_Drop\_Counter \leftarrow 0$
        end if
        if $Entropy\_Drop\_Counter = 10$ **then**
            $Flag\_DDoS\_Detected \leftarrow 1$
            Identify the suspicious IPs
            Identify suspicious switches
            Invoke DL model on identified switches
        end if
        $packet\_count \leftarrow 0$
        Clear IP_List
    end if
end while

neering techniques are applied to extract relevant features that are capable of distinguishing between normal and malicious network behaviors.

Additionally, data augmentation methods are utilized to increase diversity and robustness of the dataset, ensuring comprehensive coverage of potential attack scenarios. The test data is generated using a Mininet-emulated topology, as shown in Fig. 2. Mininet provides a realistic virtual network environment in which various network configurations and topologies can be simulated. This allows for thorough testing and evaluation of network performance and security measures, ensuring that the results apply to real-world scenarios.

3.4. Multiple Feature Selection Methods

Three feature extraction methods, namely SelectKBest, ANOVA (analysis of variance) F-value scores, and feature importance scores from a random forest (RF) classifier were employed to select the top ten features from a total of 81. Each method initially selected 20 features, and the ten most frequently occurring features across these selections were then chosen based on their rankings.

Tab. 2. Top 10 features selected from the InSDN dataset.

No.	Name of feature	Description
1	Src IP	IP address source from which the packet is sent
2	Dst IP	IP address at which the packet is intended to be sent
3	Protocol	The communication network protocol (e.g. TCP, UDP)
4	Pkt Len Std	The standard deviation of the flow's packet lengths
5	Flow ID	Distinct number for every flow
6	Bwd Pkt Len Mean	The mean packet length in the reverse direction
7	Pkt Len Var	Variation in the duration of packets within the flow
8	Src Port	Source port number used by the packet
9	Bwd Seg Avg Size	Segment size average moving backward
10	Bwd Pkt Std Len	Packet length standard deviation in the reverse direction

SelectKBest is a feature extraction method that selects the top k features based on univariate statistical tests. This method scores each feature individually using a chosen statistical test and retains the highest-scoring features, making it useful for quickly identifying the most relevant features for predictive modeling.

ANOVA F-value scores are used to assess the significance of the differences between group means in a dataset. In feature selection, ANOVA F-value scores evaluate the relationship between each feature and the target variable, selecting features that show the highest F-values, indicating a stronger discriminatory power between classes.

Feature importance scores from a RF classifier provide a measure of the significance of each feature in predicting the target variable. The random forest algorithm computes these scores by averaging the reduction in impurity across all trees in the forest for each feature. Features with higher importance scores contribute more to the model's predictive accuracy, making them valuable for feature selection.

By employing these three methods, researchers can identify the most informative features in their dataset, improving the performance and efficiency of machine learning models. The ten best-selected features are represented in Tab. 2. Algorithm 3 illustrates the steps for feature selection.

3.5. Deep Learning Module

We have used MLP, CNN and LSTM to evaluate the performance of the classification approach. The deep learning model is trained using the InSDN dataset, and testing is performed

Algorithm 3 Feature selection using multiple methods

Input: Dataset D with 81 features, target variable y
Output: Top 10 selected features

Initialize parameters:
$n \leftarrow 81$
$k \leftarrow 20$
$m \leftarrow 10$
Apply feature selection methods:
Initialize an empty list $selected\_features$
for each method in {SelectKBest, ANOVA F-value, RandomForest} **do**
    if method is SelectKBest **then**
        Perform univariate statistical tests
        Select the top k features based on their scores
    else if method is ANOVA F-value **then**
        Compute the F-value for each feature
        Select top k features based on F-value scores
    else if method is RandomForest **then**
        Train a RandomForest classifier on the dataset
        Calculate feature importance scores
        Select the top k features based on their importance scores
    end if
    Append the chosen features to $selected\_features$
end for
Combine selected features:
Initialize a dictionary $feature\_frequency$ to count the frequency of each feature
for feature in $selected\_features$ **do**
    if feature is already in $feature\_frequency$ **then**
        Increment the count of $feature$ by 1
    else
        Add $feature\_frequency$ with count 1
    end if
end for
Select top m features:
Sort the features in $feature\_frequency$ by their frequency in descending order
Select the top m features with the highest frequency
Output: Return the list of the top 10 features

using real-time traffic generated with the help of Mininet and the Ryu controller in SDN.

Algorithm 4 illustrates the steps involved in the design.

4. Results and Discussion

As shown in Fig. 2, a Mininet emulator is used to create the SDN environment. We employed a leaf and spine topology consisting of three OpenFlow virtual switches and 20 hosts. All the switches are managed by the Ryu controller, which provides centralized control and enables dynamic network management. Both normal and abnormal traffic are sent from various source IPs to multiple target machines. We also simulated spoofed source IP addresses for the attack. The `hping3` tool is used to generate the traffic.

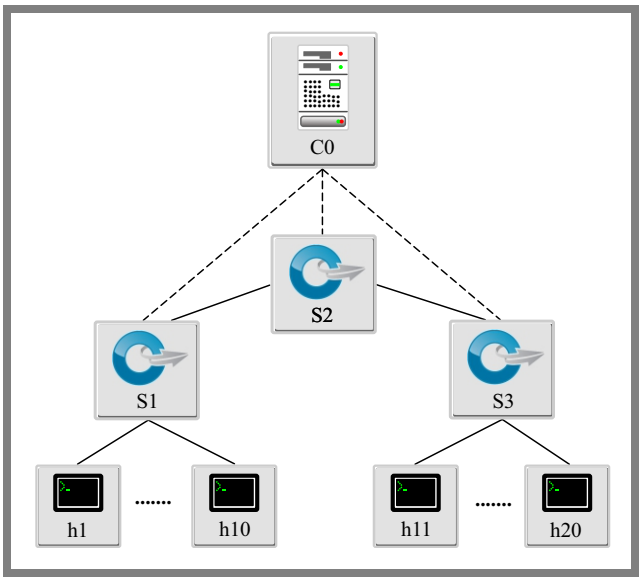

Fig. 2. Mininet network topology.

Algorithm 4 LSTM model for DDoS attacks detection

Input: Preprocessed dataset $selected\_features\_df$ with features and labels
Output: Trained LSTM model and evaluation metrics

Build LSTM model:
    Initialize a sequential model
    Add an LSTM layer with ReLU activation
    Add a dense layer with sigmoid activation function
    Compile the model using Adam optimizer and binary_cross_entropy loss function
Train the model:
for $epoch \leftarrow 1$ to $epochs$ **do**
    Train the model on X_{train} and y_{train} for one epoch with $batch\_size$
    Validate the model on X_{test} and y_{test} during training
    if $epoch$ is the final epoch **then**
        Print "Final epoch reached: Epoch" $epoch$
    else
        Print "Continuing to next epoch: Epoch" $epoch$
    end if
end for
Evaluate the model:
    Evaluate the model on the test set X_{test} and y_{test}
    Print the test accuracy

4.1. Machine Learning Results

We use machine learning (ML) algorithms to detect the attack on suspicious switches. Based on anomaly detection, traffic is captured, over a fixed time period, from suspicious switches in the SDN environment. Figure 3 illustrates the accuracy of various ML algorithms. The trained model utilized for this evaluation is based on the InSDN dataset.

We assessed the performance of machine learning algorithms using the InSDN dataset with the top 10 features selected, from an initial set of 81, as training and test data generated using the Mininet emulator and the Ryu controller. The results

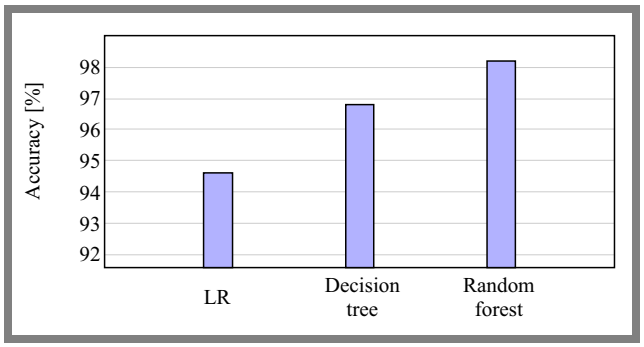

Fig. 3. Accuracy of machine learning algorithms.

showed that logistic regression (LR) achieved an accuracy of 94.6%, the decision tree classifier reached an accuracy of 96.8%, and the random forest (RF) classifier outperformed the other models with an accuracy of 98.2%.

Table 3 shows the precision, recall and F1-score levels of ML models. The results indicate that RF performs better than decision tree and LR in terms of accuracy due to its ensemble learning approach, which reduces overfitting and stabilizes predictions by averaging multiple trees. This method effectively handles complex relationships and interactions among features, making it more robust to noise and outliers. RF also evaluates feature importance across all trees, ensuring that key features are prioritized.

These factors collectively contribute to the superior accuracy of random forest, compared with other models. The receiver operating characteristic curve (ROC) and confusion matrix are represented in Figs. 4 and 5, respectively. The ROC curve plots the true positive rate (TPR) against the false positive rate (FPR), with a red line representing the performance of the RF model and a blue dashed line representing a random classifier, for comparison. The coordinates used for the RF model indicate its performance at various threshold levels, demonstrating its effectiveness in distinguishing between classes with the result of 0.99. A ROC value of 0.99 indicates that the RF model has a high TPR and a low FPR, demonstrating its effectiveness and reliability in detecting the targeted instances. Figure 5 presents a confusion matrix for a random forest model, illustrating its performance in terms of classification accuracy. This graph illustrates the model's effectiveness and the distribution of prediction errors, emphasizing its performance in distinguishing between the two classes.

Tab. 3. Performance of the ML models.

