Infocommunications Journal

A PUBLICATION OF THE SCIENTIFIC ASSOCIATION FOR INFOCOMMUNICATIONS (HTE)

December 2023 Volume XV Number 4 ISSN 2061-2079

Technically Co-Sponsored by

IEEE ComSoc™ IEEE Communications Society

hte

IEEE HUNGARY SECTION

Editorial Board

Editor-in-Chief: PÁL VARGA, Budapest University of Technology and Economics (BME), Hungary
Associate Editor-in-Chief: LÁSZLÓ BACSÁRDI, Budapest University of Technology and Economics (BME), Hungary
Associate Editor-in-Chief: JÓZSEF BÍRÓ, Budapest University of Technology and Economics (BME), Hungary

Area Editor – Quantum Communications: ESZTER UDVARY, Budapest University of Technology and Economics (BME), Hungary
Area Editor – Cognitive Infocommunications: PÉTER BARANYI, University of Pannonia, Veszprém, Hungary
Area Editor – Radio Communications: LAJOS NAGY, Budapest University of Technology and Economics (BME), Hungary
Area Editor – Networks and Security: GERGELY BICZÓK, Budapest University of Technology and Economics (BME), Hungary
Area Editor – Neural Speech Technology: TAMÁS GÁBOR CSAPÓ, Budapest University of Technology and Economics (BME), Hungary

JAVIER ARACIL, Universidad Autónoma de Madrid, Spain
LUIGI ATZORI, University of Cagliari, Italy
STEFANO BREGNI, Politecnico di Milano, Italy
VESNA CRNOJEVIÇ-BENGIN, University of Novi Sad, Serbia
KÁROLY FARKAS, Budapest University of Technology and Economics (BME), Hungary
VIKTORIA FODOR, KTH, Royal Institute of Technology, Stockholm, Sweden
JAIME GALÁN-JIMÉNEZ, University of Extremadura, Spain
EROL GELENBE, Institute of Theoretical and Applied Informatics Polish Academy of Sciences, Gliwice, Poland
ISTVÁN GÓDOR, Ericsson Hungary Ltd., Budapest, Hungary
CHRISTIAN GÜTL, Graz University of Technology, Austria
ANDRÁS HAJDU, University of Debrecen, Hungary
LAJOS HANZO, University of Southampton, UK
THOMAS HEISTRACHER, Salzburg University of Applied Sciences, Austria
ATTILA HILT, Nokia Networks, Budapest, Hungary
JUKKA HUHTAMÄKI, Tampere University of Technology, Finland
SÁNDOR IMRE, Budapest University of Technology and Economics (BME), Hungary
ANDRZEJ JAJSZCZYK, AGH University of Science and Technology, Krakow, Poland
FRANTISEK JAKAB, Technical University Kosice, Slovakia
GÁBOR JÁRÓ, Nokia Networks, Budapest, Hungary
MARTIN KLIMO, University of Zilina, Slovakia
ANDREY KOUCHERYAVY, St. Petersburg State University of Telecommunications, Russia
LEVENTE KOVÁCS, Óbuda University, Budapest, Hungary
MAJA MATIJASEVIC, University of Zagreb, Croatia
OSCAR MAYORA, FBK, Trento, Italy
MIKLÓS MOLNÁR, University of Montpellier, France
SZILVIA NAGY, Széchenyi István University of Győr, Hungary
PÉTER ODRY, VTS Subotica, Serbia
JAUDELICE DE OLIVEIRA, Drexel University, Philadelphia, USA
MICHAL PIORO, Warsaw University of Technology, Poland
ROBERTO SARACCO, Trento Rise, Italy
GHEORGHE SEBESTYÉN, Technical University Cluj-Napoca, Romania
BURKHARD STILLER, University of Zürich, Switzerland
CSABA A. SZABÓ, Budapest University of Technology and Economics (BME), Hungary
GÉZA SZABÓ, Ericsson Hungary Ltd., Budapest, Hungary
LÁSZLÓ ZSOLT SZABÓ, Sapientia University, Tirgu Mures, Romania
TAMÁS SZIRÁNYI, Institute for Computer Science and Control, Budapest, Hungary
JÁNOS SZTRIK, University of Debrecen, Hungary
DAMLA TURGUT, University of Central Florida, USA
SCOTT VALCOURT, Roux Institute, Northeastern University, Boston, USA
JÓZSEF VARGA, Nokia Bell Labs, Budapest, Hungary
ROLLAND VIDA, Budapest University of Technology and Economics (BME), Hungary
JINSONG WU, Bell Labs Shanghai, China
KE XIONG, Beijing Jiaotong University, China
GERGELY ZÁRUBA, University of Texas at Arlington, USA

Indexing information

Infocommunications Journal is covered by Inspec, Compendex and Scopus.
Infocommunications Journal is also included in the Thomson Reuters – Web of ScienceTM Core Collection, Emerging Sources Citation Index (ESCI)

Infocommunications Journal

Technically co-sponsored by IEEE Communications Society and IEEE Hungary Section

Supporters

FERENC VÁGUJHELYI – president, Scientific Association for Infocommunications (HTE)

The publication was produced with the support of the Hungarian Academy of Sciences and the NMHH

National Media and Infocommunications Authority • Hungary

Editorial Office (Subscription and Advertisements)**:**
Scientific Association for Infocommunications
H-1051 Budapest, Bajcsy-Zsilinszky str. 12, Room: 502
Phone: +36 1 353 1027 • E-mail: info@hte.hu • Web: www.hte.hu

Articles can be sent also to the following address:
Budapest University of Technology and Economics
Department of Telecommunications and Media Informatics
Phone: +36 1 463 4189 • E-mail: pvarga@tmit.bme.hu

Subscription rates for foreign subscribers: 4 issues 10.000 HUF + postage

Publisher: PÉTER NAGY

HU ISSN 2061-2079 • Layout: PLAZMA DS • Printed by: FOM Media

Key trends in applied ICT technologies for 2024

Pal Varga

WE arrived at a many-ways turbulent era at the end of 2023. Focusing on the science and technology, the evolution of Generative AI (GenAI) is at the forefront of the changes. Previously a subject of theoretical discourse, GenAI is now real and practical. This shift is not just about technological advancement but has become strategically important in various areas, from science to education and business to company operations. Enterprises are expected to implement GenAI in real-world applications, moving from a heavy emphasis on training and infrastructure costs to a more nuanced consideration of inference and operational expenses.

Another significant development is the collaboration between hyperscalers and AI models in data analytics. This partnership is destined to revolutionize how data is processed and used, with a shift toward real-time fine-tuning and the ability to adapt to current data. In current applied scientific results, we already see that such solutions drive advancements in AI applications across various industries, leading to improvements in speed, accuracy, and cost. The creation of a powerful, responsive ecosystem through this collaboration will enable the development of large-scale, complex models to address an array of industry-specific use cases.

In the cybersecurity domain, a notable trend is the convergence of IT and security teams. As the boundaries between these two traditionally separate areas blur, a more unified approach to managing digital threats is emerging. This integration is driven by the rapid advancement of technology and the evolving landscape of security risks.

Quantum computing is also set to make notable strides, particularly in the realm of quantum communications. This includes the adoption of post-quantum cryptography (PQC) to protect data against future quantum attacks and the emergence of quantum networking. Although RSA is expected to remain a robust encryption method for a while, it faces increased scrutiny and potential vulnerability from quantum computing advancements. Researchers are actively exploring strategies to leverage quantum computing to break RSA and ECC, including the possibility of hybrid attacks combining quantum and traditional computing methods. The advancements in the quantum field will be crucial for data security and processing, attracting significant research and investment, particularly from sectors with high data security demands.

Still, in the middle of these advancements, human skills remain indispensable, especially in the uptake of AI. There is a growing focus on closing skills gaps through reskilling and upskilling initiatives. The interplay between human expertise and advanced technologies will be a defining feature of the upcoming period, as the Industry 5.0 initiative also seeks to balance the benefits of automation with the special understanding that only human professionals can provide.

Having these in mind, let's briefly see the December 2023 issue of Infocommunications Journal.

In her paper, Eszter Udvary focuses on integrating Quantum Key Distribution (QKD) with high-speed optical data transmission using Dense Wavelength Division Multiplexing (DWDM) in optical fibers. This integration aims to enhance network security in a cost-effective manner. The paper explores different scenarios for optimal channel allocation and the necessary bandwidth separation between classical and quantum channels.

The paper by Ammar Al-Adhami, Yasir Al-Adhami, and Taha A. Elwi explores the integration of a 3D antenna array with solar cells for self-powered applications in modern wireless communication networks. It focuses on a cubical antenna array geometry integrated with a solar panel to achieve a selfpowered node. The design aims to enhance the performance of Multi-Input Multi-Output (MIMO) systems while considering energy efficiency.

In their current paper, Beatrix Koltai, András Gazdag, and Gergely Ács proposes a novel anomaly detection mechanism for the CAN-bus in vehicles. This mechanism integrates timeseries forecasting and signal correlation analysis to enhance detection accuracy in onboard Intrusion Detection Systems (IDS). The approach predicts sets of correlated signals collectively and identifies anomalies when the combined prediction error exceeds a predefined threshold. The paper demonstrates that this method significantly outperforms existing solutions, offering more accurate detection of a broader range of attacks with minimal delay, making it highly effective for vehicular network security.

Péter Orosz, Balázs Nagy, and Pál Varga discuss the changes in DDoS attack profiles observed in real data center infrastructures in their paper, and provide a comprehensive survey of state-of-the-art detection methods tailored to these recent attacks. The authors emphasize the significance of novel attack methods and tools, the increasing frequency, extent, and complexity of attacks, and the emergence of multi-vector attacks combining L3-L7 profiles.

But this is just the briefing – let's see the papers themselves.

Pal Varga is the Head of Department of Telecommunications and Media Informatics at the Budapest University of Technology and Economics. His main research interests include communication systems, Cyber-Physical Systems and Industrial Internet of Things, network traffic analysis, end-to-end QoS and SLA issues – for which he is keen to apply hardware acceleration and artificial intelligence, machine learning techniques as well. Besides being a member of HTE, he is a senior member of IEEE, where he is active both in the IEEE ComSoc (Communication Society) and IEEE IES (Industrial Electronics Society) communities. He is Editorial Board member in many journals, and the Editor-in-Chief of the Infocommunications Journal.

Integration of QKD Channels to Classical High-speed Optical Communication Networks

Eszter Udvary, *Member, IEEE*

***Abstract*—Integrating Quantum Key Distribution service with classical high-speed optical data transmission using a dense wavelength division multiplexing technique in a fiber is a cost-effective solution to improve the network's security. In this multichannel system, several noise sources degrade the quality of the quantum channel. The dominant degradation effect is determined by modeling in different cases. Optical filtering cannot decrease spontaneous Raman Scattering caused by the classical optical channels. So this nonlinear optical effect is investigated in detail with different system parameter setups. The optimal channel allocation and the required bandgap between the classical and quantum channels are determined.**

***Index Terms*—Optical fiber communication, Quantum communication, Quantum key distribution, Wavelength division multiplexing, Optical fiber networks, Optical fibers, Raman scattering**

I. INTRODUCTION

QUANTUM key distribution (QKD) enables more secure, quantum-safe information transmission and sharing using quantum mechanical effects. Suppose quantum communication is applied to share and generate a secret symmetric key between two parties. In that case, it is theoretically impossible for an eavesdropper to learn about the key. The technique ensures the safety of the transmitted data from all kinds of hacking and attacks, including quantum computers [1, 22]. Fiber-based QKD experiments are typically established via dedicated dark optical fibers for point-to-point applications [10, 19].

The quantum internet vision requires a full-day working, cost-effective solution, and the QKD services must be extended from point-to-point links to network configurations. Using dedicated optical fibers only to the quantum channel has a high cost and scalability problem. That means additional fibers are required for each service next to the existing infrastructure, which may need many cable sections to be swapped or built. QKD would be transmitted through the existing fiber network together with classical optical communication signals to avoid extra investments in the optical infrastructure [2] for more appropriate operation. Network integration of QKD is essential for QKD's future deployment when the quantum key distribution link successfully generates secret keys over standard single-mode fiber with co-propagating classical high-speed optical channels. Otherwise, QKD will never be a reliable and effective widespread solution outside research laboratories.

QKD integration into the existing optical fiber network infrastructure using Wavelength Division Multiplexing (WDM) is possible [3]. The main challenge of this approach is the different optical power of the different wavelength channels, as both QKD and classical communication channels are coupled to the same fiber. QKD protocols typically require a launch power of less than 1 nW. In contrast, classical signals are launched with a 1-10 mW power per channel. A small fraction of the classical signal falling to the QKD receiver is enough to decrease the quality of the quantum communication, increasing the Quantum Bit Error Rate (QBER) to a value where key extraction is impossible. In addition, the quantum communication signal cannot pass through an optical amplifier due to the no-cloning theorem. Such integration requires careful network planning [4].

Several papers related to this topic can be found in the literature, but each focuses only on its own QKD solution, primarily examining the interaction between QKD and classical channels using experimental methods. This limits the system complexity and the number of classical channels in the published studies. Therefore, the type of QKD technology most suitable for working with classical systems and which network architecture and technology can integrate the QKD channel into the network are still questions.

The typical secure communication scenarios require different expected reach lengths and serve different groups of customers. A short reach is usually between 1 and 10 km; it is mainly needed for the financial sector, typically in the access domain [8]. The 10-50 km medium reach supports the aggregation domain for public services and municipalities. Links over 50 km are called long reach, necessary for large, multinational, and industrial corporations. Different architecture solutions and different WDM techniques are required for different scenarios.

So, the feasibility of transmitting QKD with classical communication channels through the same optical fiber by employing WDM technologies at telecom wavelengths and a wide variety of scenarios is essential.

Manuscript received 25 Augustus 2023.

The work has been supported by the National Research Development and Innovation Office (NKFIH) through the OTKA Grant K 142845.

E. Udvary is with the Department of Networked Systems and Services, Budapest University of Technology and Economics, Budapest, Hungary (e-mail: udvary.eszter@vik.bme.hu).

She part time works on the ELKH-BME Information Systems Research Group, Budapest, Hungary

DOI: 10.36244/ICJ.2023.4.1

This paper overviews the applicable WDM technologies in the different optical network segments and the possible QKD integration solutions. Next, the paper presents the degradation effects of crosstalk from high-speed classical channels. The presented mathematical model was used to calculate the crosstalk originating from optical nonlinearities, leakages, and optical noises. Finally, nonlinear Raman scattering is modeled in detail, as optical filtering cannot decrease this effect.

II. WAVELENGTH ALLOCATION

The QKD protocols need a quantum channel and two directional signaling classical channels to perform error correction and privacy amplification on the QKD channel. So, traditionally, three optical fibers are required between Alice (QKD transmitter) and Bob (QKD receiver). As a first step, the signaling and synchronization channels can be multiplexed with the classical data communication channels. Sometimes, there are some limitations for the route and the distance of the service channel related to the quantum channel. So, applying the same fiber, or at least the same optical cable, is suggested.

Exploring the most efficient channel wavelength allocation for establishing an effective and reliable communication channel for QKD and encrypted data transmission in optical communication networks is essential. I suggest the following QKD integration solutions.

A. Separated QKD and classical DWDM networks

In this approach, wavelength multiplexed quantum channels are directed to a separate fiber, and the high-speed DWDM classical channels are transmitted via other fibers in the same optical cable. The solution is easy and performs well, as there is no shared transmission media for classical and quantum channels. It will be helpful if the QKD technology is widespread and many QKD channels are required in the network.

Uncoupled multicore fibers (MCFs) are also suitable for integrating the quantum channel into the classical communication network. The QKD performance may improve as the inter-core crosstalk is lower [11]. However, multicore fibers are not widely used in current telecommunication systems.

This QKD integration solution now requires additional fibers for quantum channels next to the existing infrastructure.

B. Coarse Wavelength Division Multiplexing (CWDM)

The classical and quantum channels coexist based on a two-directional Coarse Wavelength Division Multiplexing standard with a 20 nm channel grid (Fig. 1). The topology may include more optical filters to decrease the crosstalk and defend the quantum channel. The advantage is the lower crosstalk from the classical channels, but CWDM is less and less used in current and especially future optical communication networks.