ML algorithms	Precision (P)	Recall (R)	F1-score
Decision tree	0.966	0.97	0.968
Logistic regression (LR)	0.951	0.94	0.942
Random forest (RF)	0.984	0.98	0.982
Trained with the InSDN dataset and tested with data generated using Mininet emulation			

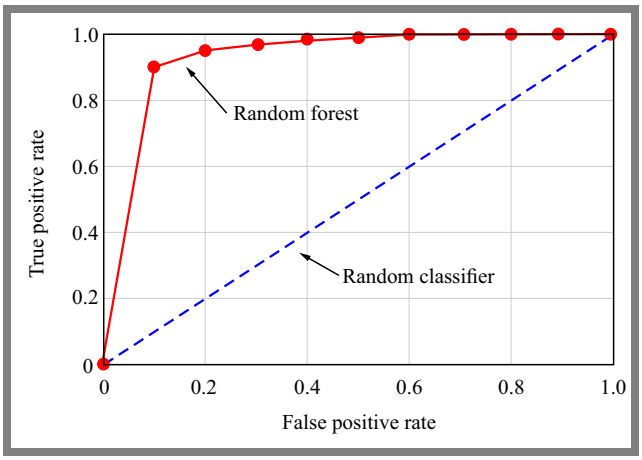

Fig. 4. Receiver operating characteristic.

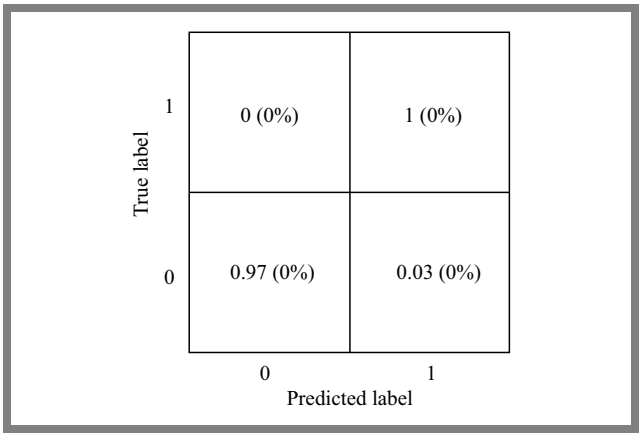

Fig. 5. Confusion matrix of random forest.

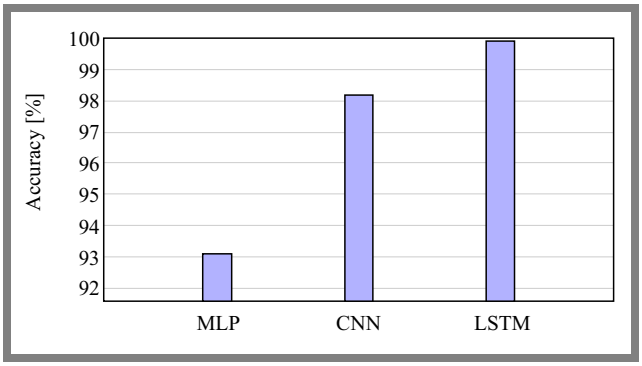

Fig. 6. Accuracy of deep learning algorithms.

4.2. Deep Learning Results

Deep learning models are essential due to their ability to capture complex patterns and relationships within data. These models handle high-dimensional data effectively, distinguishing between normal and malicious traffic, even when significant variability is at play. The scalability of deep learning allows it to be applied in large-scale network environments without performance losses.

We use three deep learning models: LSTM, CNN, and MLP, each with different hyperparameter tuning. The LSTM model consists of an LSTM layer followed by dense layers, capturing sequential patterns within the feature set, with a softmax out-

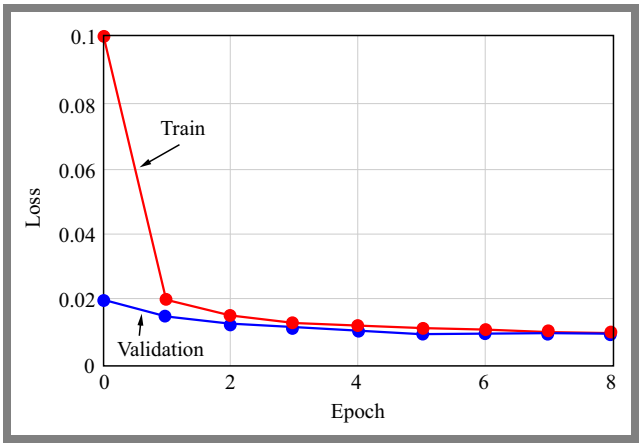

Fig. 7. Model loss of LSTM.

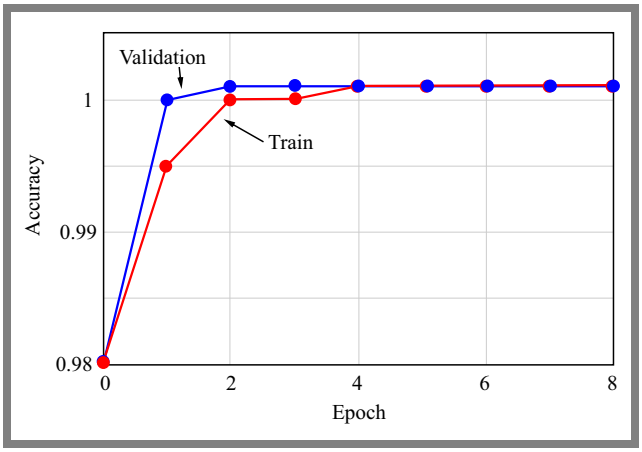

Fig. 8. Model accuracy of LSTM.

put layer. It is trained using categorical cross-entropy loss and the Adam optimizer. The MLP model is configured with an input layer matching the number of selected features, several hidden layers with ReLU activation functions, and a softmax output layer. It is trained using backpropagation with a categorical cross-entropy loss function. The CNN model includes an input layer reshaped for a 2D convolutional architecture, multiple convolutional layers with ReLU activation, max-pooling layers, and fully connected layers, culminating in a softmax output layer.

As shown in Fig. 6, LSTM achieved the highest accuracy among the three models (99.83%), outperforming both CNN and MLP models, thus highlighting its superior capability in capturing temporal relationships within the selected features.

LSTM is better at managing sequential data and capturing temporal dependencies, making them suitable for analyzing networktraffic. The built-in memory mechanism retains long-term dependencies to store anomalous traffic patterns. LSTM has allowed to enhance the process of analyzing temporal patterns and detecting anomalies within network traffic data.

Figure 7 represents the loss graph during the training phase of the dataset and the validation loss obtained after evaluating the validation dataset. The graph demonstrates that the loss experienced decreases gradually over the course of multiple epochs. On the other hand, the validation loss exhibits fluctuations at epoch 8 before stabilizing over the remaining epochs. Figure 8 depicts the accuracy graph during the training phase of the dataset and the validation accuracy obtained after evaluating the validation dataset. The results reveal that the accuracy improves gradually over the number of epochs during training. However, at epoch 8, there is a drop in validation accuracy, followed by a sudden increase, after which the accuracy remains stable over subsequent epochs.

In summary, the LSTM model achieves the highest accuracy level (99.83%) among the deep learning models. The random forest model achieves the highest accuracy (98.2%) among machine learning models. Therefore, we conclude that deep learning outperforms machine learning. Consequently, we use the LSTM model to detect DDoS attacks in data captured from suspicious switches identified by the entropy module. This collaborative approach helps reduce the DDoS attack detection lead time.

5. Conclusions

SDN is a centrally controlled and programmable network architecture that enables a flexible network environment. The primary concern in SDN networks is the issue of security. This work aims to address the security challenge by designing a method for detecting DDoS attacks in SDN. The proposed approach consists of two modules. Initially, the entropy module is deployed on the control plane to detect suspicious switches and hosts based on entropy. Then, a deep learning module is used to detect DDoS attacks using the test data captured from the suspicious switches. The deep learning module performs the important feature extraction and attack classification tasks using deep learning algorithms, such as MLP, CNN, and LSTM. Based on the experiments and evaluation results, the system achieves an average accuracy of 99.83% in categorizing various types of DDoS flooding attacks, e.g. UDP, TCP-SYN, and ICMP. The experiments indicate that the deep learning techniques yield better evaluation results, particularly in terms of model accuracy and model loss.

As future work, we plan to implement the proposed mechanism within a real-time SDN testbed based on the OpenStack cloud.

References

[1] B. Alhijawi *et al.*, "A Survey on DoS/DDoS Mitigation Techniques in SDNs: Classification, Comparison, Solutions, Testing Tools and Datasets", *Computers and Electrical Engineering*, vol. 99, art. no. 107706, 2022 (https://doi.org/10.1016/j.compeleceng.2022.107706).

[2] A.A. Bahashwan *et al.*, "A Systematic Literature Review on Machine Learning and Deep Learning Approaches for Detecting DDoS Attacks in Software-defined Networking", *Sensors*, vol. 23, no. 9, art. no. 4441, 2023 (https://doi.org/10.3390/s23094441).

[3] I.A. Valdovinos, J.A. Pérez-Díaz, K.-K.R. Choo, and J.F. Botero, "Emerging DDoS Attack Detection and Mitigation Strategies in Software-defined Networks: Taxonomy, Challenges and Future Directions", *Journal of Network and Computer Applications*, vol. 187, art. no. 103093, 2021 (https://doi.org/10.1016/j.jnca.2021.103093).

[4] H. Zhang, L. Zhou, and J. Lei, "Renyi Entropy-based DDoS Attack Detection in SDN-based Networks", *2023 IEEE 3rd International Conference on Electronic Technology, Communication and Information (ICETCI)*, Changchun, China, 2023 (https://doi.org/10.1109/ICETCI57876.2023.10176631).

[5] A. Makuvaza, D.S. Jat, and A.M. Gamundani, "Deep Neural Network (DNN) Solution for Real-time Detection of Distributed Denial of Service (DDoS) Attacks in Software Defined Networks (SDNs)", *SN Computer Science*, vol. 2, 2021 (https://doi.org/10.1007/s42979-021-00467-1).

[6] L. Tan *et al.*, "A New Framework for DDoS Attack Detection and Defense in SDN Environment", *IEEE Access*, vol. 8, pp. 161908–161919, 2020 (https://doi.org/10.1109/ACCESS.2020.3021435).

[7] K. Kalkan, L. Altay, G. Gür, and F. Alagöz, "JESS: Joint Entropy-based DDoS Defense Scheme in SDN", *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 10, pp. 2358–2372, 2018 (https://doi.org/10.1109/JSAC.2018.2869997).

[8] R.N. Carvalho, J.L. Bordim, and E.A.P. Alchieri, "Entropy-based DoS Attack Identification in SDN", *2019 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*, Rio de Janeiro, Brazil, 2019 (https://doi.org/10.1109/IPDPSW.2019.00108).

[9] S. Yu *et al.*, "A Cooperative DDoS Attack Detection Scheme Based on Entropy and Ensemble Learning in SDN", *EURASIP Journal on Wireless Communications and Networking*, 2021 (https://doi.org/10.21203/rs.3.rs-154522/v1).

[10] B. Han *et al.*, "OverWatch: A Cross-plane DDoS Attack Defense Framework with Collaborative Intelligence in SDN", *Security and Communication Networks*, 2018 (https://doi.org/10.1155/2018/9649643).

[11] A. Yadav *et al.*, "A Hybrid Approach for Detection of DDoS Attacks Using Entropy and Machine Learning in Software Defined Networks", *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Kharagpur, India, 2021 (https://doi.org/10.1109/ICCCNT51525.2021.9580057).

[12] B. Celesova, J. Val'ko, R. Grezo, and P. Helebrandt, "Enhancing Security of SDN Focusing on Control Plane and Data Plane", *2019 7th International Symposium on Digital Forensics and Security (ISDFS)*, Barcelos, Portugal, 2019 (https://doi.org/10.1109/ISDFS.2019.8757542).

[13] Y. Wang *et al.*, "SGS: Safe-guard Scheme for Protecting Control Plane Against DDoS Attacks in Software-defined Networking", *IEEE Access*, vol. 7, pp. 34699–34710, 2019 (https://doi.org/10.1109/ACCESS.2019.2895092).

[14] W.G. Gadallah, N.M. Omar, and H.M. Ibrahim, "Machine Learning-based Distributed Denial of Service Attacks Detection Technique using New Features in Software-defined Networks", *International Journal of Computer Network Information Security*, vol. 13, no. 3, pp. 15–27, 2021 (https://doi.org/10.5815/ijcnis.2021.03.02).

[15] Z. Liu, Y. He, W. Wang, and B. Zhang, "DDoS Attack Detection Scheme Based on Entropy and PSO-BP Neural Network in SDN", *China Communications*, vol. 16, no. 7, pp. 144–155, 2019 (https://doi.org/10.23919/JCC.2019.07.012).

[16] H.S. Abdulkarem and A. Dawod, "DDoS Attack Detection and Mitigation at SDN Data Plane Layer", *2020 2nd Global Power, Energy and Communication Conference (GPECOM)*, Izmir, Türkiye, 2020 (https://doi.org/10.1109/GPECOM49333.2020.9247850).

[17] R. Sanjeetha, A. Pattanaik, A. Gupta, and A. Kanavalli, "Early Detection and Diminution of DDoS Attack Instigated by Compromised Switches on the Controller in Software Defined Networks", *2019 IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics*, Manipal, India, 2019 (https://doi.org/10.1109/DISCOVER47552.2019.9007925).

[18] A. Mishra, N. Gupta, and B.B. Gupta, "Defense Mechanisms Against DDoS Attack Based on Entropy in SDN-cloud Using POX Controller", *Telecommunication Systems*, vol. 77, pp. 47–62, 2021 (https://doi.org/10.1007/s11235-020-00747-w).

[19] X. Yang, B. Han, Z. Sun, and J. Huang, "SDN-based DDoS Attack Detection with Cross-plane Collaboration and Lightweight Flow Monitoring", *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, Singapore, 2017 (https://doi.org/10.1109/GLOCOM.2017.8254079).

[20] V. Deepa, K.M. Sudar, and P. Deepalakshmi, "Detection of DDoS Attack on SDN Control Plane using Hybrid Machine Learning Techniques", *2018 International Conference on Smart Systems and Inventive Technology (ICSSIT)*, Tirunelveli, India, 2018 (https://doi.org/10.1109/ICSSIT.2018.8748836).

[21] J.E. Varghese and B. Muniyal, "An Efficient IDS Framework for DDoS Attacks in SDN Environment", *IEEE Access*, vol. 9, pp. 69680–69699, 2021 (https://doi.org/10.1109/ACCESS.2021.3078065).

[22] N. Ahuja, G. Singal, D. Mukhopadhyay, and N. Kumar, "Automated DDOS Attack Detection in Software Defined Networking", *Journal of Network and Computer Applications*, vol. 187, art. no. 103108, 2021 (https://doi.org/10.1016/j.jnca.2021.103108).

[23] D. Gadze *et al.*, "An Investigation into the Application of Deep Learning in the Detection and Mitigation of DDOS Attack on SDN Controllers", *Technologies*, vol. 9, art. no. 14, 2021 (https://doi.org/10.3390/technologies9010014).

[24] P.T. Dinh and M. Park, "BDF-SDN: A Big Data Framework for DDoS Attack Detection in Large-scale SDN-based Cloud", *2021 IEEE Conference on Dependable and Secure Computing (DSC)*, Fukushima, Japan, 2021 (https://doi.org/10.1109/DSC49826.2021.9346269).

[25] R. Durner, C. Lorenz, M. Wiedemann, and W. Kellerer, "Detecting and Mitigating Denial of Service Attacks Against the Data Plane in Software Defined Networks", *2017 IEEE Conference on Network Softwarization (NetSoft)*, Bologna, Italy, 2017 (https://doi.org/10.1109/NETSOFT.2017.8004229).

[26] N. Ahuja, G. Singal, and D. Mukhopadhyay, "DLSDN: Deep Learning for DDOS Attack Detection in Software Defined Networking", *2021 11th International Conference on Cloud Computing, Data Science and Engineering (Confluence)*, Noida, India, 2021 (https://doi.org/10.1109/Confluence51648.2021.9376879).

[27] M.S. Elsayed, N.-A. Le-Khac, and A.D. Jurcut, "InSDN: A Novel SDN Intrusion Dataset", *IEEE Access*, vol. 8, pp. 165263–165284, 2020 (https://doi.org/10.1109/ACCESS.2020.3022633).

Narayan D.G., Ph.D.
School of Computer Science and Engineering
https://orcid.org/0000-0002-2843-8931
E-mail: narayan_dg@kletech.ac.in
KLE Technological University, Hubballi, Karnataka, India
https://www.kletech.ac.in

Heena W, Ph.D.
School of Computer Science and Engineering
https://orcid.org/0009-0009-7998-3049
E-mail: heenakwalikar@gmail.com
KLE Technological University, Hubballi, Karnataka, India
https://www.kletech.ac.in

Amit K, Ph.D.
Department of MCA
https://orcid.org/0000-0002-8917-8678
E-mail: amitk@kletech.ac.in
KLE Technological University, Hubballi, Karnataka, India
https://www.kletech.ac.in

Achieving Reliability of Privacy-preserving Phantom Routing Protocols in Multi-hop Wireless Sensor Networks

Lilian C. Mutalemwa

The Open University of Tanzania, Dar es Salaam, Tanzania

https://doi.org/10.26636/jtit.2024.3.1703

Abstract — Due to the open nature of wireless channels and sensor node resource constraints, it is challenging to secure the communication in wireless sensor networks (WSNs) while simultaneously protecting the privacy of node location data. Therefore, a significant amount of research focusing on source location privacy (SLP) protocols has been conducted. The amount of research on SLP reliability, meanwhile, is insignificant. This study explores the operational features of various privacy-preserving phantom routing protocols and simulates WSNs with varied network configurations to investigate how different routing strategies affect SLP reliability. Safety period and capture ratio metrics are used to compute SLP reliability. Simulation results show that integration of phantom routing with fake packet distribution mechanisms adversely impacts SLP reliability. SLP reliability decreases also as the number of fake packet sources increases. Research proves that a protocol with many fake packet sources achieves SLP reliability for a mission duration of 940 rounds, while a protocol with no fake packet sources achieves SLP reliability for 1658 rounds.

Keywords — phantom routing, privacy preservation, reliability, source location privacy, wireless sensor network

1. Introduction

Wireless sensor networks (WSNs) employ multi-hop wireless communication mechanisms. Due to the open nature of wireless channels and limited resources available in sensor nodes, they also pose a number of challenges related to the application of WSNs [1]–[4]. For example, due to the absence of a secure infrastructure, it is challenging to preserve the locations privacy of critical sensor nodes, such as source nodes [2], [5], [6].

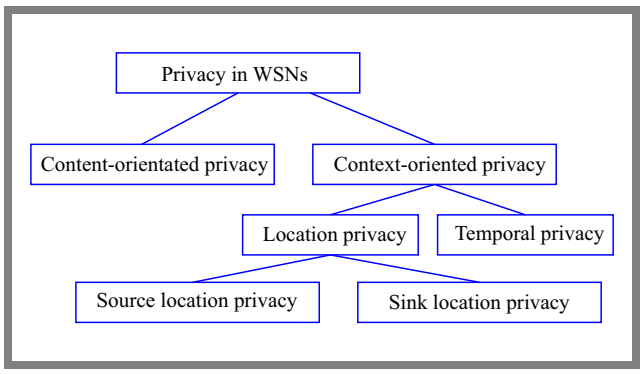

Fig. 1. Classification of privacy in WSNs.

To address this challenge, much research has been conducted on the topics of security and privacy in WSNs. This study focuses on source location privacy (SLP) in WSNs – an issue that is a specific aspect of context-oriented privacy [7]. As shown in Fig. 1, privacy in WSNs is classified into that of content-oriented and context-oriented categories. While content-oriented privacy ensures non-repudiation and confidentiality of data, context-oriented privacy is concerned with temporal and location privacy [8].

SLP is important when WSNs are deployed in safety-critical applications [3], [5], [9]. For example, in tactical military applications, SLP guarantees security and reduces the number of fatalities or victims on a battlefield. In healthcare, SLP mechanisms are employed in patient monitoring systems to ensure patient data are not disclosed to unauthorized users [5]. In wildlife monitoring, SLP techniques are employed to ensure that information on the location of endangered animal species is not exposed to illegal hunters [2], [3], [10]. SLP techniques are also important in underwater Internet of Things (UIoT) infrastructures, where WSNs are deployed for underwater exploration activities to ensure economic growth based the concept of Blue Economy [11], [12].

SLP routing protocols are used to provide SLP protection in WSNs [3], [9], [13]–[15]. In the existing literature, extensive projects on SLP routing protocols and their performance may be found. However, only a handful of papers focus on SLP reliability.