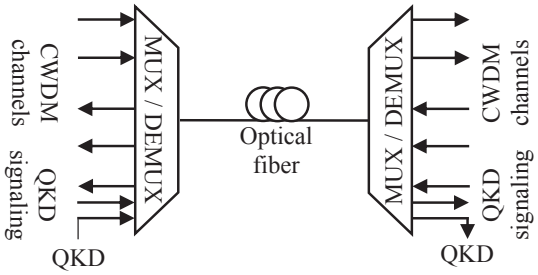


Fig. 1. CWDM Scenario, duplex communication

Nowadays, this solution may be applied in a metropolitan network, but the spread of this multicarrier standard is decreasing as the lack of an optical amplifier limits it.

C. Dense Wavelength Division Multiplexing (DWDM)

The classical and quantum channels coexist based on the Dense Wavelength Division Multiplexing standard with 0.4 or 0.8 nm channel grid (Fig. 2). The DWDM technique is applied in the core and mainly in metropolitan networks, which uses the C-band between 1530 and 1565 nm wavelength. If the QKD channel is placed in the C-band, the QKD link has low attenuation to achieve a maximum key rate and link length. However, the quantum channel is then spectrally close to the classical channels and strongly affected by stimulated Raman scattering (SRS) and four-wave mixing (FWM) in a network. A guard band between classical and quantum channels may be applied to decrease the degradation effect of FWM.

It is essential to determine the optimal position of the quantum channel in the DWDM channel allocation (lower wavelength, upper wavelength, or middle wavelength) and the required bandwidth of the guardband between the quantum and classical channels. We expected in advance that all the classical channels should be placed at wavelengths longer than the quantum channel since the Spontaneous anti-Stokes Raman scattering (SASRS) is typically weaker than the Spontaneous Stokes Raman scattering (SSRS, SRS in the following).

Additionally, amplifying quantum signals is impossible, as the quantum state can not be cloned based on the non-cloning theorem. The integrated QKD has to bypass optical amplifiers, requiring a separate multiplexer. The optical noise originating from the amplified spontaneous emission (ASE) of an erbium-doped optical amplifier (EDFA) still affects the quantum channel as it has components at the wavelength of the quantum channel. An optical notch filter can be applied to decrease the ASE noise in the QKD channel.

An optical isolator saves Alice's transmitter from the strong reflected or the backward signals. It guarantees the stable operation of the QKD transmitter.

Before Bob's receiver, an optical bandpass filter in the QKD channel is necessary for out-of-band noise reduction.

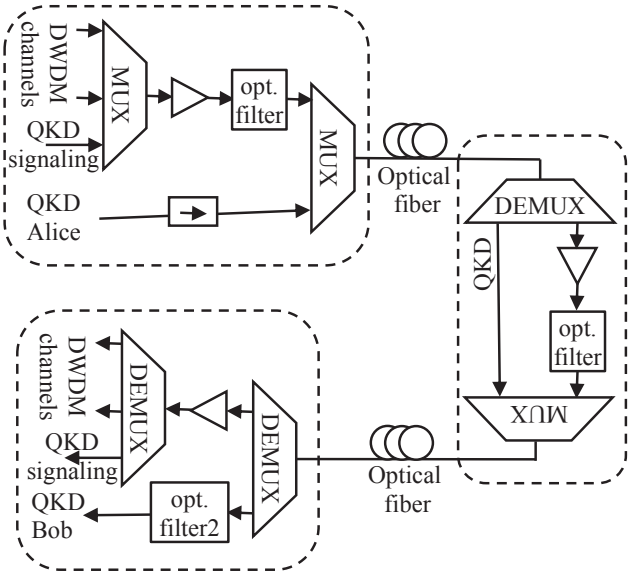


Fig. 2. One direction DWDM Scenario, simplex communication on a fiber. Duplex communication requires two fibers.
MUX: DWDM multiplexer, DEMUX: DWDM demultiplexer, opt. filter: optical notch filter for filtering out the ASE noise from the quantum DWDM channel, opt. filter2: bandpass filer for quantum channel

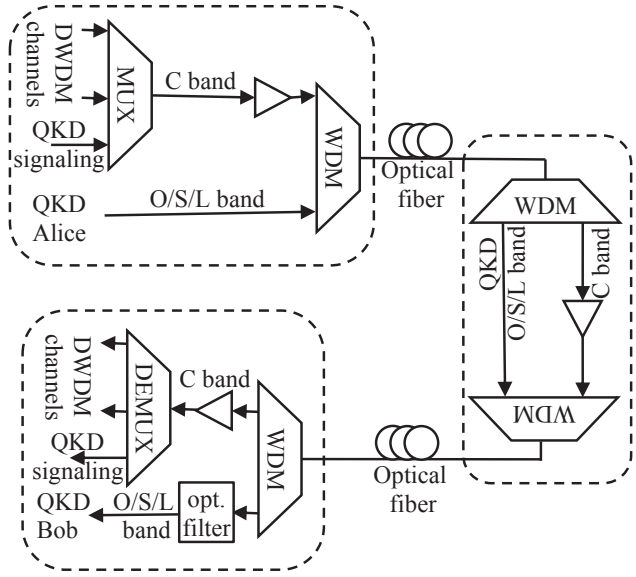


Fig. 3. Hybrid system, O/L/S-band QKD Scenario.
MUX: DWDM multiplexer, DEMUX: DWDM demultiplexer, opt. filter: bandpass filer for quantum channel, WDM: WDM coupler for separate and add the two optical bands

D. Hybrid system

In this solution, the quantum channel wavelength is placed in the O-band at 1310 nm (between 1260 and 1360 nm), and DWDM classical channels are set in the C-band (Fig. 3). In this approach, the separation is more significant than 200 nm between the QKD and C-band classical channels. By applying an optical bandpass filter on the receiver side, the effect of the crosstalk can be reduced.

The integrated QKD channel can also be placed in the S-band or L-band when the separation is smaller, but the same advantages can be realized.

The main disadvantage of this approach is that the QKD optical link loss is higher, which limits the quality of the quantum channel. However, in the current practical metropolitan optical networks, the excess loss due to fiber connections, patches, routing devices, or other passive components dominates the total loss of a link. In such an environment, the quantum channel in the O-band may be helpful since the optical loss approaches that of the C-band channels, but the noise is reduced.

E. QKD in Optical Access Networks

The optical access network is a special segment of communication networks. The access networks are short-span, point-to-multipoint, passive optical networks with a time or wavelength division multiplexing approach connecting to the end user [7]. A particular segment of optical networks consists of the optical links and systems used in mobile networks, where security is a critical issue [9]. Security and privacy are a challenge for time division multiplexing passive optical networks (TDM-PONs), where the same wavelength channel in a shared fiber link is used for broadcasting the downstream signal to all users [12].

Although time division multiplexing is used for resource sharing between classic users, integrating the QKD channel in time multiplexing is not practical and challenging. Typically, wavelength multiplexing can be used here as well. We have to apply a different wavelength allocation strategy than other network segments. The typical architecture of the access network is tree topology. A passive network requires power dividers, and the high insertion loss of the dividers dominates the optical loss over the attenuation of the optical fiber.

Different optical access standards use different wavelengths, minimum and maximum transmitter powers, and data rates to transmit upstream and downstream signals. The optical powers launched into the fiber and its center wavelength strongly influence the background noise level in the QKD channel. We assume the typical PON reaches a maximum of 20 km, of which 15 km is dedicated to the feeder fiber. QKD channels should be ideally placed spectrally far away from classical signals to avoid strong interactions and degradation. Therefore, wavelength ranges below 1310 nm are suitable.

F. Review of the published demonstrations

Table 1 summarizes the recent research published for different QKD protocols implemented in telecommunication networks. The main concept of each scenario is the same as the principles I presented in the previous subsections. But the quality-improving additional elements that I propose are not included, and usually only a connection between two points is examined, without the intermediate node. Nevertheless, it is clear that the number of channels and the link length are limited, and it is difficult to find a comprehensive simulation and experimental study in the literature.

TABLE I
REFERENCE WORKS

ref. number	QKD	Scenario	distance	channels	e/s
[12]	BB84	DWDM	80km	40	s
[13]	CV	DWDM	50km	5	s
[14]	CV	hybrid	60km	5-40	s
[15]	CV	DWDM	10km	100	e
[16]	BB84	hybrid	50km	32	e
[17]	BB84	DWDM	50km	1-4	e
[18]	BB84	hybrid	70km	9	e

channels.: number of classical channels
e/s: experiments/simulations

III. NOISE CONTRIBUTIONS

Crosstalk between channels in multichannel optical networks and optical nonlinear phenomena occurring in optical waveguides are not new or unknown. They also happen in purely classical optical networks, and we know their description [3]. However, the specialties of the quantum channel have repeatedly drawn attention to their quality degradation effects and require a specific examination of the phenomena.

The impact of classical channels on the quantum channel has been investigated. Two main sources can contribute to the noise photons in the quantum channel. The leakage of photons from classical channels is due to the imperfect isolation of WDM multiplexers and demultiplexers. The in-band noise photons are generated in optical fibers from nonlinear processes, such as four-wave mixing (FWM) and spontaneous Raman scattering (SRS). In the DWDM system, the in-band ASE photons generated by optical amplifiers are also considered. These optical noises heavily determine the quality of the QKD channel.

Suitable, cascaded optical filters can easily increase channel isolation, but the applied optical filters also increase the insertion loss.

However, nonlinear processes can create photons at the same wavelength as the quantum signal, which cannot be spectrally filtered. Depending on the wavelength of the classical and quantum channels, Raman scattering and four-wave mixing are the dominant nonlinear processes. FWM is a narrow-band effect that can be minimized by carefully choosing the quantum channel's wavelength. However, the fixed channel grid of the DWDM standard limits this opportunity. Raman noise has a broad spectrum; classical channels in the C-band create a Raman noise spectrum covering the whole band. Temporal filtering may help to reduce the impact of noise photons at the quantum channel's band. The quality of Raman noise rejection depends on the time-bandwidth product of the quantum signal and on how tight the time and spectral filtering can be implemented.

A. Leakage from the classical channels

Multiplexing and demultiplexing (MUX and DEMUX) provide filtering between the QKD and conventional signal [18]. However, photons from the classical signals will be at the quantum channel because of the finite isolation of the applied MUX/DEMUX. Two types can be distinguished. The leakage noise component can be in the quantum channel's wavelength band, called in-band noise. This happens because the signal arriving on the classical channel also has components in the QKD band, which cannot be separated with a filter after multiplexing since they fall within the wavelength range of the quantum channel.

The origin of the out-of-band leakage noise is the component at the classical channel wavelength range, attenuated by the non-infinitive isolation of the demultiplexer. That is, the attenuated signal level of the classical channel is still comparable to the unattenuated signal of the quantum channel, which is detected by a sensitive QKD receiver in the entire wavelength range. The out-of-band leakage noise component can be further reduced using spectral filters at the receiver's side.

Typically, the value of the in-band leakage noise is lower than that of the out-of-band leakage noise; because the original value of the in-band leakage noise is lower, and the multiplexer filters it out.

The typical adjacent channel isolation of a CWDM DEMUX or MUX is higher than 30dB, the non-adjacent channel isolation is higher than 45 dB, and the insertion loss is less than 3 dB. For DWDM DEMUX and MUX, these parameters are typically a bit worse, but they can also reach these values using modern technology. The average leakage photon number can be calculated from the leakage power, which is determined by the classical channel power and the isolation.

$$\langle \mathrm{N_{leakage}} \rangle = \frac{\mathrm{P_{leakage}} \cdot \Delta t}{\mathrm{h} \cdot \nu} = \frac{\mathrm{P}_{fiber\_out} \cdot \mathrm{a_{DEMUX\_ISO}} \cdot \Delta t}{\mathrm{h} \cdot \nu} \quad (1)$$

where h is the Planck constant, v is the frequency of the classical optical channel, a_{DEMUX_ISO} is the demultiplexer isolation, Δt is the time window, P_{fiber_out} is the classical channel power at the output of the fiber.

B. Amplified Spontaneous Emission (ASE)

The ASE is generated from a typical optical amplifier (EDFA or SOA). It can be considered a broadband noise source with a flat spectral power density within the spectral bandwidth of the quantum channel, as it has a broad bandwidth of tens of nm.

The ASE noise of the optical amplifier is characterized by the noise figure (NF)

$$NF = \frac{1+2 \cdot n_{sp}(G-1)}{G} \quad (2)$$

where n_{sp} is the spontaneous emission factor ($n_{sp} \geq 1$), G is the gain of the optical amplifier, and the two orthogonal polarization modes are described by factor 2. In the high gain range, the noise figure can be calculated directly, $NF \approx 2 \cdot n_{sp}^2$.

ASE has a broad bandwidth and components in the quantum channel. So, it will contribute to in-band noise after the demultiplexer, but the isolation of the demultiplexer attenuates it. The in-band ASE photon number per spatiotemporal mode originating from the transmitter side optical amplifier, at the input of the quantum receiver, is given by

$$\langle N_{ASE} \rangle = 2 \cdot n_{sp} \cdot (G-1) \cdot a_{MUX\_ISO} \cdot a_{fiber} \cdot a_{DEMUX} \quad (3)$$

where a_{MUX_ISO} is the multiplexer isolation, a_{fiber} is the attenuation of the optical fiber between the optical amplifier and the QKD detector, a_{DEMUX} is the attenuation of the demultiplexer.

In practical cases, the value of n_{sp} is often unknown, but optical noise power can be measured. The mean number of the noise photon for a narrow optical band can be calculated from the measured optical noise power (P_{opt}):

$$\langle \mathrm{N} \rangle = \frac{P_{opt} \cdot \Delta t}{\frac{h \cdot \nu}{\lambda}} \tag{4}$$

and

$$\langle N_{ASE} \rangle = \frac{P_{ASE} \cdot \lambda \cdot \Delta t}{h \cdot \nu} \cdot a_{Notch} \cdot a_{MUX\_ISO} \cdot a_{fiber} \cdot a_{DEMUX} \tag{5}$$

where a_{Notch} is the attenuation of the optical notch filter, P_{ASE} is the ASE optical noise power.

The laser diode has similar broadband optical noise, which can be modeled with the same approach. However, the noise level is lower in the in-band leakage noise analysis.

C. Optical nonlinearities

As the strong classical signals propagate along the optical fiber, various nonlinear optical processes will generate noise photons at different wavelengths. If the wavelength of the noise photons falls into the wavelength band of the quantum channel, it is in-band noise that can not be filtered at the side of the receiver.

1) Spontaneous Stokes Raman scattering (SRS)

The level of the SRS effect depends on the propagation direction of the classical and quantum channels. If the signal counter-propagates to the quantum signal, the Raman noise level is higher than the co-propagating channels because of the isotropic nature of Raman scattering and the higher power at the receiver.

The SRS noise power within $\Delta\lambda$ optical bandwidth at the output of the optical fiber is given by

$$P_{SRS} = P_{fiber_{out}} \cdot \beta_{SRS} \cdot z \cdot \Delta\lambda = P_{fiber\_in} \cdot \beta_{SRS} \cdot z \cdot \Delta\lambda \cdot a_{fiber} \tag{6}$$

where β_{SRS} is the spontaneous Raman scattering coefficient depending on the optical wavelengths, P_{fiber_in} is the classical channel power at the input of the fiber, P_{fiber_out} is the classical channel power at the output of the fiber, z is the fiber length, and a_{fiber} is the loss of the optical link.

The average SRS photon number can be calculated from the SRS noise power within a bandwidth of $\Delta\lambda$. However, the mode number corresponding to the $\Delta\lambda$ optical bandwidth and a time window (Δt) must be considered [3].

$$N_{mode} = |\Delta\nu \cdot \Delta t| = \frac{c}{\lambda^2} \Delta\lambda \cdot \Delta t \tag{7}$$

where c is the speed of light in a vacuum.