The study described in [16] reported that a lot of the existing literature focuses on analyzing connectivity-oriented and flow-oriented reliability. For example, the recent studies on reliability, as presented in [17]–[20], did not analyze SLP reliability. Although the authors of [20] studied privacy issues, their reliability evaluations focused mainly on data transmission reliability. SLP reliability was determined for the first time in [9].

To address this gap, this study analyzes SLP reliability and evaluates the energy efficiency of protocols, as inefficient use of energy in WSNs affects SLP reliability [9]. Moreover, to ensure higher operational efficiency of WSNs and SLP

This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) License.
For more information, see https://creativecommons.org/licenses/by/4.0/

protocols, it is important to balance energy distribution inside the WSN [21]–[23].

Three representative SLP protocols are considered: tree-based SLP protocol (TRP) [24], probabilistic SLP protocol (PRP) [25], and ring-based SLP protocol (RNP) [26]. The selection of TRP, PRP, and RNP protocols is based on the fact that TRP, PRP, and RNP employ phantom routing techniques. However, each protocol involves different routing strategies, as shown in Tab. 1. Therefore, it is interesting to observe the effects that each routing strategy has on privacy preservation and SLP reliability. Performance-related gains and limitations of each routing strategy are investigated as well.

The contribution of this study is summarized as follows:

- the paper explores operational features of various privacy-preserving phantom routing protocols and identifies the key similarities and differences in the routing strategies of TRP, PRP, and RNP,
- performance of TRP, PRP, and RNP is evaluated against an estimating adversary that uses the hidden Markov model,
- the impact of the routing strategies in TRP, PRP, and RNP on SLP reliability is investigated and energy efficiency for different network configurations is analyzed to identify the SLP protocol with superior performance features.

2. Related Work

Ozturk *et al.* introduced, in [27], a phantom routing strategy for SLP. Many other studies have explored phantom routing techniques for SLP, including those presented in [22], [26], [28]–[33]. Operational features of the phantom routing protocols are presented below.

The phantom routing protocol described in [32] is based on multiple sinks and a dynamic multipath routing strategy. It is designed to provide SLP protection for Internet of Things systems. The protocol generates a candidate area and selects phantom nodes to ensure the consumption of energy is more significant in the non-hotspot regions. Then, to increase dissemination randomness during the transmission, the protocol employs a packet segmentation technique to ensure the packets are segmented into many portions. The portions are transmitted through multiple dynamic paths. Furthermore, the protocol forms a closed-loop from the sink node to improve SLP protection.

Tab. 1. Summary of the key similarities and differences in the techniques of TRP, PRP, and RNP.

Routing technique	Protocol		
	TRP	PRP	RNP
Phantom routing	✓	✓	✓
Two-level phantom routing	×	×	✓
Fake packet distribution	✓	✓	×
Distributes fake packets in the vicinity of phantom node	✓	×	×
Distributes fake packet traffic in the sink node region	×	✓	×
Generates multiple sources of fake packets	✓	×	×
Isolates fake sources from real source nodes	×	✓	×
Long fake packet routing paths	✓	×	×

The protocol presented in [29] uses phantom nodes, rings, and fake paths. It is useful in networks that deploy multiple source or sink nodes. Fake packets are employed to mimic the behavior of real packets and perplex the adversary. The protocol transmits packets in four phases. In phase 1, the source node detects an event and generates real packets. The real packets are conveyed to the phantom node. In phase 2, the phantom node relays the real packets to the circular region. In phase 3, the real packets are conveyed clockwise in the circular region and arrive at the sink proxy node. The fourth phase involves the sink proxy node transmission process, where real packets are transmitted to the destination sink node.

The backbone-based protocol [26] relies on a routing technique to address the challenges faced by protocols that broadcast heavy traffic. The protocol introduces multiple levels of adversary confusion. It also employs a random backbone route between the sink node and the neighboring node of the phantom nodes, and packets are routed using a directed random-walk routing strategy.

It was highlighted in [28] that many of the existing phantom routing protocols are characterized by a high communication overhead and high energy consumption during the process of selecting the phantom node. Consequently, SLP protection and network lifetime are affected. To address this challenge, grid-based protocols [28] use powerful sink nodes to select phantom nodes. When the phantom nodes are selected by sink nodes, the energy of the ordinary sensor nodes is conserved.

The phantom walkabouts protocol described in [31] is a more general version of the phantom routing approach and presents phantom routes of variable lengths. It is different from the traditional phantom routing protocols, in which nodes have two sets of neighboring nodes. Here, the phantom walkabouts protocol divides each node's neighbors into four sets in different directions and creates a long random walk to generate a safeguarding distance between the phantom nodes and the real source node. As a result, the protocol achieves high levels of SLP protection and outperforms the baseline phantom routing protocol.

The phantom with angle protocol from [22] preserves the SLP and guarantees energy efficiency by regulating the traffic load in the network. The protocol locates phantom nodes at a safeguarded distance, away from the source nodes. Furthermore, it allocates multiple unique regions for phantom node selection. All the regions are assigned equal probability of selection. To ensure a new phantom node is selected for each successive packet, the protocol assigns a phantom selection factor.

This study evaluates SLP reliability of phantom routing protocols that were proposed in [24]–[26]. To route packets,

the TRP protocol proposed in [24] establishes a backbone route from the sink node to the boundaries of the network. Then, it establishes many redundant diversionary routes, serving as branch routes of the backbone route. Fake packets are distributed throughout the diversionary routes, while source nodes send packets to the sink node through the phantom nodes. To ensure strong SLP, TRP safeguards the location of phantom nodes located at a certain distance from the source nodes.

The PRP protocol [25] considers exposed regions and employs the directed random walk technique to ensure that selected phantom nodes are at a safe distance away from the source nodes. However, unlike in the case of TRP, PRP generates fake packet sources in the vicinity of the sink node and allows only one node to act as a fake source, for a fixed period of time. Also, PRP employs dynamic packet routes that vary based on source-sink distance and transmits both real and fake packets simultaneously. Real packets are routed through the selected phantom nodes. Fake packets are routed through rings and arcs facing different directions.

The RNP protocol presented in [26] does not broadcast fake packets. To ensure strong SLP, RNP relies on highly randomized packet routing paths that are less predictable to the adversaries. To achieve that goal, RNP constructs two levels of phantom nodes that provide strong adversary obfuscation effects.

3. Models

3.1. Network Model

Here, a WSN with numerous sensor nodes and a sink node is considered. The sink node is located at the center of the WSN. The sensor nodes are set to detect events. Once a sensor node detects an event within its sensing range, it becomes the source node and starts sending packets to the sink node using the multi-hop communication technique. If a sensor node fails to detect an event, it follows the sleeping pattern.

All sensor nodes are static and are characterized by the same memory capacity, initial energy, and computing ability. Neighboring sensor nodes are within each other's communication range and are able to communicate and exchange data. Figure 2 shows the structure of such a WSN. More details of this network model are presented in [8], [25].

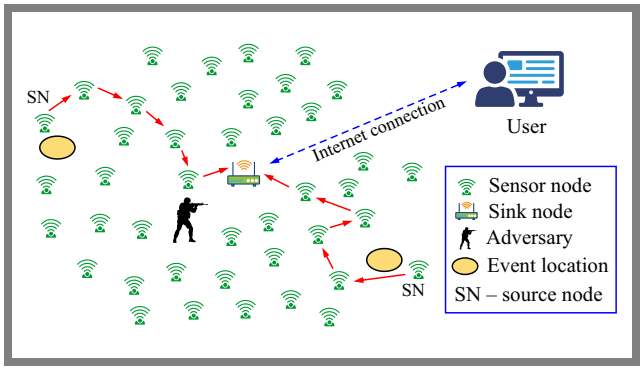

Fig. 2. Structure of the considered WSN.

3.2. Adversary Model

The adversary model uses an estimating adversary that eavesdrops on the communication between sensor nodes. It analyzes the packet traffic and checks the packet type by scanning the header of each packet. Thereafter, it employs the hidden Markov model (HMM) to estimate the likely state of the source nodes at a particular time.

The use of HMM is advantageous to the adversary because it helps make more precise estimations of the source state and predictions of the source location. Additional details of the adversary model are presented in [25].

4. Performance Evaluation

4.1. Simulation Environment

With the use of Matlab simulation software, a WSN was simulated with randomly distributed sensor nodes. The network configuration was adopted from [34]. Table 2 shows a summary of the network simulation parameters.

4.2. Performance Metrics

The performance of TRP, PRP, and RNP protocols was evaluated in terms of energy efficiency, SLP protection, and SLP reliability. The baseline phantom single-path protocol (PHP) [35] was used as a reference solution, for comparison purposes. The following metrics were used to analyze the performance:

- the energy ratio (ER) metric was used to measure energy efficiency,
- node utilization ratio (NUR) and sensor node residual energy metrics were used to evaluate energy distribution in the WSN,
- safety period (SP) and capture ratio (CR) metrics were used to determine the level of SLP preservation,
- safety period reliability (SPR) and capture ratio reliability (CRR) metrics were used to assess SLP reliability.

Tab. 2. Parameters used for network simulations.

Parameter	Value
Network side length	2000 m
Number of sensor nodes	3000
Number of sink nodes	1
Sensor node communication range	40 m
Adversary hearing range	40 m
Adversary attack strategy	Traffic analysing with HMM
Packet size	1024 bits
Source packet rate	Varied between 1 and 4 packet/s
Sensor node initial energy	0.5 J

ER, NUR, SP, SPR, CR, and CRR were computed using the following equations:

$$ER = \frac{E_{600}}{E_t}\,, \tag{1}$$

where E_{600} is the energy used in 600 rounds and E_t is total energy.

$$NUR = \frac{SN_{600}}{N_{sn}}\,, \tag{2}$$

where SN_{600} defines the number of sensor nodes participating in data transmission for 600 rounds and N_{sn} stands for total number of sensor nodes.

$$SP = N_{hops}\,, \tag{3}$$

where N_{hops} is the number of hops during the adversary backtracing attack.

$$CR = \frac{N_e}{T_e}\,, \tag{4}$$

Here, N_e is the number of experiments where the adversary ends in locating the source node while T_e is the total number of experiments.

$$SPR = \begin{cases} 1 & \text{if } e^{\Delta_{SP}} \geqslant 1 \\ 0 & \text{otherwise} \end{cases}\,, \tag{5}$$

$$CRR = \begin{cases} 1 & \text{if } e^{\Delta_{CR}} \geqslant 1 \\ 0 & \text{otherwise} \end{cases}\,. \tag{6}$$

The relationship between ER and energy efficiency is that high ER correlates with low energy efficiency. Also, the relationship between SP, CR, and SLP preservation is that long SP corresponds to high levels of SLP preservation, while low CR corresponds to high levels of SLP preservation [36]. In Eq. (5), Δ_{SP} is the difference between the achieved SP and the application-specific required SP. When $e^{\Delta_{SP}} \geqslant 1$, SPR becomes 1 to indicate that SLP reliability is guaranteed. Otherwise, SPR becomes 0 to indicate that SLP reliability is not guaranteed [9]. Same principles apply to compute CRR, as shown in Eq. (6).