The average SRS photon number at the output of the optical fiber can be calculated.

$$\langle N_{SRS\_fiber\_out} \rangle = \frac{P_{SRS} \cdot \Delta t}{h \cdot \nu \cdot N_{mode}} \tag{8}$$

The SRS signal goes via the DEMUX; the insertion loss of the DEMUX attenuates the average SRS photon number at the input of the optical receiver.

$$\langle N_{SRS} \rangle = \frac{P_{SRS} \cdot \Delta t}{h \cdot \nu \cdot N_{mode}} \cdot a_{DEMUX} \tag{9}$$

So, the average SRS photon number at the output of the optical fiber is given by

$$\langle N_{SRS} \rangle = \frac{\lambda^3}{h \cdot c^2} \cdot P_{fiber_{out}} \cdot \beta_{SRS}(\lambda) \cdot z \cdot a_{DEMUX} \tag{10}$$

2) Four Wave Mixing (FWM)

FWM is a third-order nonlinear optical process, which can be observed during two or more optical signal propagation through an optical fiber and originates from the Kerr effect.

Three optical channels at frequencies f_x, f_y, f_z mix through the fiber's third-order susceptibility, generating a new optical carrier at $f_{xyz}=f_x+f_y-f_z$ frequency.

The effect requires phase matching. So, the level is high at sort fiber length or zero-dispersion optical fiber. However, it is weaker in long-distance standard single-mode fiber due to the chromatic dispersion. The effect can be decreased by applying a guard band between the classical and quantum channels. Also, FWM can be minimized by channel wavelength optimization in a flex grid system. Polarization multiplexing also suppresses this effect. Based on it, 2-2 channel guardbands were applied, and FWM was neglected in this investigation.

3) Cross-Phase Modulation (XPM)

Cross-phase modulation also originates from the Kerr effect in optical fibers, where intensity modulation of one optical carrier can modulate the phases of other transmitting optical carriers in the same fiber. Phase noise is not a direct source of performance degradation in intensity modulation and direct detection systems. However, together with the dispersion, it already has a negative effect. In addition, the quality-deteriorating effect is also direct in a coherent optical system.

When integrating a quantum channel, its effect strongly depends on the applied QKD technology. It is more significant for Phase-modulated quantum information, especially in the case of CV-QKD [5]. In this case, the QKD structure is similar to classical coherent detection; but the noise level must be kept sufficiently low so that the effect of quantum phenomena can be observed. Therefore, its effect must be examined independently of the other noise components.

IV. SIMULATION RESULTS

Optical scattering and crosstalk between classical and quantum channels are critical in the one-direction DWDM scenario (Chapter II.C). I investigated the noise and crosstalk from the classical channels in this topology because it is the more challenging scenario. During the detailed noise analysis, a simplified system structure is used, where we implement classical and quantum communication between two points, and the classical channels are amplified (Fig. 4).

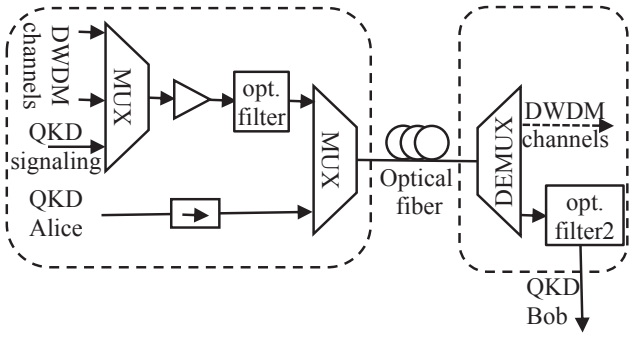


Fig. 4. One direction DWDM simulation setup. Simplex communication is used on a fiber, Duplex communication requires two fibers.
MUX: DWDM multiplexer, DEMUX: DWDM demultiplexer, opt. filter: optical notch filter for filtering out the ASE noise from the quantum DWDM channel, opt. filter2: bandpass filer for quantum channel

This topology includes every important effect, and the main conclusions can be drawn from it. The level of crosstalk from the classical to the quantum channel, practically the number of noise photons, is determined. The out-of-band classical signal leakage noise, the nonlinear Raman scattering, the ASE noise of the optical amplifier, and the laser noise were taken into account; I presented all of them in detail in the previous chapter. Filtering the QKD channel from DWDM channels is performed by both the multiplexer and the demultiplexer [18]. So, the in-band leakage noise was neglected, as the multiplexer filters it out, and the value is lower than the out-of-band leakage noise.

Table II summarizes the applied simulation parameters.

TABLE II
SIMULATION PARAMETERS

Symbol	Parameter name	Value
P	Classical transmitter output optical power	0dBm
RIN	Laser relative intensity noise	-160dBc/Hz
a_{MUX}	Multiplexer insertion loss	3dB
a_{MUX_ISO}	Multiplexer isolation	40dB
a_{filter}	Notch optical filter insertion loss	1dB
NF	Optical amplifier noise figure	5dB
α_{fiber}	Fiber attenuation	0.2dB/km
β_{SRS}	Nonlinear Raman scattering coefficient	$2 \cdot 10^{-9}$/km/nm
a_{DEMUX}	Demultiplexer insertion loss	3dB
a_{DEMUX_ISO}	Demultiplexer isolation + optical bandpass filter	85dB
a_{filter}	Optical bandpass filter insertion loss	1dB
$\Delta\lambda$	Channel bandwidth	0.4nm
Δt	Gating time window	1ns

Fig. 5 presents the noise photons in the quantum channel. Similar to practical optical systems, in this first simulation, the optical gain is adaptive, compensating for the attenuation of the channel, and therefore, has a variable value depending on the length of the connection. The dominant noise type depends on the system parameters and varies over the communication link.

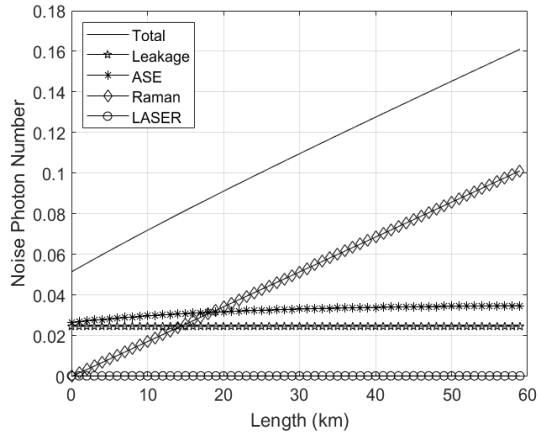


Fig. 5. Noise contributions versus optical fiber. Adaptive optical amplification, it compensates the optical channel loss.

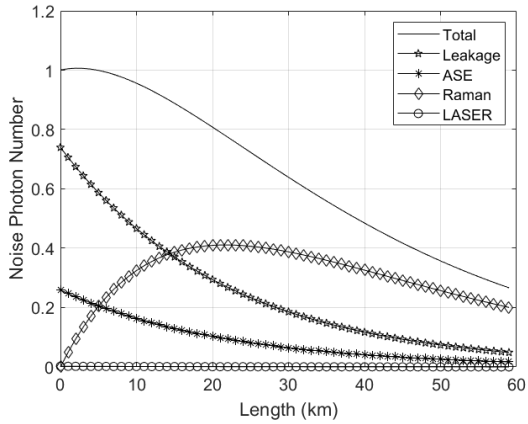


Fig. 6. Noise contributions versus optical fiber. Constant optical amplification.

The situation changes somewhat if the optical gain is not adaptive, but has a constant value regardless of the connection length (Fig. 6). The level of different types of noise changes differently along the connection length. Therefore, the dominant noise type in the total noise level may differ from section to section. But typically, the Raman scattering exceeds the value of the other noise components for longer lengths.

Correctly setting the parameters makes it possible to make one type of noise dominant in the total noise value. The broad-spectrum noise level of the optical amplifier and laser source can be reduced with better isolation of the multiplexer and better suppression of the optical notch filter (Fig. 7). The isolation of the multiplexer and the characteristic of the proposed optical notch filter have similar effects. The isolation of the multiplexer decreases the in-band leakage from the classical channels, including also the ASE noise photons at the input of the second multiplexer. The suggested optical notch filter decreases the ASE noise in the QKD band, but it simultaneously decreases all the in-band noises at the optical amplifier output. Fig. 7. represents the consequences of these elements, which mainly decrease the ASE noise in the QKD channel as in-band leakage noise is much less than ASE noise.

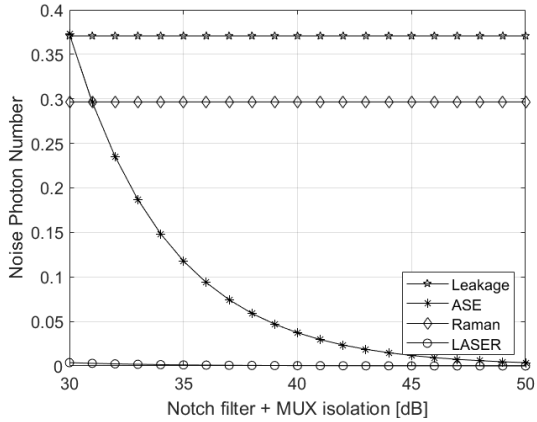


Fig. 7. Noise contributions versus the isolation of the multiplexer and suppression of the optical notch filter. Fiber length is 40km.

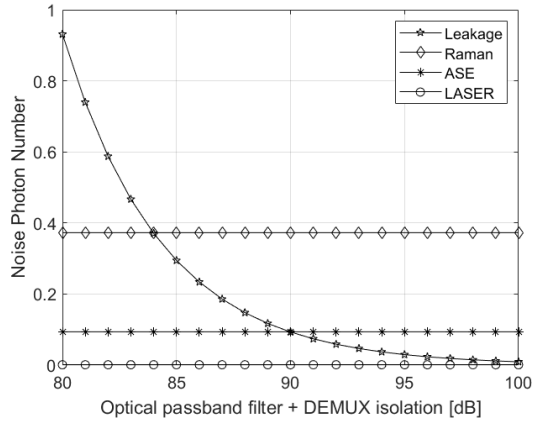


Fig. 8. Noise contributions versus the demultiplexer isolation and the suppression of the optical bandpass filter. Fiber length is 20km.

The figure emphasizes the importance of the proposed notch filter. A MUX's typical adjacent channel isolation is higher than 30 dB. This is why the graph starts from 30dB (MUX isolation) and presents the additional effect of the proposed notch filter.

By improving the demultiplexer isolation and the suppression of the optical bandpass filter, out-band leakage noise (typically leakage) can be reduced (Fig. 8). The demultiplexer and the proposed bandpass optical filter for the QKD band have similar effects. They decrease the out-of-band classical channels. 30dB typically MUX's adjacent channel isolation is insufficient; the bandpass filter can never be omitted. Therefore, the figure concentrates on the essential operating range above 80 dB.

However, optical filtering cannot remove the noise components due to nonlinear phenomena, even using an ideal filter. Therefore, the nonlinear Raman scattering effect is investigated in detail.

Actually, the Raman scattering coefficient is not constant depending on the wavelengths of the applied quantum and classical channels. For more accurate calculations, the wavelength of the channels must be known and the exact coefficient determined. Based on the measurement results found in the literature, this coefficient is a non-symmetrical V-shaped function versus the wavelength [21]. In addition, the coefficient is different in the case of Stokes and anti-Stokes, as in the case of co- and counter propagation. If there are several classical channels in the DWDM system, then the effect of their Raman scattering is added. It is worth using at least 2 guard bands in addition to the quantum channel to reduce the impact of four-wave mixing.

Fig. 9 represents the Raman scattering-originated noise photon number versus the position of the quantum channel. One quantum channel and 93 classical channels were considered in the C Telecommunication band, with 2-2 channel wide guardbands before and after the quantum channel. The ITU DWDM standard determines the IDs of the DWDM channels in this paper. Namely, ID 1 means optical carrier with a wavelength of 1567.13nm, ID 96 means 1529.16nm, and the channel spacing is 0.4nm. The optimal quantum channel position is the DWDM channel with ID 61-62. The Raman noise photon number can be halved by carefully choosing the wavelengths. It can improve the connection length or the quantum key rate.

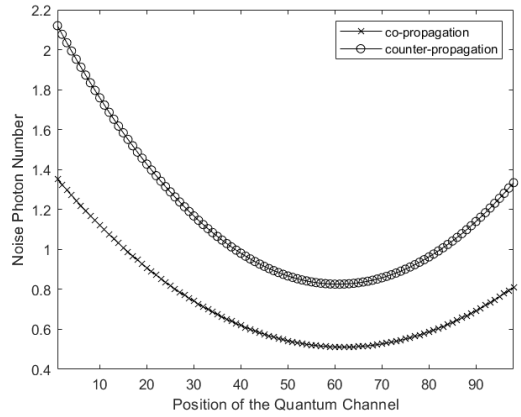


Fig. 9. Noise Photon Number originating from the Nonlinear Raman Scattering versus the DWDM ID of the quantum channel. One quantum channel with 2-2 guardbands, in case of co- and counter propagation, 40km fiber length

V. CONCLUSION

This paper deals with the sensitive QKD channel integration to classical high-speed optical communication networks for improving network security.

The first part of the paper overviews the typical technologies in the different optical network segments. It also suggests effective QKD integration solutions, which differ from segment to segment. The main ideas of the system structures are presented using the same optical fiber for QKD and classical channels. The paper recommends additional elements to maintain the quality of the QKD channel, such as optical bandpass filter, optical notch filter, and optical isolator.

The second part of the paper presents the optical noise calculations and simulations. In the proposed multichannel system, several noise sources degrade the quality of the quantum channel. The dominant degradation effect was determined by modeling in adaptive and fixed optical amplification depending on the system parameters. The results

show that the broadband optical noise of the laser source has a negligible effect compared to other noise sources. The broad-spectrum noise level of the optical amplifier and laser source can be reduced with better isolation of the multiplexer and better suppression of the optical notch filter. Improving the demultiplexer isolation and suppressing the optical bandpass filter can reduce out-band leakage noise (typically leakage). Optical filtering cannot decrease spontaneous Raman Scattering caused by the classical optical channels. So this nonlinear optical effect was investigated in detail with different system parameter setups. The optimal channel allocation and the required bandgap between the classical and quantum channels were determined.

VI. ACKNOWLEDGMENT

The author would like to thank Ms. Flóra Kárpát, Mr. Bálint Pákai, and Mr. András Sóki, undergraduate students from the Budapest University of Technology and Economics who are involved in the published work.

REFERENCES

[1] Chandra. D., et al., "On the Road to Quantum Communications", *Infocommunications Journal*, Vol. XIV, No 3, September 2022, pp. 2–8., **DOI**: 10.36244/ICJ.2022.3.1

[2] Ciurana, A. et al., "Entanglement Distribution in Optical Networks," *IEEE Journal of Selected Topics in Quantum Electronics*, Vol. 21, No 3, May – June 2015. **DOI**: 10.1109/JSTQE.2014.2367241

[3] Bing Qi et al., "Feasibility of quantum key distribution through a dense wavelength division multiplexing network," *New Journal of Physics*, Vol. 12, 2010. https://doi.org/10.1088/1367-2630/12/10/103042

[4] Martelli, P. et al., "Integration of QKD in WDM networks" *International Conference on Optical Network Design and Modeling (ONDM)*, 28 June 2021 – 01 July 2021. **DOI**: 10.23919/ONDM51796.2021.9492394

[5] Karinou, F. et al., "Toward the Integration of CV Quantum Key Distribution in Deployed Optical Networks" *IEEE Photonics Technology Letters*, Vol. 30, No 7, pp. 650–653, April 2018. **DOI**: 10.1109/LPT.2018.2810334

[6] David Kobor and Eszter Udvary, "Optimisation of Optical Network for Continuous-Variable Quantum Key Distribution by Means of Simulation", *Infocommunications Journal*, Vol. XII, No 2, July 2020, pp. 18–24. **DOI**: 10.36244/ICJ.2020.2.3

[7] Fröhlich, B. et al., "Quantum secured gigabit optical access networks," *Scientific Reports*, Vol. 5, No 18121, pp. 1–7, 2016. **DOI**: 10.1038/srep18121

[8] Pham, T. A. et al., "Quantum Key Distribution over Hybrid Fiber-Wireless System for Mobile Networks" *International Symposium on Information and Communication Technology*, pp. 236–241, December 2019. **DOI**: 10.1145/3368926.3369670

[9] Zavitsanos, D. et al., "On the QKD Integration in Converged Fiber/Wireless Topologies for Secured, Low-Latency 5G/B5G Fronthaul," *Applied Sciences*, Vol. 10, No 15, pp. 1–21, 2020. **DOI**: 10.3390/app10155193

[10] Zhang Q. et al., "Large scale quantum key distribution: challenges and solutions [Invited]," *Optics Express*, Vol. 26, No 18, pp. 24 260–24 273, 2018. **DOI**: 10.1364/OE.26.024260

[11] Gaglianoa, A. et al., "Integration of the QKD Layer in Fibre Networks Using Multicore Fibres", *IEEE International Conference on Transparent Optical Networks (ICTON)*, Bucharest, Romania, 02-06 July 2023. **DOI**: 10.1109/ICTON59386.2023.10207357

[12] Aleksic, S. et al., "Quantum Key Distribution Over Optical Access Networks" *IEEE European Conference on Network and Optical Communications (NOC)*, 10-12 July 2013, Graz, Austria. **DOI**: 10.1109/NOC-OCI.2013.6582861

[13] Lin, R. and Chen, J., "Modeling and Minimizing Spontaneous Raman Scattering for QKD Secured DWDM Networks" *IEEE Communications Letters*, Vol. 25, No. 12, pp. 3918–3921, Dec. 2021. **DOI**: 10.1109/LCOMM.2021.3116755

[14] Du, S. Tian, Y. and Li, Y. "Impact of Four-Wave-Mixing Noise from Dense Wavelength-Division-Multiplexing Systems on Entangled-State Continuous-Variable Quantum key Distribution", *Physical Review Applied*, Vol. 14, No 2, August 2020. **DOI**: 10.1103/PhysRevApplied.14.024013

[15] Vorontsova, I. et al., "Theoretical analysis of quantum key distribution systems when integrated with a DWDM optical transport network", *Journal of the Optical Society of America B*, Vol. 40, No 1, pp. 63–71, Jan. 2023. **DOI**: 10.1364/JOSAB.469933

[16] Eriksson, T.A., et al., "Wavelength division multiplexing of continuous variable quantum key distribution and 18.3 Tbit/s data channels", *Communications Physics*, Vol. 2, No 9, 2019. **DOI**: 10.1038/s42005-018-0105-5

[17] Wang, L.-J. et al., "Long-distance copropagation of quantum key distribution and terabit classical optical data channels", *Physical Review A*, Vol. 95, No 012301, 2017. **DOI**: 10.1103/PhysRevA.95.012301

[18] Patel, K. A. et al., "Quantum key distribution for 10 Gb/s dense wavelength division multiplexing networks", *Applied Physics Letters*, Vol. 104, No 5, Febr. 2014. **DOI**: 10.1063/1.4864398

[19] Pistoia, M., et al., "Paving the way toward 800 Gbps quantum-secured optical channel deployment in mission-critical environments", *Quantum Science and Technology*, Vol. 8, N 3, 035015, 2023. **DOI**: 10.1088/2058-9565/acd1a8

[20] Czermann, M. et al., "Demonstrating BB84 Quantum Key Distribution in the Physical Layer of an Optical Fiber Based System", *Infocommunications Journal*, Vol. XIII, No 3, September 2021, pp. 45–55., **DOI**: 10.36244/ICJ.2021.3.5

[21] Rui Lin and Jiajia Chen, "Modeling and Minimizing Spontaneous Raman Scattering for QKD Secured DWDM Networks", *IEEE Communication Letters*, Vol. 25, No 12, 2021, pp. 3918–3921, **DOI**: 10.1109/LCOMM.2021.3116755

[22] Eraerds P., et al., "Quantum key distribution and 1 Gbps data encryption over a single fibre", *New Journal of Physics*, Vol. 12, June 2010, 063027, **DOI**: 10.1088/1367-2630/12/6/063027

[23] András Mihály, and László Bacsárdi, "Optical transmittance based store and forward routing in satellite networks", *Infocommunications Journal*, Vol. XV, No 2, June 2023, pp. 8–13., **DOI**: 10.36244/ICJ.2023.2.2

Eszter Udvary received her Ph.D. degree in electrical engineering from the Budapest University of Technology and Economics (BME), Budapest, Hungary, in 2009. She is currently an Associate Professor at BME, Mobile Communication and Quantum Technologies Laboratory. Dr. Udvary's research interests are in the broad area of optical communications, including microwave photonics, optical access network, visible light communication, and quantum communication.

A 3D Antenna Array based Solar Cell Integration for Modern MIMO Systems

Ammar Al-Adhami, Yasir Al-Adhami, and Taha A. Elwi

***Abstract*—In this work, a design of a 3D array antenna based solar panel integration for self-powered applications in modern wireless communication network. Such array configuration is proposed for Multi Input Multi Output (MIMO) applications. The proposed antenna array is structured as a cubical geometry integrated to a solar panel. Such integration is employed to achieve a self-powered node. The proposed antenna is designed to perform an excellent size reduction at sub-6GHz frequency bands when designed with the aid of the metamaterial (MTM) structures. The antenna performance is enhanced by Moore fractal geometry based on electromagnetic band gap (EBG) defects in the ground plane. The proposed antenna is found to provide a moderate gain at 3.6GHz, 3.9GHz, and 4.9GHz. The antenna array shows low coupling effects, below -20dB, due the array configuration in a cubical shape arrangement. The proposed work is extended to evaluate effectively the bit error rate (BER) and channel capacity (CC), when the proposed antenna system is located in real world communication environments. Therefore, QPSK modulation scheme is considered to suite the applications of 5G systems. The amount of the harvested solar energy is considered the limit to manage the total signal to noise ratio (SNR) in that is applied to the proposed communication scheme. This work is considered for practical aspect issues that is related with amount for the generated power from such integration with solar panel and the total generated SNR. Finally, the comparison between measured and simulated data reveals an excellent agreement between them.**

***Index Terms*—SNR, MIMO, self-powered, MTM.**

I. INTRODUCTION

Fifth generation (5G) was developed technology of mobile broadband communication that is planned to use after 2020 [1]. The frequency spectrum of 5G systems is used two ranges, sub-6 GHz that is including (3 GHz -to- 6 GHz) and above-6 GHz according to study of physical properties [2]. Different 5G antenna designs were proposed for 5G specifications [3]. The authors attempted to produce the exhaustive review that covered vary important antenna design and its parameters and performances enhancement such as gain, bandwidth and radiation patterns [4]. A monopole antenna array design based on fractal geometry was proposed for MIMO applications [4]. In [5], an antenna structure was consistent of a traditional monopole antenna; this antenna was attached to a fractal patch and mounted on an FR4 substrate to provide broadside radiation patterns as the result of fractal geometry addition with excellent bandwidth.

Manuscript received June 19, 2022; revised July 26, 2023. Date of publication 2023. Date of current version August 1, 2023.

First author is with Al-karkh University of Science, Baghdad, Iraq. (e-mail: ammarisam@kus.edu.iq).

Second author is with the Institute of Technology - Baghdad, Middle Technical University, Baghdad, Iraq. (e-mail: yasir_isam@mtu.edu.iq).

Third author is with the International Applied and Theoretical Research Center (IATRC), Baghdad Quarter, Iraq. (e-mail: taelwi82@gmail.com).

DOI: 10.36244/ICJ.2023.4.2

A multiband fractal at sub-6 GHz microstrip antenna utilizing circle and triangle fractals for different wireless applications to cover three frequency bands of (1.8-2.9GHz, 3.4–4.6 GHz, and 5–5.6 GHz) was proposed in [6]. by loading Koch fractal geometry on edges of a square microstrip patch antenna to enhance the antenna bandwidth and gain by suppressing the surface wave effects [7]. A multiband reconfigurable antenna for sub-6GHz 5G wireless communication networks was designed to control the frequency resonance using pin diodes [8] at two frequency bands.

In another concept, the introduction of the solar panel to the modern antenna designs has become the most urgent for the self-powered wireless systems [9]. For this, MTM based antennas were integrated with solar panels in different publications [10]-[13]. For instance, the authors in [10] proposed an antenna design based on MTM for gain enhancements by integration their antenna design with the solar panel as a substrate. Furthermore, another design was suggested in [11] to integrate the solar panel as a reflector for a Vivaldi antenna based corrugated grating patch with MTM. In [12], a design of MTM structure based on plasmonic antenna array for modern applications; in that design, the authors printed their antenna array on a flexible solar panel. The design of an antenna based MTM in [13] was introduced for self-powered wireless systems by attaching the antenna ground plane to the solar panel.

Later, researchers explored another technique to harvest the green energy from an arterial resource based on RF energy bands [14]. In such technology, the antenna design was introduced with certain specifications such as: High gain-bandwidth products to improve reception efficiency [15], unidirectional radiation patterns to ensure the locus independency [16], and size reduction to ensure system embedding and compatibility with miniaturized electronic circuits [17]. Nonetheless, the matching impedance between the proposed antenna with respect to the rectifier circuit is an important issue to ensure the rectification efficiency of the resulted structure [18].

Recently, several researchers have exploited fractal geometries in microwave circuits, such as antenna designs [19] and microwave filters [2]. To make microwave resonators, fractal geometries may be constructed from a variety of forms [21]. One of the most well-known is the Koch-fractal geometry, which has been adopted as a standard in microwave applications [22]-[29]. In [22-23], a modified Koch-fractal geometry was used to create microstrip band-stop filters. In [24-27], Koch and Minkowski showed how to make MIMO microstrip antenna arrays using fractal geometries. The Koch fractal geometry was utilized to uncouple side lengths of a triangular patch resonator to build a compacted dual-mode

band reject filter [28]. A 1D Koch fractal electromagnetic bandgap structure was proposed by [29].

To replace typical ground plane hole etching, a 1D Koch fractal electromagnetic bandgap structure was presented in [30]. The self-similar and space-filling properties of Koch fractal geometry were used to build a monopole antenna [31] based on a tri band for MIMO antenna applications. An octagonal shaped fractal Ultra-Wide Band (UWB) MIMO antenna was designed employing Minkowski fractal geometry of an octagonal form to achieve high downsizing factor with a bandwidth enhancement [32]. Microwave antenna designs might be employed in 5G [33] applications that need high antenna gain [34]. Due to propagation losses, several governments are unwilling to allot microwave bands for 5G potential uses; as a response, high-gain antenna designs have been proposed to avoid this difficulty with little propagation loss [35]. As a result, the sub-six-GHz band has been largely recognized as the preferred frequency range for 5G cellular communication systems [36]. Endfire antenna designs were one of the top competitors to satisfy 5G potential criteria due to their high gain bandwidth offers [37].

The proposed work is organized as following: In section II, the antenna array details are presented with all relative geometrical dimensions. The two scenario of array design two dimensions and three dimensions methodology are discussed in section III, section IV showed the effects of solar cell on antenna performance and channel performance. Finally, the conclusion of this work is presented in section V.

II. Antenna Array Design Details

The proposed antenna array is designed based on a fractal geometry that is proposed for self-powered applications [1]. The antenna array structure is consistent of four printed circuit antennas. The antenna elements are arranged as a cubic array as be shown in Fig. 1. The individual antenna element design is based on four Moore fractal inclusions fed with a coplanar waveguide (CPW) microstrip line. The resulted design is matched to a circuit load with a transmission line as seen in Fig. 1(a). The advantages of the Moore fractal, see Fig. 1(a), introduction are to increase the antenna size reduction and realize multiband resonances as MTM inclusion in the patch [13]. The antenna is printed on an FR4 Techtronic epoxy substrate with dimensions of 22×22×30mm^3. A bandwidth enhancement is achieved within a limited area by magnifying surface current that is afforded on electrically long path [14]. The proposed CPW is used to avoid the capacitive coupling between antenna element and the ground plane [15]. Nevertheless, such design maintains the other surface of the antenna substrate without any printed circuit to be used for busing electronic circuit [16]. The transmission line is designed with width of (0.84mm) to realize 50Ω match circuit and to transfer the surface current motion to the proposed Moore structure [17]. The air gap on the transmission line is introduced to force the current motion toward the fractal geometry as explained in [16]. The antenna ground plane is covered with EBG layer to avoid surface wave retardation that could be created in an opposite interference and inductive loss due to the use of the transmission line junction that can be removed by the capacitive effects of the matching circuit [18]. The proposed EBG is introduced to suppress the surface waves from the antenna edges those effects negatively on the radiation efficiency [19]. The microstrip patch is printed on the FR4 substrate without a ground plane. The substrate dielectric constant is 4.3 with height of 1.67mm. This antenna provides the broadside radiation patterns to cover the tangential components of the antenna radiations [2]. The proposed antenna for this study; three different configurations as 1D, 2D, and 3D as shown in Figs. 1(b), 1(c), and 1(d). The antenna is fabricated on the FR-4 substrate as shown in Fig. 1(e). The antenna ground planes are connected to avoid charges accumulations on the antenna body after introducing the solar panel [20]. The solar panel is mounted on top of the antenna structure as shown in Fig. 1(f).

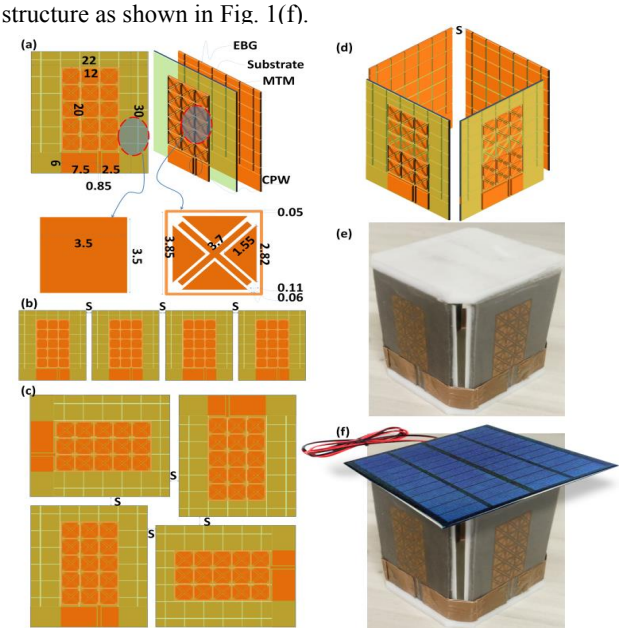


Fig. 1. Antenna array design (a) Front and back side (b) Moore fractal unit cell (c) Cubic antenna array (d) Feed port

III. Array Design Methodology

A- Single Antenna Design

The proposed antenna structure is consistent of four main parts: The first one is patch stricture without MTM circuitries as (case_1). The second case (case_2) is based on the MTM based on Moore geometry introduction with the patch layer. The introduction of the partial ground plane structure to the proposed design is called (case_3).

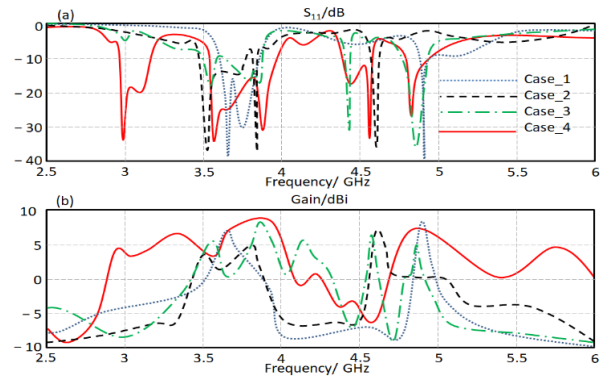


Fig. 2. Antenna performance based on a single antenna element: (a) S_{11} and (b) Gain

The proposed antenna structure based on EBG layer in Fig. 1(a) is presented in (case_4) after the matching circuit introduction. The proposed cases performances are calculated numerically using CST MWS in terms of S_{11} and gain spectra as shown in Fig. 2. Based on the evaluated results, it is found that the proposed antenna bandwidth is enhanced significantly after introducing are matching circuit to design in case_4 with maximum gain of 7.5dBi at 3.83GHz and 4.75GHz. However, the first mode is generated, at 2.96GHz, from the Moore fractal structure introduction. The other two modes were generated from the proposed designs in case_2 and case_3. The fundamental modes are generated by case_1 due to the current motion in the fractal geometry [22].