5. Results and Discussions

5.1. Energy Efficiency

Energy efficiency determines the reliability of routing protocols in WSNs [21], [37]–[42]. The ER metric was used to measure energy efficiency. The relationship between ER and energy efficiency is that high ER correlates with low energy efficiency. The energy consumption model presented in [2], [9], [21], [25], [43]–[45] was used.

Figure 3 shows the ER of specific protocols at varied source packet rates. It shows that the ER of TRP is significantly higher than the ER of PRP and RNP. The baseline PHP incurs the lowest ER. TRP and PRP distribute fake packet traffic that increases the energy consumption and ER. Moreover, TRP generates multiple fake packet sources. In the transmission of every real packet, TRP distributes many fake packets, resulting in an increased ER.

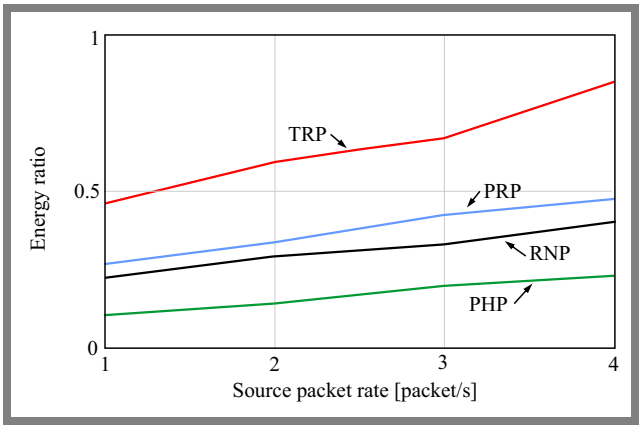

Fig. 3. Energy efficiency of the protocols.

PRP generates a single fake packet source at a time. Therefore, the ER of PRP is significantly lower than the ER of TRP. RNP does not distribute fake packet traffic. As a result, the ER in RNP is lower than in TRP and PRP. However, despite the fact that RNP does not distribute fake packet traffic, the ER of RNP is close to the ER of PRP. This is because RNP employs a two-level phantom routing strategy which creates long routing paths. Long routing paths increase the ER.

The results shown in Fig. 3 suggest that the energy efficiency of TRP is significantly lower. Thus, PRP and RNP outperform TRP in terms of energy efficiency.

5.2. Energy Distribution

The use of multi-hop packet routing technique results in unbalanced energy distribution (ED) in the WSN domain [46], [47]. Unbalanced ED affects the operation of WSNs and results in short-term SLP protection [3]. In systems requiring reliable data transmission or in which as many sensor nodes as possible are required to operate simultaneously, unbalanced ED results in a system failure [48].

To effectively measure the ED of TRP, PRP, and RNP, NUR and residual sensor node energy levels were analyzed. NUR increases along with an increase in randomness and length of the routing paths [32]. Additionally, NUR increases with the grooving number of routing paths and the amount of packet traffic. The relationship between NUR and ED is that if different regions in the WSN incur different levels of NUR, the ED is affected as well. Thus, an area with a high NUR depletes the sensor node's residual energy at a fast rate and uneven distribution of sensor node energy occurs.

In the experiments, NUR and ED were observed in different regions of the WSN: hotspot (HT) regions where the hop distance between a sensor node and the sink node was $\leqslant 20$ hops, and non-hotspot (NHT) regions where the hop distance was > 20.

Figure 4 shows that with the exception of TRP, all the other protocols achieve higher NUR in HT regions due to the increased amount of packet traffic. The unbalanced packet traffic causes unbalanced ED. The NUR of TRP in NHT is

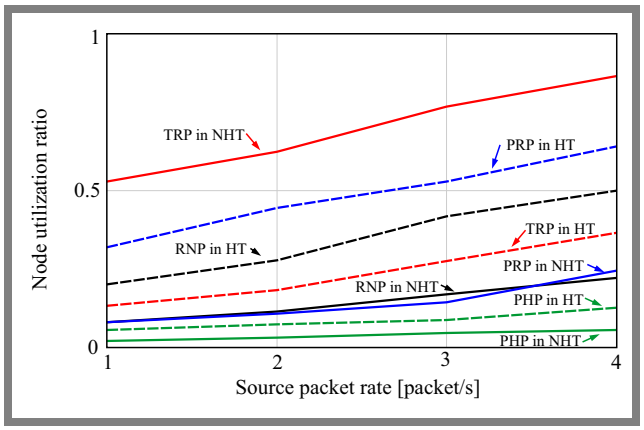

Fig. 4. Node utilization ratio of the protocols.

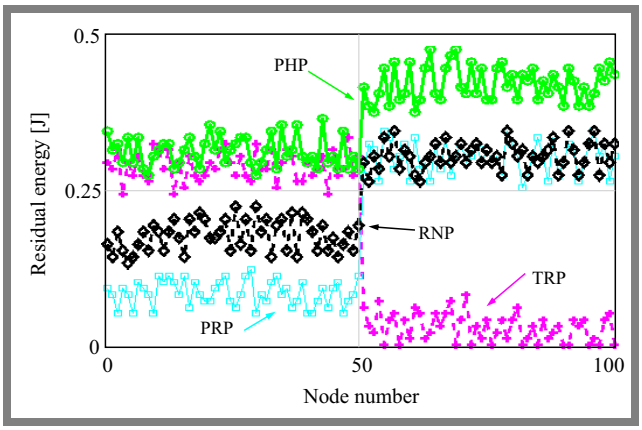

Fig. 5. Energy distribution of the protocols under consideration.

very high, because TRP generates multiple fake packet sources in NHT regions and for every real source node packet TRP distributes many fake packets. Consequently, many routing paths are created to route packet traffic in NHT regions and a high NUR value is incurred.

It is also shown in Fig. 4 that the NUR of PRP is significantly higher in HT regions. This is because PRP distributes fake packet traffic only in HT regions. Therefore, PRP creates a large number of routing paths in HT regions, which results in high NUR.

RNP creates few routing paths in the regions. As a result, in HT regions, the NUR of RNP is lower than the NUR of PRP. On the other hand, in HT regions, RNP creates longer routing paths and incurs higher NUR than TRP. In NHT regions, the NUR of RNP and PRP is comparable, because RNP and PRP employ similar routing techniques for source nodes in NHT regions. The results of experiments conducted indicate that RNP achieves an improved ED, outperforming TRP and PRP.

Figure 5 shows the residual energy of 100 randomly selected sensor nodes at the mission duration of 900 rounds. Sensor nodes number 1 to 50 were located in HT regions, while sensor nodes number 51 to 100 were located in NHT regions.

It is shown in Fig. 5 that for TRP, the residual energy of many sensor nodes in HT regions is above 0.25 J. However, in NHT regions, many of the sensor nodes have depleted their energy. These results confirm that TRP suffers from exhaustive energy consumption in NHT regions. Thus, TRP incurs unbalanced ED.

On the other hand, there is a smaller difference in the residual energy of RNP for HT and NHT regions. The results confirm that RNP outperforms TRP and PRP in terms of ED.

5.3. Safety Period

Figure 6 shows that TRP achieves the longest safety period (SP), while PRP achieves the shortest SP. The results indicate that the privacy preservation in TRP is the strongest. Figure 7 shows the SP of the protocols, as the number of sensor nodes increases. In the experiments, the source nodes were assumed to be at a source-sink distance of 30 hops. The packet rate was fixed at 1 packet/s. This shows that the SP of RNP increases more rapidly than the SP of the remaining protocols.

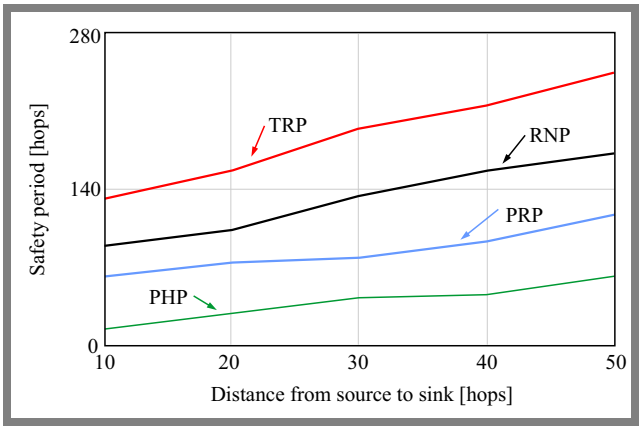

Fig. 6. Privacy preservation of the protocols while varying distance from source to sink.

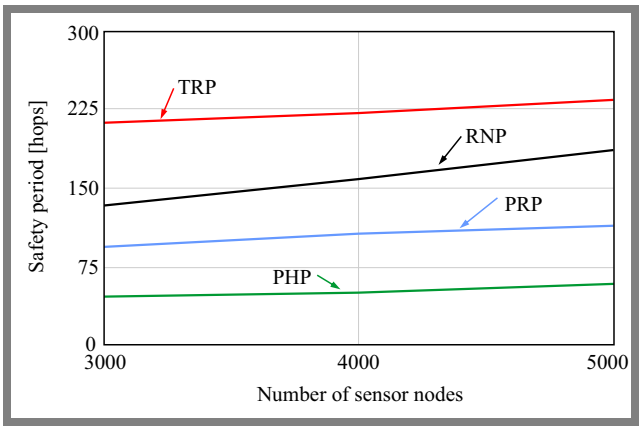

Fig. 7. Privacy preservation of the protocols for different number of sensor nodes.