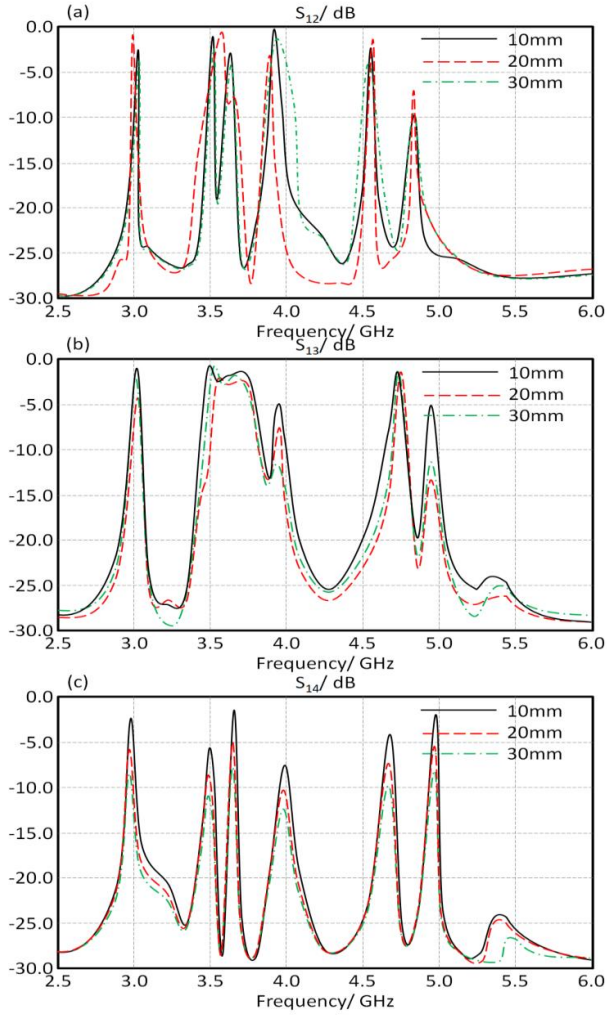


Fig. 3. Coupling effects of 1D configuration in terms of: (a) S_{12}, (b) S_{13}, and (c) S_{14} spectra.

B- Array Design

The proposed antenna is arranged with 1D, 2D, and 3D array scenarios. At beginning, four antenna elements are arranged with different separation distances between the antenna elements as presented in Fig. 1(b). The proposed scenarios at different separation distance are included to realize the effects of that between the mutual coupling spectra only in terms of S_{12} in Fig. 3(a), S_{13} in Fig. 3(b), and S_{14} Fig. 3(c). Therefore, the separation distances (s) are changed from 10 to 30mm with a step 10mm. It is found after reaching 30mm distance, the mutual coupling reaches -10dB for most S_{12}, S_{13}, and S_{14} at the frequency band of interest. Reaching this reduction in the mutual coupling could be good, however, such distance adds a great impact on the antenna array size. Therefore, the authors conducted their study to the next section.

Next, the proposed antenna is arranged based on 2D configuration. The separation distances between the antenna elements are organized as presented in Fig. 1(c). The antenna array configuration is oriented sequentially, to reduce the effectively the mutual coupling between antenna elements. The proposed scenarios at different separation distance are included to realize the effects of that between the mutual coupling spectra only in terms of S_{12} in Fig. 4(a), S_{13} in Fig. 4(b), and S_{14} Fig. 4(c). Therefore, the separation distances (s) are changed from 10 to 30mm with a step 10mm. It is found after reaching 30mm distance, the mutual coupling reaches -10dB for most S_{12}, S_{13}, and S_{14} at the frequency band of interest. Reaching this reduction in the mutual coupling could be good, however, such distance adds a great impact on the antenna array size. Therefore, the authors conducted their study to the next section.

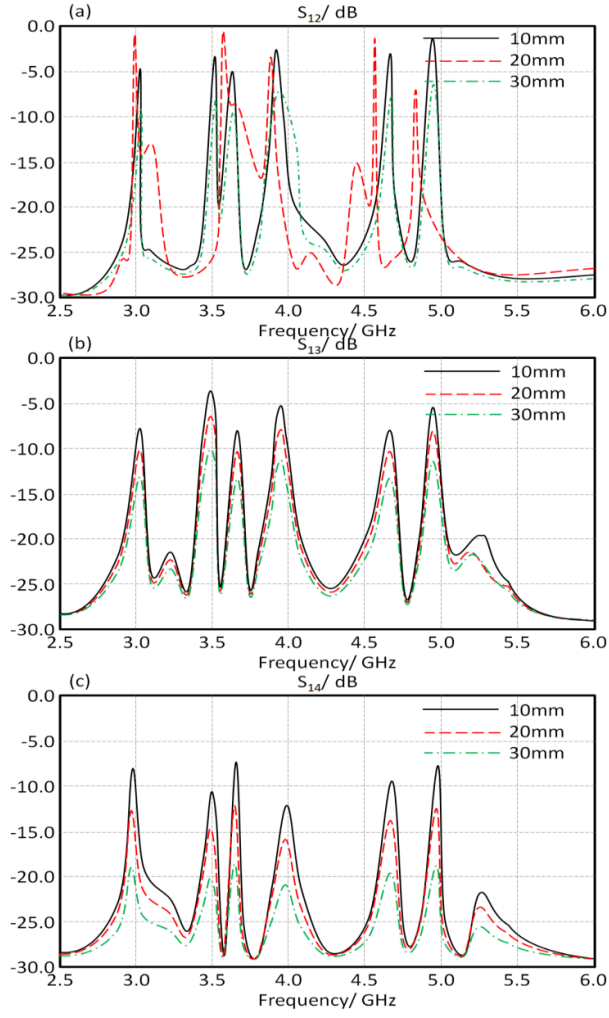


Fig. 4. Coupling effects of 2D configuration in terms of: (a) S_{12}, (b) S_{13}, and (c) S_{14} spectra.

Finally, the proposed antenna array is formed in shape of 3D configuration as shown in Fig. 1(d). Therefore, the proposed configuration with different separation distance (s) to monitor the mutual coupling between the antenna elements, S_{12}, S_{13}, S_{14}, as seen in Fig. 5(a)-5(c). The parameter (s) is changed from 1mm to 3mm with a step 1mm. It is recognized when s=2mm, the coupling in general is less than -15dB. However, when s=3mm, the coupling is reduced to less than -20dB. After reaching s=30mm, the mutual coupling becomes much less than -27dB for S_{12}, S_{13}, and S_{14} spectra at the frequency band of interest. Therefore, the authors considered s=3mm is enough for the design specification to provide mutual coupling of -20dB. Therefore, it is concluded according to the achieved results, the proposed antenna array based on 3D configuration shows excellent performances over other suggested configurations. This motivated us to move this configuration to the next step of the research.

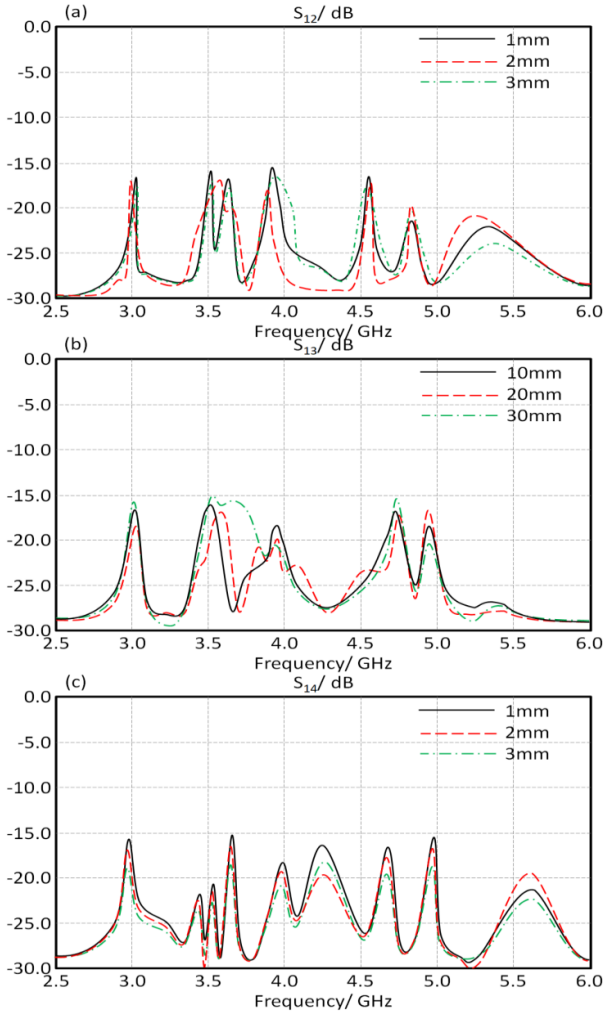


Fig. 5. Coupling effects of 2D configuration in terms of: (a) S_{12}, (b) S_{13}, and (c) S_{14} spectra.

IV. Solar Cell Effects

In this section, the authors presented the effects of introducing the solar cell on the proposed antenna array performance in terms of S-parameters and gain spectra experimentally. The simulation study is not invoked in this work due to the knowledge limitations in the considered solar panel electromagnetic constitutive parameters. Thus, it is very important to point out that all previous simulations are attempted without considering the solar panel introduction. The proposed MIMO antenna system is introduced to the solar panel as shown in Fig. 1(f). It is found after applying the experimental study, the antenna performance is found to be insignificantly affected with the solar panel introduction. Indeed, the antenna gain is enhanced slightly at the desirable frequency bands. This is achieved by reducing the surface wave effects due to introducing the antenna array normally to the solar panel structure which realizes less surface wave effects [4].

Now, the solar panel effects on the proposed antenna array performance are described in term of S_{11} and gain spectra with and without the solar panel introduction. The antenna array position with respect to the solar panel is presented in Fig. 1(f). In Fig. 6(a), the obtained S_{11} from the proposed antenna array with and without solar panel is shown; it is found that matching bandwidth, $S_{11} \leq -10$dB, is insignificantly affected. The antenna gain spectra of the proposed antenna array are evaluated as seen in Fig. 6(b). It is found that the proposed antenna array gain is significantly not affected after the solar panel introduction in specific at 3.6GHz and 3.9GHz to realize a moderate gain that is suitable for 5G and other modern applications [13]. The effective correlation effects is not significantly increased after introducing the solar panel structure as shown in Fig. 6(c). The total affective reflection coefficient and total channel capacity losses are not changed at all as depicted in Figs. 6(d) and 6(e). The evaluated antenna diversity spectra are not increased after the solar panel introduction.

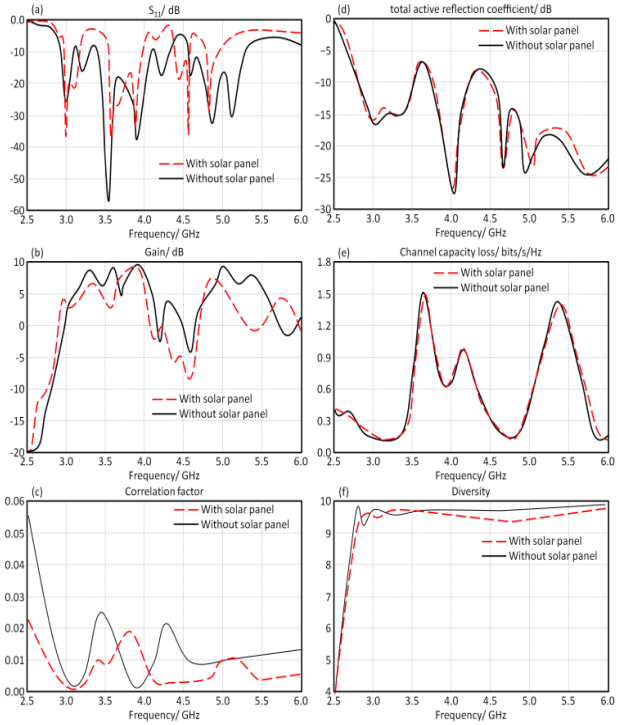


Fig. 6. Antenna array performance with and without solar panel introduction; (a) S_{11}, (b) gain, (c) correlation factor spectra, (d) total effective gain spectra, (e) Channel capacity losses, and (f) antenna diversity spectra.

V. EXPERIMENTAL VALIDATION

In this section the proposed antenna array performance are measured using a vector network analyzer (37347A) and RF chamber. The proposed antenna array is fabricated, as seen in Fig. 1(f), and validated experimentally. The obtained results from the numerical analysis are compared to the experimental measurements. As seen in Figs. 7(a) and 7(b), the measured S_{11} and gain spectra are presented. It is found that the experimental results agree very well with those obtained from CST MWS. The other relative measurements of the S-parameters in terms of S_{12}, S_{13}, and S_{14} spectra as seen in Fig. 7(c)-7(e). It is validated that the maximum coupling is about -20dB after the solar panel introduction with excellent gain and bandwidth enhancements. The harvested solar energy of the proposed antenna with and without antenna introduction is presented in Fig. 7(f). The proposed MIMO antenna is put to the test in conjunction with a solar panel. As a result, before and after the solar panel integration, the measured I-V characteristics are given in Fig. 7(f). The I-V characteristics are not greatly impacted after and before the solar panel integration since the solar panel are placed properly to the antenna array. After the planned antenna is installed, the I-V characteristics of the solar panels are measured. This comparison is made to determine the impacts of the antenna on the solar panel in question. As a result, minimal effects on solar panel performance are discovered; see Fig. 7(f), after and before the antenna insertion. This is due to the antenna construction being located on the solar panel's rear panel. The solar energy is defined in the photo-current and photo-voltage curve.

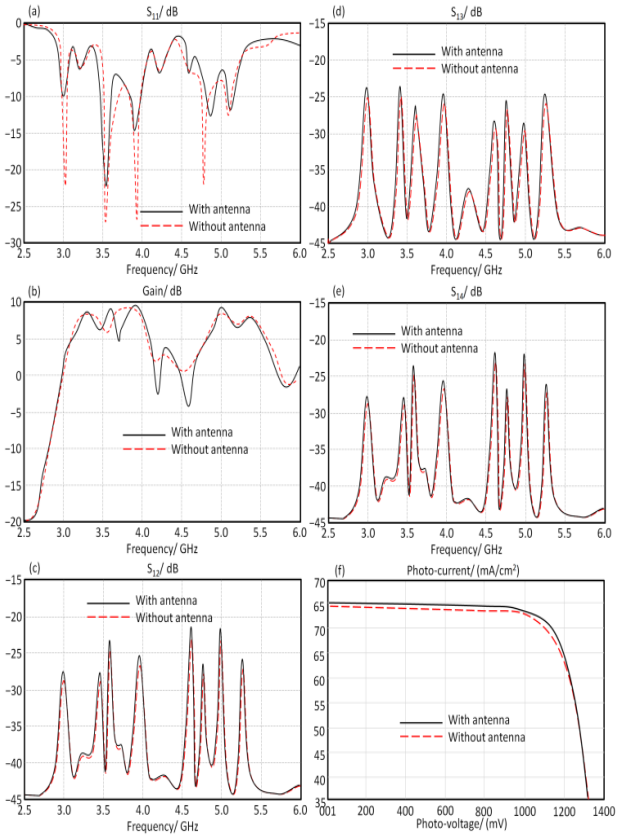


Fig. 7. Antenna performance with solar panel introduction; (a) S_{11}, (b) gain, (c) S_{12}, (d) S_{13}, (e) S_{14} spectra, and (f) I-V characteristics.