The rapid increase in SP of RNP is caused by the fact that when the number of nodes in the network is growing, the number of neighboring nodes and candidate phantom nodes increases as well. Consequently, a larger set of phantom nodes is generated and the path diversity increases. Also, RNP employs a bias random number to ensure a dynamic phantom node selection process. The use of the bias random number guarantees that for each successive packet, the second level phantom node is selected from different regions of the WSN domain.

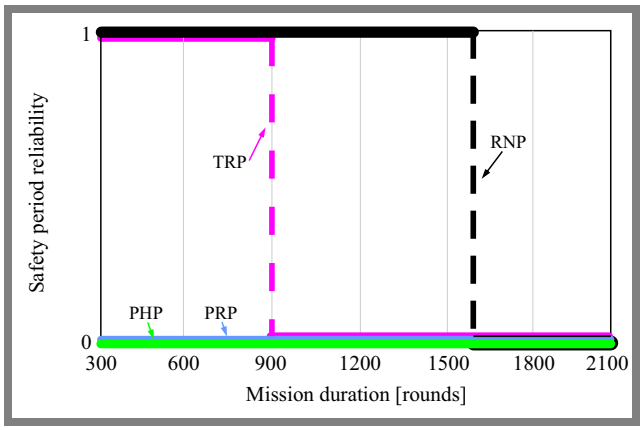

Fig. 8. Safety period reliability of the protocols.

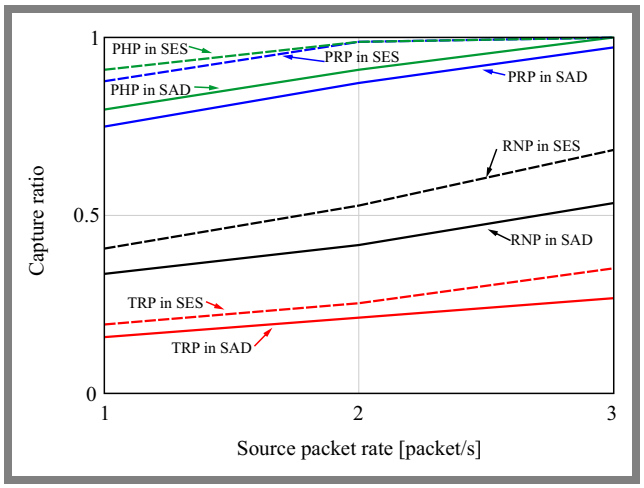

Fig. 9. Capture ratio of the protocols.

Moreover, RNP selects a new phantom node for each source node packet transmission. Therefore, when a larger set of phantom nodes is available, the path diversity and adversary obfuscation effects become more significant, resulting in longer SP.

5.4. Safety Period Reliability

Experiments were performed to observe safety period reliability (SPR) for a mission duration of 2100 rounds. Source nodes were located at a source-sink distance of 40 hops. The source packet generation rate was 1 packet/s. Similarly to [9], the minimum required SP was assumed to be 140 hops.

Figure 8 shows that PHP and PRP do not ensure SPR. This is because PHP and PRP are not able to achieve the required SP. On the other hand, TRP and RNP ensure SPR because they are able to achieve the required SP. However, the SPR of TRP is of the short-term variety. TRP is not capable of ensuring SPR beyond 950 rounds, because it is energy inefficient, as shown in Fig. 3. Only RNP is capable of providing SPR at 1500 rounds.

The results shown in Fig. 8 suggest that RNP provides long-term SPR to outperform TRP. RNP achieves better performance, as it is more energy efficient and achieves better ED than TRP.

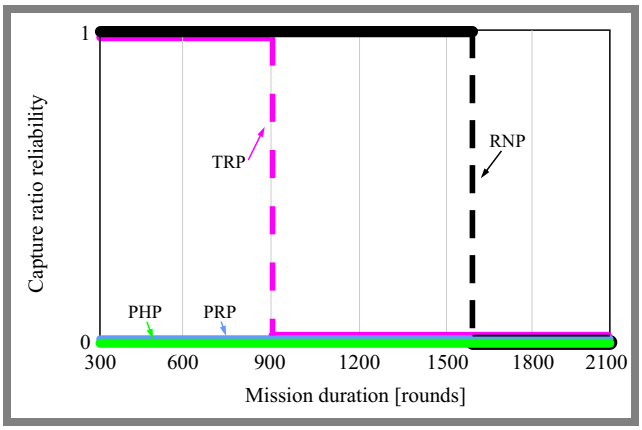

Fig. 10. CRR reliability of the protocols.

5.5. Capture Ratio

Figure 9 shows the capture ratio (CR) when the number and location of source nodes is varying. In the experiments, two scenarios were considered: “scenario east of the sink node” (SES) and “scenario all directions” (SAD). Four source nodes were deployed. In SES, the source nodes were randomly distributed on the east side of the sink node, at a source-sink distance of 25 hops. The distance between the source nodes was between 3 and 6 hops. In SAD, one source node was positioned at a source-sink distance of 25 hops, on the north side of the sink node.

Similarly, one source node was deployed east, west, and south of the sink node. Source packet generation rate was 1 packet/s. It is shown in Fig. 9 that for all the protocols, CR in SES is higher than in SAD. This is because in SES, the generated source packet traffic is concentrated on one side of the network domain. Therefore, the adversary performs a more focused attack and improves its CR. On the other hand, when the source nodes are distributed in SAD, the packet traffic arrives at the sink node from all directions. Consequently, the traffic analyzing attack becomes more complex and CR is reduced.

Figure 9 also shows the differences in the level of CR in SES and SAD. For TRP, the difference is small. This is because TRP distributes fake packet traffic between long diversionary routes. Therefore, it is able to effectively obfuscate the adversary in both SES and SAD. Furthermore, TRP distributes fake packet traffic in the vicinity of the phantom nodes. Therefore, even when the adversary is able to locate the phantom nodes, the success of its attack is hindered and CR improves at a slow rate.

The difference in the level of CR in SES and SAD for PRP is high, because unlike TRP, PRP employs short fake packet routes and it does not distribute fake packet traffic in the vicinity of the phantom nodes. Therefore, when the adversary performs a more focused attack in SES, it is able to improve CR.

The difference in the level of CR in SES and SAD for RNP is lower than for PRP, because RNP is characterized by high path diversity. The use of a dynamic two-level phantom node selection process ensures that packets arrive at the sink node through highly randomized packet routes, both in SES and

Tab. 3. Summary of the observations.

Routing strategy	Impact on the performance of the protocols
Two-level phantom routing	– Increases the path diversity and adversary obfuscation effect in RNP to improve SLP preservation
Fake packet distribution	– Increases the adversary obfuscation effect in TRP to improve SLP preservation – Reduces energy efficiency, which results in short-term SLP protection and reduced SLP reliability
Distributes fake packet traffic in the vicinity of phantom node	– Increases the adversary obfuscation effect in TRP and improves SLP preservation – Reduces energy efficiency, which results in short-term SLP protection and reduced SLP reliability
Distributes fake packet traffic in the vicinity of sink node	– Reduces energy efficiency – Insignificant effect on SLP preservation for PRP
Generates multiple fake packet sources at a time period	– Increases the adversary obfuscation effect in TRP to improve SLP preservation – Reduces energy efficiency, which results in short-term SLP protection and reduced SLP reliability – Unbalanced energy distribution in TRP
Isolates fake source nodes from real source nodes	– Reduces SLP preservation in PRP
Long fake packet routes	– Increases the adversary obfuscation effect in TRP to improve SLP preservation – Reduces energy efficiency, which results in short-term SLP reliability

in SAD. Therefore, the adversary obfuscation effect remains high, even in SES. The routing paths of the baseline PHP are less random. Therefore, CR is significantly higher when the adversary performs a more focused attack in SES.

In addition, Fig. 9 shows that CR in SES and SAD increases at higher source packet rates. This is because at higher data rates the adversary captures an increased number of packets and makes, within a short period of time, significant progress in its attacks. Consequently, CR increases. In particular, when the packet rate is high in SES, the region where the source nodes are located becomes an obvious hotspot region. Therefore, the attack becomes less complex and the adversary improves its CR. In SES, at the rate of 2 packets/s, CR for both PRP and PHP is 100%.

5.6. Capture Ratio Reliability

In the experiments aiming to compute capture ratio reliability (CRR), the value of the parameter was measured for a mission duration of 2100 rounds. A single source node was deployed. The maximum required CR equaled 0.35. Figure 10 shows that PHP and PRP do not ensure the CRR. This is because PHP and PRP are not capable of achieving a CR below 0.35.

On the other hand, TRP and RNP ensure CRR, because they are able to achieve the required CR. However, the CRR of TRP is of the short-term variety, because TRP is energy inefficient and it causes the sensor nodes to deplete their battery power at a fast rate. These results indicate that RNP outperforms TRP and PRP in terms of long-term CRR.

Table 3 summarizes the observations from the experiments. Based on the observations, it is established that RNP offer superior performance. RNP achieves long-term SLP reliability, outperforming TRP and PRP protocols.

6. Conclusion

There are key similarities and differences in the routing strategies of privacy-preserving phantom routing protocols, namely TRP, PRP, and RNP. The routing strategies exert different impacts on SLP reliability, energy distribution, and energy efficiency of the protocols under investigation. Protocols that integrate phantom and fake packet routing strategies improve SLP performance. However, distribution of fake packet traffic degrades the performance of the protocols in terms of SLP reliability and energy efficiency.

The TRP protocol is based on phantom and fake packet routing strategies that achieve short-term SLP reliability. The PRP protocol is more energy efficient than TRP, but it achieves low levels of SLP protection and reliability. On the other hand, RNP is based on a ring distribution of phantom nodes and evades the broadcasting of fake packets. As a result, RNP achieves improved SLP reliability, outperforming TRP and PRP protocols.

As part of future work, SLP protocols for underwater sensor networks (USNs) will be investigated, as previous studies have shown that SLP preservation still remains a challenge in USNs. Studies focusing on SLP are necessary due to the fact that USNs and UIoT systems have become rather powerful and may potentially support various applications, such as maritime security, natural disaster prediction and control, oil and gas exploration, and marine life observation.