Based on the achieved results, the antenna performance is evaluated when introduced to real communication system based on QPSK modulation scheme. In this experiment, the total bit error rate (BER) and channel capacity are evaluated as shown in Fig. 8. The signal to nose ratio (SNR) are changed according to amount of the obtained power from the solar panel. According to our measurements, the total generated voltage was in the range of 0mV to 1000mV with constant current around 65mA as shown in Fig. 7(f). This gave us an indication of the total harvested energy from the considered solar panel. In such case, SNR is changed between 0dB to 60dB. It is found from the results in Fig. 8, a high SNR provide low BER, while, a low SNR would increase BER. In such case, this acquires an efficient solar panel for long use effectively. Nevertheless, the link budget effects could be significantly affected according to such knowledge. With the introduction of 5G the energy efficiency (Watt/Mbyte) of the 5G mobile radios many improvements were achieved. However, the total amount of the consumed energy is still increasing due to the sharply increasing data traffic that cause headache for the mobile network operators, since most of them operating not only 5G. It is obvious, frequency bands with higher gain are better than those with low gain to provide less chance of errors due to the noise effects.

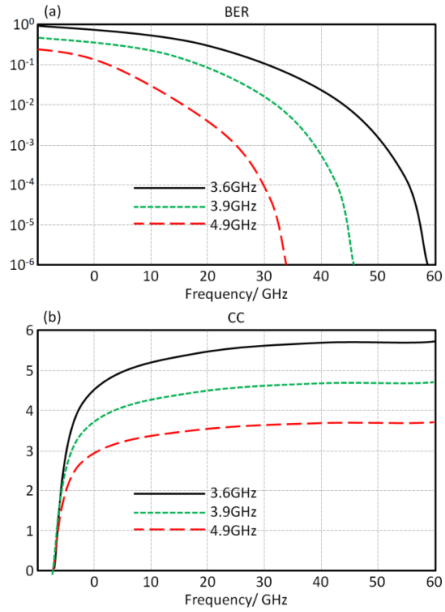


Fig. 8. Channel performance according to the solar panel I-V characteristics at different frequency bands. (a) BER and (b) CC calculations.

The proposed antenna radiation patterns are measured at 3.6GHz, 3.9GHz, and 4.9GHz as shown in Fig. 9. The measurements are given in both co- and cross-polarizations. It is concluded from the obtained results that the antenna is effectively very directive toward the broadside direction. Also, the introduction of the solar panel has no significant effects on the antenna directivity. It is compared in this section between the proposed antenna and other published results in Table I. It is found that the proposed antenna provides a highest miniaturized profile in comparison to other antenna designs based on 3D structures.

TABLE I
RESULTS COMPARISON WITH RESPECT OTHER PUBLISHED RESULTS.

Reference	Size/mm	Freq./GHz	Gain/dBi
23	70×30	3.78	-1.2
		8.24	0.56
24	100×100	2.45	3.45
25	130×130	5.6	5.8
26	60×70	2.45	3.4
27	90×90	5.8	6.7
28	60×40	2.45	2.3
29	90×90	2.45	1.2
30	22×30×72	2.45	2
32	40×50	2.45	2.1
33	20×30	3.5, 5	3, 3.5
This work	20×30	3.6	6.3
		3.9	7.9
		4.9	9

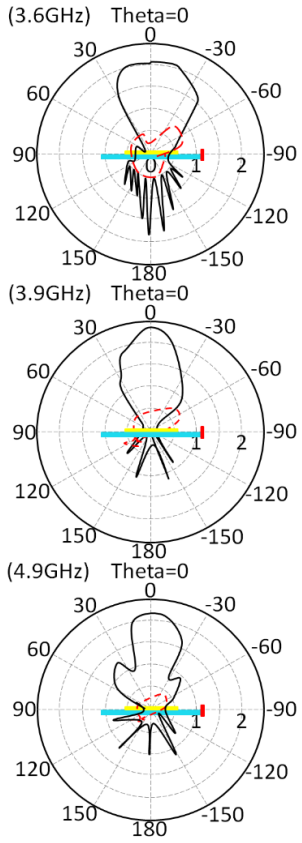


Fig. 9. Antenna radiation patterns measurements only after the solar panel introduction. Note: The discrete black line for co-polarization results and the red line for the cross-polarization results.

VI. CONCLUSION

In this work, a design of a 3D array antenna based solar panel integration for self-powered applications in modern wireless communication network. The proposed antenna array is structured as a cubical geometry with solar panel integration. Such integration is employed to achieve a self-powered node. The proposed antenna is designed to perform an excellent beam reconfiguration at sub-6GHz frequency bands. The proposed antenna is found to provide a moderate gain at 3.6GHz, 3.9GHz, and 4.9GHz. The proposed antenna array is found to provide coupling between the adjacent antenna elements bellow -20dB at the frequency bands of interest. The proposed antenna is fabricated and compared to the simulated results to achieve an excellent agreement between them. A solar panel is introduced to the proposed antenna array for self-powered applications. It is reviled that the proposed antenna system integration with solar panel shows no negative effects on the antenna performance. The antenna is introduced to a real wireless QPSK communication scheme. With such environments, we calculated the BER and CC with respect to the total obtained solar energy. This is done by introducing the amount of achieved power from the solar panel as SNR level to feed the QPSK scheme. We obtained within the limit of the harvested solar energy, acceptable BER and CC values. The antenna performance is validated theoretically and experimentally to arrive to an excellent agreement between the obtained results.

REFERENCES

[1] Abdulsattar, R. K.; Elwi, T. A.; Abdul Hassain, Z. A. A New Microwave Sensor Based on the Moore Fractal Structure to Detect Water Content in Crude Oil. *Sensors* 2021, 21, 7143. DOI: 10.3390/s21217143

[2] Alaukally M. N. N.; Elwi T. A.; Atilla D. C. Miniaturized flexible metamaterial antenna of circularly polarized high gain-bandwidth product for radio frequency energy harvesting. *Int J Commun Syst*. 2022; 35(3):e5024. DOI: 10.1002/dac.5024

[3] Yehia, S., Fatah, A.; Taher, F.; Elwi, T. A.; Fathy, M., Sree, A.; ... Limiti, E. (2023). Design of Compact Flexible UWB Antenna Using Different Substrate Materials for WBAN Applications. In *2023 Photonics and Electromagnetics Research Symposium (PIERS)*. Proceedings (373–378). DOI: 10.1109/PIERS59004.2023.10221357

[4] Ahmed Imad Imran, Taha Ahmed Elwi, and Ali J. Salim, "On the Distortionless of UWB Wearable Hilbert-Shaped Metamaterial Antenna for Low Energy Applications," *Progress In Electromagnetics Research M*, Vol. 101, pp. 219–239, 2021, DOI: 10.2528/PIERM20113008.

[5] S. M. Obaid, T. A. Elwi, and M. Ilyas, "Fractal Minkowski-Shaped Resonator for Noninvasive Biomedical Measurements Blood Glucose Test," *Progress in Electromagnetics Research C*, Volume 107, pp. 143–156, December 2020, http://dx.doi.org/10.2528/PIERC20072603.

[6] Elwi T. A., Al-Saegh A. M. Further realization of a flexible metamaterial-based antenna on indium nickel oxide polymerized palm fiber substrates for RF energy harvesting. *International Journal of Microwave and Wireless Technologies*. 2021;13(1):67–75. DOI: 10.1017/S1759078720000665

[7] Y. Alnaiemy, T. A. Elwi, L. Nagy, "An end fire printed monopole antenna based on electromagnetic band gap structure," *Automatika*, volume 61, issue 3, pp. 482–495, DOI: 10.1080/00051144.2020.1785783.

[8] T. A. Elwi, "Remotely Controlled Reconfigurable Antenna for Modern Applications", *Microwave and optical letters*, Volume 6, Issue 1, pp. 1–19, April 2020, DOI: 10.1002/mop.32505.

[9] A. Abdulmjeed, T. A. Elwi, and S. Kurnaz, "Metamaterial Vivaldi Printed Circuit Antenna Based Solar Panel for Self-Powered Wireless Systems," *Progress In Electromagnetics Research M*, Vol. 102, 181–192, 2021, http://dx.doi.org/10.2528/PIERM21032406

[10] M. A. Jawad, M. A. Elwi, E. Y. Salih, T. A. Elwi, and Zulkifly Abbas, "Monitoring the Dielectric Properties and Propagation Conditions of Mortar for Modern Wireless Mobile Networks," *Progress in Electromagnetic Research Letters*, Volume 89, pp. 91–97, January 2020, http://dx.doi.org/10.2528/PIERL19090912.

[11] P. Mishra and S. S. Pattnaik, "Metamaterial loaded fractal based interdigital capacitor antenna for communication systems," *Progress In Electromagnetics Research*, Vol. 70, pp. 127–134, 2018, http://dx.doi.org/10.2528/PIERM18032801.

[12] Y. Alnaiemy, T. A. Elwi, L. Nagy, "Mutual Coupling Reduction in Patch Antenna Array Based on EBG Structure for MIMO Applications", *Periodica Polytechnica Electrical Engineering and Computer Science*, Volume 1, number 4, pp. 1–11, September 2019, DOI: 10.3311/PPee.14379,

[13] T. A. Elwi, "Further Investigation on Solant-Rectenna based Flexible Hilbert-Shaped Metamaterials", *IET Nanodielectrics*, Volume 4, Issue 12, pp. 1–12, March 2020, DOI: 10.1049/iet-nde.2020.0013.

[14] H. M. Al-Sabbagh, T. A. Elwi, Y. Al-Naiemy, and H. M. Al-Rizzo, "A Compact Triple-Band Metamaterial-Inspired Antenna for Wearable Applications", *Microwave and Optical Technology Letters*, Volume 11, Number 2, October 2019, DOI: 10.1002/mop.32067.

[15] Z. A. Abdul Hassain, A. R. Azeez, M. M. Ali, and T. A. Elwi, "A Modified Compact Bi-Directional UWB Tapered Slot Antenna with Double Band-Notch Characteristics", *Advanced Electromagnetics*, Volume 8, number 4, September 2019, http://dx.doi.org/10.7716/aem.v8i4.1130.

[16] T. A. Elwi, Z. A. AL-Hussain, O. Tawfeeq, "Hilbert Metamaterial Printed Antenna based on Organic Substrates for Energy Harvesting", *IET Microwaves, Antennas & Propagation*, volume 12, number 4, pp. 1–8, June 2019, DOI: 10.1049/iet-map.2018.5948.

[17] H. S. Ahmed and T. A. Elwi, "On the design of a reject band filter for antennas mutual coupling reduction", *International Journal of RF and Microwave Computer-Aided Engineering*, volume 11, number 3, pp.1–11, April 2019, DOI: 10.1002/mmce.21797.

[18] Y. Al Naiemy, T. A. Elwi, and L. Nagy, "An end fire printed monopole antenna based on electromagnetic band gap structure," *Automatika*, vol. 61, no. 3, pp. 482–495, 2020, DOI: 10.1080/00051144.2020.1785783.

[19] T. A. Elwi, D. A. Jassim, H. H. Mohammed, "Novel miniaturized folded UWB microstrip antenna-based metamaterial for RF energy harvesting," *International Journal of Communication Systems*, Volume 1, Issue 2, January 2020, DOI: 10.1002/dac.4305.

[20] R. R. K. Al-Taie et al., "On the Performance of a Composite Right Left Hand Electromagnetic Bandgap Structure," *2022 9th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, Jakarta, Indonesia, 2022, pp. 420–423, DOI: 10.23919/EECSI56542.2022.9946487.

[21] Hussein, H.; Atasoy, F.; Elwi, T.A. Miniaturized Antenna Array-Based Novel Metamaterial Technology for Reconfigurable MIMO Systems. *Sensors* 2023, 23, 5871. DOI: 10.3390/s23135871.

[22] Zainab S. Muqdad, Mohammad Alibakhshikenari, Taha A. Elwi, Zaid A. Abdul Hassain, Bal.S. Virdee, Richa Sharma, Salahuddin Khan, Nurhan Türker Tokan, Patrizia Livreri, Francisco Falcone, Ernesto Limiti, "Photonic controlled metasurface for intelligent antenna beam steering applications including 6G mobile communication systems", *AEU - International Journal of Electronics and Communications*, Vol. 166, 2023, 154652, ISSN 1434-8411, DOI: 10.1016/j.aeue.2023.154652.

[23] M. R. I. Faruque, M. T. Islam, and N. Misran, "Design analysis of new metamaterial for EM absorption reduction," *Progress In Electromagnetics Research*, vol. 124, pp. 119–135, 2012, http://dx.doi.org/10.2528/PIER11112301.

[24] H. Nakano, Low-profile Natural and Metamaterial Antennas: Analysis Methods and Applications. John Wiley & Sons, 2016.

[25] A. R. Al-tameemi et al., "A Novel Conformal MIMO Antenna Array based a Cylindrical Configuration for 5G Applications," *2022 9th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, Jakarta, Indonesia, 2022, pp. 446–451, DOI: 10.23919/EECSI56542.2022.9946617.

[26] S. H. Ghadeer, T. A. Elwi and S. K. A. Rahim, "Compact MIMO Antenna Array for 5G Applications," *2022 9th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, Jakarta, Indonesia, 2022, pp. 399–402, DOI: 10.23919/EECSI56542.2022.9946554.

[27] Ismail, M. M.; Elwi, T. A.; I Salim, A. J. (2022). A Miniaturized Printed Circuit CRLH Antenna-based Hilbert Metamaterial Array. *Journal of Communications Software and Systems*, 18 (3), 236–243, DOI: 10.24138/jcomss-2022-0030.

[28] Y. Alnaiemy, T. A. Elwi and N. Lajos, "Enhancing the Microstrip Antenna Gain Using a Novel EBG Lens Based on a Single Layer," *2018 11th International Symposium on Communication Systems, Networks & Digital Signal Processing (CSNDSP)*, Budapest, Hungary, 2018, pp. 1–4, DOI: 10.1109/CSNDSP.2018.8471786.

[29] Y. Alnaiemy, T. A. Elwi, L. Nagy, and T. Zwick, "A systematic analysis and design of a high gain microstrip antenna based on a single EBG layer," *Infocommunications Journal*, vol. 10, no. 4, pp. 22–30, 2018, http://dx.doi.org/10.36244/ICJ.2018.4.4.

[30] Z. Al-Dulaimi, T. A. Elwi, and D. C. Atilla, "Design of a meander line monopole antenna array based hilbert-shaped reject band structure for MIMO applications," *IETE Journal of Research*, pp. 1–10, 2020, DOI: 10.1080/03772063.2020.1743207.

[31] Jwair, Marwah Haleem and Elwi, Taha A. (2023) Metasurface Antenna Circuitry for 5G Communication Networks. *Infocommunications Journal: A Publication of the Scientific Association for Infocommunications (HTE)*, 15 (2). pp. 2–7. ISSN 2061-2079, DOI: 10.36244/ICJ.2023.2.1.

[32] H. Almizan, Z. A. A. Hassain, T. A. Elwi and S. M. Al-Sabti, "Controlling Gain Enhancement Using a Reconfigurable Metasurface Layer," *2021 12th International Symposium on Advanced Topics in Electrical Engineering (ATEE)*, Bucharest, Romania, 2021, pp. 1–6, DOI: 10.1109/ATEE52255.2021.9425037.

[33] H. Almizan, Marwah Haleem Jwair, Yahiea Al Naiemy, Zaid A. Abdul Hassain, Lajos Nagy and Taha A. Elwi, "Novel Metasurface based Microstrip Antenna Design for Gain Enhancement RF Harvesting", *Infocommunications Journal*, Vol. XV, No 1, March 2023, pp. 2–8., DOI: 10.36244/ICJ.2023.1.1.

[35] Y. A. Jassim, M. Çevik and T. A. Elwi, "10GHz Printed Circuit Antenna for Wireless Power Transfer Applications," *2023 5th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, Istanbul, Turkiye, 2023, pp. 1–4, DOI: 10.1109/HORA58378.2023.10156795.

[36] Ali, L., Ilyas, M., & Elwi, T. A. (2023). A metamaterial-based compact MIMO antenna array incorporating hilbert fractal design for enhanced 5G wireless communication networks. *Mathematical Modelling of Engineering Problems*, 10(3), 930–936, https://hdl.handle.net/20.500.12939/3672.

[37] M. H. Jwair et al., "Intelligent Metasurface Layer for Direct Antenna Amplitude Modulation Scheme," in *IEEE Access*, vol. 11, pp. 77 506–77 517, 2023, DOI: 10.1109/ACCESS.2023.3297264.

Ammar Al-Adhami was born in Baghdad, Iraq, in 1993. He received the B.Sc. degree in Electrical Engineering from Al-Ma'amon University, Iraq in 2015, He received M.Sc degree from Gaziantep University, Turkey 2017. He received PhD degree from Gaziantep University, Turkey 2022. His research interests include wireless communication systems, microwave circuits design, wearable antennas, and antenna design.

Yasir Al-Adhami was born in Baghdad, Iraq, in 1990. He received the B.Sc. degree in Computer communication Engineering from Al-Mansour University, Iraq in 2012, He received M.Sc degree from Çankaya University, Turkey 2014. He received PhD degree from Gaziantep University, Turkey 2018. He is currently Lecturer in Electrical and Electronics Engineering Department of Gaziantep University. His research interests include Harvest Solar Energy and RF Energy for wireless applications, wireless communication systems, microwave circuits design, and antenna design.

Taha A. Elwi received his B.Sc. in Electrical Engineering Department (2003) (Highest Graduation Award), and Postgraduate M.Sc. in Laser and Optoelectronics Engineering Department (2005) (Highest Graduation Award) from Al-Nahrain University Baghdad, Iraq. From April 2005 to August 2007, he worked with Huawei Technologies Company, in Baghdad, Iraq. On January 2008, he joined the University of Arkansas at Little Rock and he obtained his Ph.D. in December 2011 in system engineering and Science. He is considered of Stanford University's top 2% scientists in 2022. His research areas include wearable and implantable antennas for biomedical wireless systems, smart antennas, WiFi deployment, electromagnetic wave scattering by complex objects, design, modeling, and testing of metamaterial structures for microwave applications, design and analysis of microstrip antennas for mobile radio systems, precipitation effects on terrestrial and satellite frequency re-use communication systems, effects of the complex media on electromagnetic propagation and GPS. His research is conducted to consider wireless sensor networks based on microwave terminals and laser optoelectronic devices. The nano-scale structures in the entire electromagnetic spectrum are a part of his research interest. Also, his work is extended to realize advancements in reconfigurable intelligent surfaces and control the channel performance. Nevertheless, the evaluation of modern physics phenomena in wireless communication networks including cognitive radio networks and squint effects is currently part of his research. His research interests include pattern recognition, signal and image processing, machine learning, deep learning, game theory, and medical image analysis-based artificial intelligence algorithms and classifications. He serves as an editor in many international journals and publishers like, MDPI, IEEE, Springer, and Elsevier. He is currently the head of the International Applied and Theoretical Research Center (IATRC), Baghdad Quarter, Iraq. Also, he has been a member of the Iraqi scientific research consultant since 2016. He is leading three collaborations around the world regarding biomedical applications using microwave technology. He is the supervisor of many funded projects and Ph.D. theses with corresponding of more than 150 published papers and holding 10 patents. He can be contacted at email: taelwi82@mail.com.

Improving CAN anomaly detection with correlation-based signal clustering

Beatrix Koltai, András Gazdag, and Gergely Ács

Abstract—**Communication on the Controller Area Network (CAN) in vehicles is notably lacking in security measures, rendering it susceptible to remote attacks. These cyberattacks can potentially compromise safety-critical vehicle subsystems, and therefore endanger passengers and others around them. Identifying these intrusions could be done by monitoring the CAN traffic and detecting abnormalities in sensor measurements. To achieve this, we propose integrating time-series forecasting and signal correlation analysis to improve the detection accuracy of an onboard intrusion detection system (IDS). We predict sets of correlated signals collectively and report anomaly if their combined prediction error surpasses a predefined threshold. We show that this integrated approach enables the identification of a broader spectrum of attacks and significantly outperforms existing state-of-the-art solutions.**

Index Terms—**CAN, Anomaly Detection, TCN, Correlation.**

I. INTRODUCTION

SECURING vehicular communication networks is becoming crucial as the automotive industry rapidly evolves and increasingly adopts connectivity. Applying Intrusion Detection Systems (IDS) in specific domains is becoming essential for identifying and mitigating threats to vehicular networks. One such domain is the vehicles' inner communication on the Controller Area Network (CAN).

The CAN bus is a complex network of Electronic Control Units (ECUs) that collaborate to provide the necessary functions of the vehicle. Cyber attacks targeting these ECUs can have dire consequences for safety-critical subsystems such as brakes, the engine, or the steering wheel. A malfunctioning vehicle not only endangers passengers and others around it but also impacts the VANET (Vehicular Ad-hoc Network). Compromising data used in Vehicle-to-Everything (V2X) communication, an attacker could spread malicious information and alter the behavior of others, which could cause congestion or severe accidents in an urban environment. An attacker can have financial motivation besides deteriorating reliability and driving safety. Gaining control over the vehicle could allow theft, stealing sensitive data, and sabotaging the system.

Laboratory of Cryptography and System Security Department of Networked Systems and Services, Budapest University of Technology and Economics
E-mail: {bkoltai, agazdag, acs}@crysys.hu

Since the CAN protocol does not implement any security measures [1], an attacker can potentially attack the ECUs by making communication inaccessible, injecting new malicious messages, or even modifying valid messages. DoS (Denial-of-Service) attacks disable the benign CAN communication by flooding the network with the highest priority messages. However, this attack can be easily detected because the network load is significantly increased during the attack. Message injection can also affect specific vehicle functions, but these attacks are also easy to detect, with simple statistical methods, as injected messages cause a recognizable change in the regular arrival times.

The most challenging issue is message modification attacks that do not introduce new messages to the network, only the data contents are changed. This attack is the hardest to detect due to the variability in traffic patterns, lack of authentication or encryption, the existence of stealthy attack techniques, and the lack of attack signatures. In general, only the continuously changing message data can be used for identifying anomalies that requires general, accurate methods to differentiate between normal and malicious behavior.

After extracting signals from the message data, the detection of malicious message modifications follow two main approaches: time-series forecasting [2], [3], [4] and signal correlation analysis [5], [6]. In time-series forecasting, a machine learning model is trained per signal that predicts the next, expected signal value. Anomaly is reported when there is a substantial deviation between the prediction and the actual value. Unfortunately, this method is incapable of identifying modifications that fall within the usual, non-anomalous range of signal values, even if they constitute an attack. For instance, this limitation is evident when the speed value is modified, causing it to marginally fall below the speed limit. To overcome this shortcoming, the deviation of the correlation between each pair of signals is checked, where correlation is calculated based on the most recent few minutes' worth of signal data [5], [6]. Indeed, increasing the speed should naturally result in a corresponding increase in the RPM signal; otherwise their correlation would appear anomalous. Consequently, to evade detection, an attacker would need to maintain the original correlation intact and simultaneously modify all correlated signals, which could be prohibitively expensive in practice. Nonetheless, unlike time-series forecasting, this purely correlation-driven approach is

DOI: 10.36244/ICJ.2023.4.3

unable to identify malicious alterations in signals that lack any correlation between them.

Our proposal combines the merits of both time-series forecasting and correlation analysis. We simultaneously forecast multiple correlated signals and flag an anomaly if the cumulative difference between the predicted values and the actual values of all correlated signals exceeds a specified threshold. The underlying idea is that, as a single model forecasts multiple highly correlated signals, any alteration in one signal will inevitably influence the predictions of all other correlated signals. In other words, we leverage signal correlation not only for more accurate prediction, but also to induce detectable deviation of the predicted signals from the actual ones even if only one of them is maliciously modified. For example, the larger the speed the larger the RPM value, which means that increased speed with constant RPM is likely to produce a noticeable cumulative prediction loss over *both* signals if they are predicted jointly by a single model. Furthermore, unlike pure correlation-based approaches, our method is capable of identifying malicious alterations in signals, even those that lack correlation, when their predicted values deviate significantly from their actual values. Additionally, it can detect attacks in which the attacker modifies correlated signals simultaneously without altering their correlation, yet still induces abnormal behavior.

Our contributions in this work are as follows:

- We employ a combination of time-series forecasting and signal correlation analysis to identify anomalies in the vehicular CAN bus. Our unsupervised method relies solely on *unlabeled* CAN traces for training and calibration prior to deployment. It operates by simultaneously predicting correlated signals that allows a more accurate detection of abnormal behaviour.
- We assess the effectiveness of our approach using a dataset comprising eight distinct message modification attack types. Our results demonstrate a substantial performance improvement over the state-of-the-art: we achieve a detection rate of 95% (compared to 68%) with a precision of 80% (versus 30%). Additionally, our method exhibits a minimal average detection delay of just 0.38 seconds.
- Finally, we show that in addition to modification attacks, our solution also effectively identifies injection attacks, allowing the identification of both types of attacks by a single algorithm.

The rest of the paper is organized as follows: Section II briefly covers prior research and developments in anomaly detection in Controller Area Networks. Section III summarizes the relevant background of the CAN bus and vehicular intrusion detection solutions. The attacker model is introduced in Section IV. Section V describes the proposed anomaly detection mechanism, the training process, and the detection process. Section VI evaluates the performance of the method on real-world CAN data. Finally, in Section VII we conclude our paper.

II. RELATED WORK

Intrusion detection systems used in in-vehicle networks differ from those used on the Internet because there are limited known attack signatures. Most research results are based on unsupervised learning, as the available data can only be used appropriately to describe the benign state of the systems. Following this approach, papers have been published on detecting message injection and modification attacks.

IDS systems often rely on measuring and monitoring the timestamp of message arrivals to detect injection attacks. Due to the periodical timing of CAN data messages in a benign state, timing-based detection methods can effectively detect message insertions and drops [7], [8].

Attackers, however, cannot only inject messages into the bus, but it is also possible for them to modify messages, as described in Section IV.

In [9], the proposed method can detect these modification attacks by utilizing the transient state at the beginning of a modification attack. For a short time missing messages could indicate a suspension attack as a preparation step for a modification attack. However, if this phase is not detected in time, the rest of the attack will be successful.

In recent years, many papers have been published on identifying modification attacks based only on the message data contents. Among others, researchers tackled the problem by continuously measuring the relationship between data fields, forecasting future data values and later identifying deviations between the predictions and actual values.

CAN signal correlation analysis is proposed in [5] to identify modification attacks. Even though this approach is robust against attacks that target highly correlated signals, its effectiveness is generally limited. In [6], the authors extend correlation analysis with hierarchical clustering. Their results are demonstrated on a dataset, but it is not compared to other baseline results. As the presented framework can only handle entire traffic logs, it is not applicable as a real-time detector for the CAN bus but only as a forensics tool.

Time series forecasting is also used to predict future values in CAN communication, either on message or signal level. These predictive methods can identify possible modification attacks by measuring deviations between predicted and actual measured values.

Using a neural network for anomaly detection has been proposed in CANet [2]. Although this approach exploits relations between signals for detection, this information is not directly used in the network structure. In [3], the INDRA framework was proposed, which analyzes temporal patterns and behavior of messages using Gated Recurrent Unit (GRU) based recurrent autoencoders. The authors show that INDRA outperforms CANet in accuracy and false positive rate. In [4], the authors introduce a Temporal Convolutional Network based detection system. Their approach separates CAN signals and builds individual predictor models for each signal, similar to CANet and INDRA. However, as TCN networks are smaller and faster than previous neural networks, such as LSTMs, their solution outperforms all previous results. In this paper, we improve on the TCN-based approach by introducing signal

clustering to improve detection results while reducing the mechanism's footprint.

III. BACKGROUND

This section provides an overview of the CAN network's operation within vehicles, and introduces the application of Temporal Convolutional Neural Networks (TCNs) along with signal correlation analysis as part of our proposed anomaly detection approach.

A. CAN

Modern-day vehicles have a complex internal control system comprised of ECUs, each assigned to manage a specific function. These ECUs are interconnected via networks, the most important being the Controller Area Network. While this system has proven reliable over the years, external interfaces have exposed it to potential attacks [10].

On the CAN bus information is transmitted in frames. A CAN frame contains header, payload, and trailer segments. The actual data to be transmitted is in the payload segment.

Within the data part, various digital and analog signals are encoded. Manufacturers do not disclose how the signals are encoded, but they can be reverse-engineered using methods previously proposed in the literature [11].

B. Temporal Convolutional Networks

Convolutional Neural Networks (CNNs) and Temporal Convolutional Networks (TCNs) are deep learning architectures widely used for various tasks, including image recognition and natural language processing. They offer significant benefits when applied to time series data, making them suitable for detecting anomalies in the Controller Area Network (CAN) [4].

CNNs are designed to process grid-like data, such as images, by applying convolutional filters to extract spatial features. In the case of time series data, 1-dimensional causal convolutions can be used to identify local patterns and dependencies within the data.

To process sequences in parallel, TCNs use dilated convolutions, which enable them to capture long-range dependencies efficiently. This ability is critical in identifying anomalies that may occur over extended periods or exhibit complex temporal behaviors. Additionally, TCNs stack multiple layers for hierarchical feature extraction.

TCNs can handle large volumes of data, making them suitable for analyzing extensive CAN message traffic. This architecture can be optimized for real-time processing, allowing immediate anomaly detection and response in safety-critical CAN systems.

IV. ATTACKER MODEL

This section discusses the attacker model and the attack surface of a CAN network. We describe the capabilities and goals of an attacker and classify the potential attacks that an attacker may perform on CAN messages.

We assume that the attacker can gain access to the vehicle using the most common attack vectors [10]. The goal of the attacker is to send forged data to an ECU, forcing it into a corrupt state. This could cause problems anywhere between showing invalid values on the dashboard to making the vehicle completely unusable or stealing it[1], depending on the target ECU. This goal can be achieved in multiple ways. For example, vehicles with wireless interfaces, such as Bluetooth, WiFi, or a 3G/4G/5G connection, can also be attacked remotely. Once an attacker has the capability to interact with the CAN bus, there are multiple possible attack strategies, including DoS, message injection, and message modification. The latter two are also referred to as a fabrication and a masquerade attack.

We focus on the most challenging problem, which is the message modification attack. During these attacks the repetition times of the messages are unchanged, as there are no new messages introduced to the network. Hence, messages arrive at their expected time but with a modified data content. Carrying out such an attack requires strong technical skills, nevertheless, its feasibility has already been demonstrated in [12]. A practical implementation of such an attack exploits the error handling mechanism of the CAN protocol. If a device detects an error during transmission, an error signal bit can be used to inform the sender about the problem. Repeated error signals can force an ECU into an error state. In this state all further message transmissions are suspended, allowing an attacker to take the place of the ECU in the communication and send modified messages. Therefore, identifying modification attacks based only on meta-data (e.g., the number or timing of CAN messages) is not possible. In this paper, we present a novel anomaly detection mechanism, designed to detect such attacks.

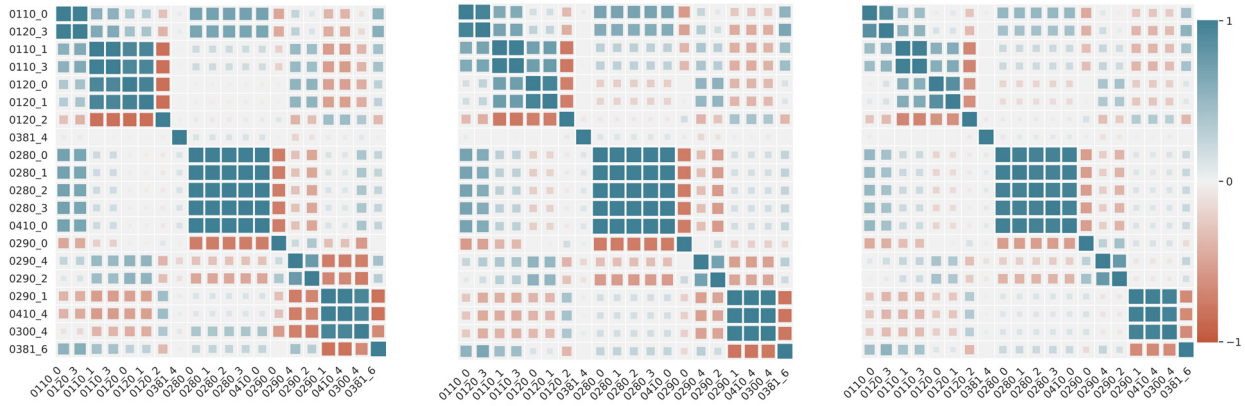
V. PROPOSED SOLUTION

Our solution has three main components: after extracting signals from the raw CAN traffic, (1) correlated signals are grouped together using clustering, (2) a separate and independent supervised forecasting model per group predicts the next value of all correlated signals within a group, and finally (3) an anomaly is reported if at least one of the forecasting model's predictions deviate significantly from the true, observed values of the predicated signals. We detail the operation of each component as follows.