References

[1] K.P.R. Krishna and R. Thirumuru, "Energy Efficient and Multi-hop Routing for Constrained Wireless Sensor Networks", *Sustainable Computing: Informatics and Systems*, vol. 38, art. no. 100866, 2023 (https://doi.org/10.1016/j.suscom.2023.100866).

[2] M. Rathee *et al.*, "Ant Colony Optimization Based Quality of Service Aware Energy Balancing Secure Routing Algorithm for Wireless Sensor Networks", *IEEE Transactions on Engineering Management*, vol. 68, no. 1, pp. 170–182, 2021 (https://doi.org/10.1109/TEM.2019.2953889).

[3] N. Wang, J. Fu, J. Li, and B.K. Bhargava, "Source-location Privacy Protection Based on Anonymity Cloud in Wireless Sensor Networks", *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 100–114, 2020 (https://doi.org/10.1109/TIFS.2019.2919388).

[4] A. Chmielowiec, L. Klich, and W. Woś, "Energy Efficient ECC Authenticated Key Exchange Protocol for Star Topology Wireless Sensor Networks", *Journal of Telecommunication and Information Technology*, vol. 1, pp. 1–10, 2024 (https://doi.org/10.26636/jtit.2024.1.1389).

[5] I. Butun, P. Osterberg, and H. Song, "Security of the Internet of Things: Vulnerabilities, Attacks, and Countermeasures", *IEEE Communications Surveys & Tutorials*, vol. 22, no. 1, pp. 616–644, 2020 (https://doi.org/10.1109/COMST.2019.2953364).

[6] N. Hrovatin, A. Tošić, M. Mrissa, and J. Vičič, "A General Purpose Data and Query Privacy Preserving Protocol for Wireless Sensor Networks", *IEEE Transactions on Information Forensics and Security*, vol. 18, pp. 4883–4898, 2023 (https://doi.org/10.1109/TIFS.2023.3300524).

[7] G. Han *et al.*, "CPSLP: A Cloud-based Scheme for Protecting Source Location Privacy in Wireless Sensor Networks Using Multi-sinks", *IEEE Transactions on Vehicular Technology*, vol. 68, no. 3, pp. 2739–2750, 2019 (https://doi.org/10.1109/TVT.2019.2891127).

[8] N. Jan and S. Khan, "Energy-efficient Source Location Privacy Protection for Network Lifetime Maximization against Local Eavesdropper in Wireless Sensor Network (EeSP)", *Emerging Telecommunications Technologies*, vol. 33, no. 2, art. no. 3703, 2022 (https://doi.org/10.1002/ett.3703).

[9] L.C. Mutalemwa and S. Shin, "Novel Approaches to Realize the Reliability of Location Privacy Protocols in Monitoring Wireless Networks", *IEEE Access*, vol. 9, pp. 104820–104836, 2021 (https://doi.org/10.1109/ACCESS.2021.3099499).

[10] L.C. Mutalemwa, "On the Use of Wireless Technologies for Wildlife Monitoring: Wireless Sensor Network Routing Protocols", *Tanzania Journal of Engineering and Technology*, vol. 42, no. 2, pp. 113–133, 2023 (https://doi.org/10.52339/tjet.v42i2.837).

[11] G. Han, R. Xia, H. Wang, and A. Li, "Source Location Privacy Protection Algorithm Based on Polyhedral Phantom Routing in Underwater Acoustic Sensor Networks", *IEEE Internet of Things Journal*, pp. 8459–8472, 2023 (https://doi.org/10.1109/JIOT.2023.3318567).

[12] S.A.H. Mohsan, A. Mazinani, N.Q.H. Othman, and H. Amjad, "Towards the Internet of Underwater Things: A Comprehensive Survey", *Earth Science Informatics*, vol. 15, no. 2, pp. 735–764, 2022 (https://doi.org/10.1007/s12145-021-00762-8).

[13] M. Singh and M.P. Singh, "Congestion Avoidance with Source Location Privacy Using Octopus-based Dynamic Routing Protocol in WSN", *Wireless Networks*, vol. 29, pp. 729–748, 2023 (https://doi.org/10.1007/s11276-022-03165-9).

[14] G. Kumar *et al.*, "Dynamic Routing Approach for Enhancing Source Location Privacy in Wireless Sensor Networks", *Wireless Networks*, vol. 29, pp. 2591–607, 2023 (https://doi.org/10.1007/s11276-023-03322-8).

[15] M. Kamarei, A. Patooghy, A. Alsharif, and V. Hakami, "SiMple: A Unified Single and Multi-path Routing Algorithm for Wireless Sensor Networks with Source Location Privacy", *IEEE Access*, vol. 8, pp. 33818–33829, 2020 (https://doi.org/10.1109/ACCESS.2020.2972354).

[16] S. Chakraborty, N.K. Goyal, S. Mahapatra, and S. Soh, "Minimal Path-based Reliability Model for Wireless Sensor Networks with Multistate Nodes", *IEEE Transactions on Reliability*, vol. 69, no. 1, pp. 382–400, 2020 (https://doi.org/10.1109/TR.2019.2954894).

[17] P. Mishra *et al.*, "Reliability Evaluation of a Wireless Sensor Network in Terms of Network Delay and Transmission Probability for IoT Applications", *Contemporary Mathematics*, pp. 309–325, 2024 (https://doi.org/10.37256/cm.5120242906).

[18] F. Dan *et al.*, "An Accuracy-aware Energy-efficient Multipath Routing Algorithm for WSNs", *Sensor*s, vol. 24, no. 1, art. no. 285, 2024 (https://doi.org/10.3390/s24010285).

[19] N. Sonnappa and K. Muniyegowda, "Privacy-aware Secured Discrete Framework in Wireless Sensor Network", *International Journal of Electrical and Computer Engineering*, vol. 14, no. 1, pp. 75–85, 2024 (https://doi.org/10.11591/ijece.v14i1.pp75-85).

[20] X. Xu, J. Tang, and H. Xiang, "Data Transmission Reliability Analysis of Wireless Sensor Networks for Social Network Optimization", *Journal of Sensors*, vol. 2022, art. no. 3842722, 2022 (https://doi.org/10.1155/2022/3842722).

[21] A.M. Alabdali, N. Gharaei, and A.A. Mashat, "A Framework for Energy-efficient Clustering with Utilizing Wireless Energy Balancer", IEEE Access, vol. 9, pp. 117823–117831, 2021 (https://doi.org/10.1109/ACCESS.2021.3107230).

[22] L.C. Mutalemwa and S. Shin, "Energy Balancing and Source Node Privacy Protection in Event Monitoring Wireless Networks", *2021 International Conference on Information Networking (ICOIN)*, Jeju Island, South Korea, 2021 (https://doi.org/10.1109/ICOIN50884.2021.9333901).

[23] C.-M. Yu *et al.*, "BRATRA: Balanced Routing Algorithm with Transmission Range Adjustment for Energy Efficiency and Utilization Balance in WSNs", *IEEE Internet of Things Journal*, vol. 10, no. 2, pp. 1096–1111, 2023 (https://doi.org/10.1109/JIOT.2022.3206316).

[24] J. Long, M. Dong, K. Ota, and A. Liu, "Achieving Source Location Privacy and Network Lifetime Maximization Through Tree-based Diversionary Routing in Wireless Sensor Networks", *IEEE Access*, vol. 2, pp. 633–651, 2014 (https://doi.org/10.1109/ACCESS.2014.2332817).

[25] H. Wang *et al.*, "A Probabilistic Source Location Privacy Protection Scheme in Wireless Sensor Networks", *IEEE Transaction on Vehicular Technology*, vol. 68, no. 6, pp. 5917–5927, 2019 (https://doi.org/10.1109/TVT.2019.2909505).

[26] L.C. Mutalemwa and S. Shin, "Secure Routing Protocols for Source Node Privacy Protection in Multi-hop Communication Wireless Networks", *Energies*, vol. 13, no. 2, art. no. 292, 2020 (https://doi.org/10.3390/en13020292).

[27] C. Ozturk, Y. Zhang, and W. Trappe, "Source-location Privacy in Energy-constrained Sensor Network Routing", *Proceedings of the 2nd ACM Workshop on Security of Ad Hoc and Sensor Networks*, pp. 88–93, 2004 (https://doi.org/10.1145/1029102.1029117).

[28] Q. Wang, J. Zhan, X. Ouyang, and Y. Ren, "SPS and DPS: Two New Grid-based Source Location Privacy Protection Schemes in Wireless Sensor Networks", *Sensors*, vol. 19, no. 9, art. no. 2074, 2019 (https://doi.org/10.3390/s19092074).

[29] Z. Xiong *et al.*, "A Ring-based Routing Scheme for Distributed Energy Resources Management in IIoT", *IEEE Access*, vol. 8, pp. 167490–167503, 2020 (https://doi.org/10.1109/ACCESS.2020.3023260).

[30] T. Hussain *et al.*, "Improving Source Location Privacy in Social Internet of Things Using a Hybrid Phantom Routing Technique", *Computers & Security*, vol. 123, art. no. 102917, 2022 (https://doi.org/10.1016/j.cose.2022.102917).

[31] C. Gu, M. Bradbury, and A. Jhumka, "Phantom Walkabouts: A Customisable Source Location Privacy Aware Routing Protocol for Wireless Sensor Networks", *Concurrency and Computation, Practice and Experience*, vol. 31, no. 20, 2019 (https://doi.org/10.1002/cpe.5304).

[32] G. Han *et al.*, "A Dynamic Multipath Scheme for Protecting Source-location Privacy Using Multiple Sinks in WSNs Intended for IIoT", *IEEE Transactions on Industrial Informatics*, vol. 16, no. 8, pp. 5527–5538, 2020 (https://doi.org/10.1109/TII.2019.2953937).

[33] N. Jan, A. Al-Bayatti, N. Alalwan, and A. Alzahrani, "An Enhanced Source Location Privacy Based on Data Dissemination in Wireless Sensor Networks (DeLP)", *Sensors*, vol. 19, no. 9, art. no. 2050, 2019, DOI: (https://doi.org/10.3390/s19092050).

[34] L. Mutalemwa, "Location Privacy Protection and Coverage Hole Effects in Event Monitoring Wireless Networks for Internet of Things Applications", *Journal of ICT Systems*, vol. 1, no. 2, pp. 52–70, 2023 (https://doi.org/10.56279/jicts.v1i2.50).

[35] P. Kamat, Y. Zhang, W. Trappe, and C. Ozturk, "Enhancing Source-location Privacy in Sensor Network Routing", *25th IEEE International Conference on Distributed Computing Systems (ICDCS'05)*, Columbus, USA, 2005 (https://doi.org/10.1109/ICDCS.2005.31).

[36] C. Gu, M. Bradbury, J. Kirton, and A. Jhumka, "A Decision Theoretic Framework for Selecting Source Location Privacy Aware Routing Protocols in Wireless Sensor Networks", *Future Generation Computer Systems*, vol. 87, pp. 514–526, 2018 (https://doi.org/10.1016/j.future.2018.01.046).

[37] W.-K. Yun and S.-J. Yoo, "Q-Learning-based Data-aggregation-aware Energy-efficient Routing Protocol for Wireless Sensor Networks", *IEEE Access*, vol. 9, pp. 10737–10750, 2021 (https://doi.org/10.1109/ACCESS.2021.3051360).

[38] N.R. Patel, S. Kumar, and S.K. Singh, "Energy and Collision Aware WSN Routing Protocol for Sustainable and Intelligent IoT Applications", *IEEE Sensors Journal*, vol. 21, no. 22, pp. 25282–25292, 2021 (https://doi.org/10.1109/JSEN.2021.3076192).

[39] D. Thomas, R. Shankaran, M.A. Orgun, and S.C. Mukhopadhyay, "SEC 2: A Secure and Energy Efficient Barrier Coverage Scheduling for WSN-Based IoT Applications", *IEEE Transactions on Green Communications Networking*, vol. 5, no. 2, pp. 622–634, 2021 (https://doi.org/10.1109/TGCN.2021.3067606).

[40] Y. Liu *et al.*, "An Improved Energy-efficient Routing Protocol for Wireless Sensor Networks", *Sensors*, vol. 19, no. 20, art. no. 4579, 2019 (https://doi.org/10.3390/s19204579).

[41] G. Sudha and C. Tharini, "Trust-based Clustering and Best Route Selection Strategy for Energy Efficient Wireless Sensor Networks", *Automatika*, vol. 64, no. 3, pp. 634–641, 2023 (https://doi.org/10.1080/00051144.2023.2208462).

[42] V. Narayan, A.K. Daniel, and P. Chaturvedi, "E-FEERP: Enhanced Fuzzy Based Energy Efficient Routing Protocol for Wireless Sensor Network", *Wireless Personal Communications*, vol. 131, pp. 371–398, 2023 (https://doi.org/10.1007/s11277-023-10434-z).

[43] T.M. Behera *et al.*, "CH Selection via Adaptive Threshold Design Aligned on Network Energy", *IEEE Sensors Journal*, vol. 21, no. 6, pp. 8491–8500, 2021 (https://doi.org/10.1109/JSEN.2021.3051451).

[44] S. K. Chaurasiya *et al.*, "An Energy-efficient Hybrid Clustering Technique (EEHCT) for IoT-Based Multilevel Heterogeneous Wireless Sensor Networks", *IEEE Access*, vol. 11, pp. 25941–25958, 2023 (https://doi.org/10.1109/ACCESS.2023.3254594).

[45] Z. Qu *et al.*, "An Energy-efficient Dynamic Clustering Protocol for Event Monitoring in Large-scale WSN", *IEEE Sensors Journal*, vol. 21, no. 20, pp. 23614–23625, 2021 (https://doi.org/10.1109/JSEN.2021.3103384).

[46] Q. We *et al.*, "A Cluster-based Energy Optimization Algorithm in Wireless Sensor Networks with Mobile Sink", *Sensors*, vol. 21, no. 7, art. no. 2523, 2021 (https://doi.org/10.3390/s21072523).

[47] T. Shafique *et al.*, "A Review of Energy Hole Mitigating Techniques in Multi-hop Many to One Communication and its Significance in IoT Oriented Smart City Infrastructure", *IEEE Access*, vol. 11, pp. 121340–121367, 2023 (https://doi.org/10.1109/ACCESS.2023.3327311).

[48] X. Liu and J. Wu, "A Method for Energy Balance and Data Transmission Optimal Routing in Wireless Sensor Networks", *Sensors*, vol. 19, no. 13, art. no. 3017, 2019 (https://doi.org/10.3390/s19133017).

Lilian C. Mutalemwa, Ph.D.
Faculty of Science, Technology and Environmental Studies
https://orcid.org/0000-0003-4342-5562
E-mail: lilian.mutalemwa@out.ac.tz
The Open University of Tanzania, Dar es Salaam, Tanzania
https://out.ac.tz

Information for Authors

Journal of Telecommunications and Information Technology (JTIT) is published quarterly since 2000. It comprises original contributions, dealing with a wide range of topics related to telecommunications and information technology. **All papers are subject to peer review.** Topics presented in the JTIT report primary and/or experimental research results, which advance the base of scientific and technological knowledge about telecommunications and information technology.

JTIT is dedicated to publishing research results which advance the level of current research or add to the understanding of problems related to modulation and signal design, wireless communications, optical communications and photonic systems, voice communications devices, image and signal processing, transmission systems, network architecture, coding and communication theory, as well as information technology.

We encourage submissions from a diverse range of authors from across all countries and backgrounds.

Manuscript

Latex files are preferred and Editorial Office provides a style to prepare the material along with the documentation. We also accept Microsoft Word and PDF files. A typical article is 10 pages long (approximately 6,000 words) and must include the following contents:

- Authors' names and affiliations in the following format:
  First name and surname (last name), academic title,
  Position held,
  ORCID number,
  E-mail address from the University's domain,
  Faculty and name of the University,
  Link to University website.
- Abstract (150-200 words). The abstract should contain statement of the problem, assumptions and methodology, results and conclusion or discussion on the importance of the results. Abstracts must not include mathematical expressions or bibliographic references.
- Keywords related to the content of the article. About four keywords or phrases in alphabetical order should be used, separated by commas.
- The content of the article in a typical structure, i.e.: introduction, related work, conducted research, conclusions, references.

Figures, Tables and Photos

Together with the article, please send files with graphics with the highest resolution available, 150 dpi or more in bitmap resolution (jpg, png) and vector (cdr, svg, ps, pdf) formats are welcomed.

References

We use four main citation styles for a journal article, for an Internet article, for a conference paper, and for a book. Below are examples of citations. In each item, the DOI number or link to the PDF of the cited article should be provided.

[1] R.K. Meyers and A.H. Desoky, "An implementation of the blowfish cryptosystem", *2008 IEEE International Symposium on Signal Processing and Information Technology*, 2008 ((https://doi.org/10.1109/IS-SPIT.2008.4775664).

[2] K. Nowicki and T. Uhl, *Ethernet End-to-End*, 1st ed. Germany, Shaker-Publisher, 2008 (ISBN: 978383832271404).

[3] C. Shorten and T.M. Khoshgoftaar, "A survey on image data augment-tation for deep learning", *Journal of Big Data*, vol. 6, no. 1, pp. 1–48, 2019 (https://doi.org/10.1186/s40537-019-0197-0).

[4] S. Wong *et al*., "Traffic forecasting using vehicle-to-vehicle communication", *3rd Annual Conference on Learning for Dynamics and Control*, pp. 917–929, 2021 (https://arxiv.org/pdf/2104.05528).

Submission

The paper with full PDF version and anonymous PDF version for the blind review process should be submitted on the JTIT website https://www.jtit.pl/jtit/about/submissions.

Reviewing Process

The article is initially approved by the Editor-In-Chief and if the decision is positive, is then sent to the reviewers. Depending on the subject of the article, it takes few weeks. In the next step, reviews are showed to authors who have 2 weeks to correct the article. Finally, the corrected text can be re-presented to the reviewer for reevaluation, which will take another 2 weeks.

As a result, after about 3 months, we are able to send the text for publication in the upcoming issue of JTIT.

When the reviews are inconsistent, additional corrections are necessary, or the reviewer expects additional verification because the corrections ordered by the author are insufficient or additional problems arise, the review of the article may be extended by another month or more.

Editorial Work

Positively reviewed and corrected article is next prepared by the editorial office for publication. At the end of this process the author receives an copyedited version for approval.

Licensing

Manuscript submitted to JTIT should not be published or simultaneously submitted for publication elsewhere. By submitting a manuscript the author grants license to the National Institute of Telecommunications, for the use of the paper in the fields of exploitation: reproducing and fixing the paper, distributing the paper by means of introduction to trade, letting for use or rental of the original or copies, and distributing the paper by means of public exhibition, screening, presentation and broadcast as well as rebroadcast, and making the paper publicly available in such a manner that anyone could access it at a place and time selected thereby, or by making it available in a way not allowing selection of time or place, including by means of Internet or other networks.

Ghostwriting Declaration

We require formal declaration that the process of writing the paper was not influenced by any third party. In the article, all the contributions of other people are clearly indicated. The theories presented, methods used, analysis and research, as well as the copyrights to the drawings, photographs and other figures belong to the authors or are clearly credited in the text. The author must also indicate whether his work has received financial support and if the realization of the whole project was possible thanks to the permission and cooperation with scientific institutions, associations and others.

Other Information

- The JTIT being an Open Access Journal (OAJ) has no article processing charges (APCs). The published articles can be downloaded freely without payment.
- JTIT supports open access and using continuous publishing "publish-as-you-go" scheme. This means that we no longer wait to accumulate several articles into a quarterly issue before publication. Rather, articles are continuously added to current issues after acceptation. Publish-as-you-go reduces publication lag for our authors, and make the newest research available quickly. After completing the review process, an article is published online in the current issue with DOI registration. When the issue period ends, a new issue is activated. So accepted articles are published without waiting for the quarterly issue end.

National Institute of Telecommunications

Editorial Office

National Institute
of Telecommunications
Szachowa st 1
04-894 Warsaw, Poland
https://www.gov.pl/web/instytut-lacznosci

phone +48 22 512 81 83
fax +48 22 512 84 00

e-mail: journal@jtit.pl
www.jtit.pl

JOURNAL OF TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY 3/2024