A. Preprocessing of CAN Traffic

All signals from the available CAN messages are extracted using the manufacturer's specification or any state-of-the-art automatic extraction tool [11]. As not all extracted signals are equally useful for anomaly detection, a subset K of all extracted signals are retained while the rest are dropped. Indeed, useless signals are extracted from unused parts of the CAN messages (i.e., there is no device in the vehicle that uses that part of the message), or carry constant values with no

[1] https://arstechnica.com/information-technology/2023/04/crooks-are-stealing-cars-using-previously-unknown-keyless-can-injection-attacks



(a) Pearson correlation. (b) Spearman correlation. (c) Kendall's Tau correlation.

Fig. 1: Heatmaps of different correlation metrics used to determine similarity between signals (Pearson, Spearman and Kendall correlation form left to right). To better illustrate the magnitude of the correlation, we also varied the size of each point on the heatmaps, which is proportional to the darkness of the color.

predictive power. This filtering process also helps minimize the size of the forecasting model detailed in Section V-C. Finally, all retained signals are normalized by dividing each signal value by their theoretical maximum that is either specified by the manufacturer, or computed as $\lceil 2^s \rceil$ where s is the number bits used to store the signal in the CAN message.

B. Grouping of Correlated Signals

All retained K signals are clustered into C groups based on their pairwise correlation values. Although our approach is not restricted to any specific similarity measure or clustering technique, we show in this section that hierarchical clustering with Pearson correlation is the most effective combination. Specifically, each signal is first assigned to a separate cluster and then the closest clusters are iteratively merged until the number of clusters attains C. The distance of two clusters with centroids c_i, c_j is measured by $1 - |corr(c_i, c_j)|$, where $corr$ denotes the Pearson correlation.

1) Correlation analysis: We have analyzed Pearson, Spearman, and Kendall correlation metrics.

Pearson's correlation coefficient [13] measures the linear relationship between two continuous variables, suitable for typical analog signals on the CAN bus, such as speed, PRM, etc. It is sensitive to outliers, which means that extreme values can significantly influence the correlation value.

Spearman Rank Correlation [13] measures the strength and direction of the monotonic relationship between two variables by calculating the correlation based on the ranks of data points. Other than a monotonic relationship, it does not assume linearity or follow any specific distribution. The ranking property of this metric makes it robust to outliers.

Like Spearman, Kendall's Tau, also known as Kendall's rank correlation coefficient [13], does not assume linearity or follow any specific distribution. Kendall's Tau is often considered more robust than Spearman's.

A heatmap for each correlation method displaying the pairwise correlation between signals is shown in Figure 1. This visualization reveals that all three correlation methods exhibit nearly identical dependencies. However, some significant differences occur in the case of signals `0290_1`, `0410_4`, and `300_4`, which correlate only with signals `0290_4` and `0290_2` according to their Pearson coefficients. As Figure 2 shows, signal `0290_1`, `0410_4`, and `300_4` are indeed more similar, even though their Spearman and Kendall correlation values are significantly smaller.

2) Clustering of signals: We compared four distinct clustering algorithms on our dataset - DBSCAN, Affinity Propagation, Hierarchical Clustering, and Mean Shift Clustering[2]. We chose only clustering techniques that do not require the number of clusters to be specified in advance, as we do not know the optimal number of groups.

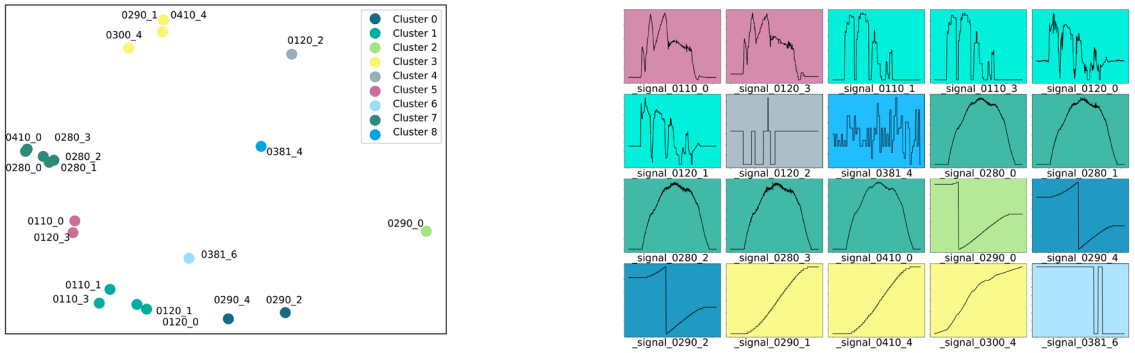
DBSCAN (Density-Based Spatial Clustering of Applications with Noise) is a density-based clustering algorithm that groups data points based on their density within the dataset [14]. It can discover clusters of arbitrary shapes and is robust to outliers called noise points.

Affinity Propagation is an exemplar-based clustering algorithm that selects a set of data points as exemplars and assigns the rest of the points to the nearest exemplar [15]. Affinity Propagation can be sensitive to the choice of similarity or distance metric, and the number of exemplars can significantly affect the results.

Hierarchical clustering builds a tree-like hierarchy of clusters, often represented as a dendrogram [14]. It can be agglomerative or divisive. Agglomerative clustering starts with individual data points as clusters. It merges them iteratively, while divisive clustering begins with a single set containing all data points and splits them into smaller groups.

Mean Shift clustering is a mode-seeking clustering algorithm that aims to find the modes or peaks of data density [16]. It is beneficial for finding clusters with non-uniform shapes or densities. Mean Shift is sensitive to the bandwidth parameter, which affects the size and shape of the groups.

[2] https://github.com/CrySyS/CAN-Message-Modification-Detection



(a) Clustering illustrated in a 2D representation, each point corresponds to a signal.

(b) Result of the clustering, background color indicates the cluster.

Fig. 2: Hierarchical clustering with Pearson correlation.

We visually compared the combinations of all clustering methods with each correlation metrics. In case of DBSCAN, we noticed that some signals are assigned to separate groups, even though they apparently belong together. Moreover, the result was sensitive to the clustering parameters. We also found Affinity Propagation method too sensitive to its parameters, and even with the best settings, it grouped signals that did not belong together. MeanShift and Hierarchical clustering essentially gave the same results. We opted to use the Hierarchical clustering algorithm with Pearson correlation for ease of use. Figure 2 illustrates the result, where the signals are represented in 2D space while preserving their pairwise similarities with Multidimensional Scaling (MDS).

C. Signal Forecasting

We train C supervised models on the clustered CAN data in order to predict the next upcoming signal value: all retained K signals are divided into equally-sized overlapping segments using a sliding window with size w, and each segment serves as input to the forecasting model to predict the subsequent signal value immediately following the segment.

More precisely, let a signal with ID s be represented as a time series $(T_1^s,\dots,T_n^s)$ after pre-processing, and $\mathbf{O}^G = [(T_1^{g_j},T_2^{g_j},\dots,T_n^{g_j})] \in \mathbb{R}^{|G|\times n}$ denotes the time series of all correlated signals in group G, where $G=\{g_1,\dots,g_{|G|}\}$ are the set of signal IDs belonging to G. For any signal group G, a forecasting model f_G simultaneously predicts the next element of each signal of the group: given the most recent w signal values $\mathbf{O}^G_{t-w:t} = [(T^{g_j}_{t-w},T^{g_j}_{t-w+1},\dots,T^{g_j}_{t-1})] \in \mathbb{R}^{|G|\times w}$ as input, the forecasting model predicts the next value $\mathbf{O}^G_{t:t+1} = (T_t^{g_1},T_t^{g_2},\dots,T_t^{g_{|G|}})^\top \in \mathbb{R}^{|G|}$ of every signal in G. Before deployment, all forecasting models are trained on CAN data that comes from the same or sufficiently similar distribution as the actual CAN traffic after deployment.

D. Decision

We compare the prediction made by every forecasting model with the actual, observed values of the signals, and report anomaly if the deviation of the prediction is too large for any group.

More precisely, let $\mathbf{O}^G_{t:t+1}$ denote the actual, observed value of the signals at time t in group G after performing the pre-processing steps detailed in Section V-A. The prediction error for group G at time t is defined as

$$\mathrm{err}_G(t) = \frac{1}{|G|}||f_G(\mathbf{O}^G_{t-w:t}) - \mathbf{O}^G_{t:t+1}||_2^2 \qquad (1)$$

which measures the mean squared error (MSE) between the actual signal values and the values predicated by f_G from the last w observed values of the signal. Note that $\mathbf{O}$ denotes the true value of the signal that is observed on-line after the deployment of the trained forecasting model f_G.

A naive method of detection is to directly compare the prediction error with a threshold τ, and report anomaly if $\mathrm{err}_G(t) \geq \tau$ for any group G. However, since the variance of $\mathrm{err}_G(t)$ can be large depending on the accuracy of the forecasting model f_G, this approach can yield large detection error: any value of τ would induce either too many false positives (for smaller τ) or false negatives (for larger τ). To mitigate such effect of forecasting inaccuracy, we rather compare the mean of the last ℓ error values with the threshold, that is, report anomaly if $(1/\ell)\sum_{i=t-\ell}^{t-1}\mathrm{err}_G(i) \geq \tau$ for any group G. This approach also more reliably detects stealthier attacks that span multiple time slots and involve insignificant modification of the signal value per slot, but surpass the threshold when aggregated.

To adjust τ, we follow the standard three-sigma rule and set τ to three times the standard deviation of $(1/\ell)\sum_{i=t-\ell}^{t-1}\mathrm{err}_G(i)$ plus its expected value on normal (attack-free) traffic [17]. The underlying assumption is that, without adversarial manipulation, the cumulative prediction error lies within three standard deviations of its mean that has a probability of 0.9973 if it is normally distributed (which is the case if ℓ is sufficiently large). The three-sigma rule is applicable even without access to attacked traffic before deployment, otherwise an optimal calibration of τ follows from the Neyman-Pearson lemma.

We performed experiments to determine the value of ℓ. We analyzed different values of ℓ ranging from 1 to 500, where $\ell = 1$ means that only the present error value is considered, and 500 corresponds to the mean computed over roughly 0.8 seconds of data. Too large values of ℓ can smooth out

shorter attacks potentially increasing false negative rate after deployment. On the other hand, too small values of ℓ yields larger variance of $(1/\ell)\sum_{i=t-\ell}^{t-1} \mathrm{err}_G(i)$ which can increase false positive rate.

Since attacked traces are usually not available during training, the value of ℓ is adjusted to minimize false positives only on benign signals. This is what we will do to evaluate our proposal in Section VI, and set the value of ℓ to 200.

In general, the values of τ and ℓ would depend on the manufacturer's priorities. For instance, a manufacturer may prefer to minimize false positives to detect and respond to attacks quickly or to investigate all suspicious cases. However, this can lead to missing some stealthy and short-duration attacks.

E. Discussion

1) Why Grouping Correlated Signals: The joint forecasting of correlated signals offers several advantages for anomaly detection. First, it allows a single model per group to leverage the inherent interdependencies among group members, resulting in more accurate forecasts for each signal within the group. Second, any malicious modification of a signal is likely to impact the predictions of all group members, thereby increasing the cumulative prediction error as described in Eq. (1). This enhances the detectability of attacks compared to prior methods in the literature, as demonstrated in Section VI. Finally, instead of creating a stand-alone model for each individual signal as in [4], our approach requires the construction of only K forecasting models, rendering it a more appealing choice in resource-constrained environments.

2) Cost Analysis: The cost of our approach is dominated by that of the forecasting models. Apart from the C forecasting models, $K \cdot w$ signal values are stored for forecasting and $K \cdot \ell$ error values for decision purposes. The forecasting models are trained off-line in parallel, and the trained models are deployed in the vehicle. Therefore, the computational cost is dominated by the inference time of the forecasting models, where the inference processes of models are parallelizable.

VI. EVALUATION

A. Dataset

We use two CAN datasets for evaluation: Dataset-1 introduced in [4], and Dataset-2 introduced in [18].

Dataset-1 contains seven short (<1 minute) traces of specific driving and traffic scenarios, and a longer trace (∼25 minutes). Dataset-2 contains nine short traces and eleven longer traces.

As the datasets originate from the same vehicle type, both have 20 message IDs and 1-6 signals per ID. Similarly, both datasets contain message injection and message modification attacks. As our main objective is to detect modification attacks, first we only use the corresponding traces.

We evaluate our mechanism on Dataset-1 to compare its performance to the chosen baseline described in Section VI-C. Since the two datasets are based on very similar CAN traffic from the same vehicle type, and most attacks follow the same strategy (only the RANDOM and DELTA attacks are not included in both), we present only the joint results.

The attacks have been performed using 6 different signal modification strategies:

- ADD-DECR - Modify with decrement value: a decrease per message is subtracted from the original value.
- ADD-INCR - Modify with increment: increases the original value by one increment per message.
- CONST - Change to constant: constant value replaces the original value.
- NEG-OFFSET - Modify with delta: a given value is subtracted from the original data value.
- POS-OFFSET - Modify with delta: a given value is added to the original data value.
- REPLAY - Replace the original data value with a previous value.
- DELTA - An attacker chosen value is added to the original value.
- RANDOM - The original value is replaced by a new random value in every attacked message.

B. Model Architecture and Parameters

For evaluation, we instantiate our proposal described in Section V. We create two datasets for training and testing purposes. A total number of 3.2 million CAN messages were used to create a training dataset for signal forecasting and calibrating all parameters of our approach (i.e., K, C, w, ℓ). Our calibrated model is tested on 1.3 million benign and malicious test messages (67 attacked traces and 9 benign traces), each containing one attacked signal. Both datasets undergo the same pre-processing steps with the same parameters that were computed exclusively on the training data.

a) Pre-processing: We use a signal mask based on the bit flip rate to extract relevant signals. We retain $K = 20$ of the $N = 77$ extracted signals that describe the state of the vehicle and likely to have sufficient predictive power for signal forecasting[3]. The retained signals are normalized as described in Section V-A.

b) Signal grouping: We conduct a correlation analysis on the signals and identify groups of correlated signals. We utilize hierarchical clustering with Pearson correlation as a similarity measure, and group linearly dependent signals together accordingly. We identify $C = 9$ clusters of the 20 signals in our dataset.

c) Signal forecasting: For forecasting, we use multi-channel Temporal Convolutional Networks (TCN). We apply an input sliding window of size $w = 1750$, equivalent to roughly 3 seconds, and each TCN has a receptive field with the same size w. Each channel of the multi-channel model corresponds to an individual signal in the group. The output of the TCN layers is then forwarded to a fully connected linear layer which generates the prediction of the upcoming signal values. Each multichannel TCN layer has four dilatation layers with a logarithmic offset of 2 $(1,2,4,8)$. The kernel size is fixed at 16. We train each forecasting model with Adam optimizer and MSE loss using early stopping.

[3] Note that this information is already known to a car manufacturer

The total size of all forecasting models, capable of handling all message IDs together in groups, is approximately 15 MB and contains 4.157 million parameters.

d) Decision: We average the last $\ell = 200$ prediction error values of our forecasting models and compare with threshold τ which is calibrated according to the three-sigma rule on the training data as described in Section V-D. In other words, we do *not* use the attacked traces in our dataset to adjust τ because it is unlikely to have sufficiently representative data about all possible attacks in practice.

C. Comparison with Baselines

The most relevant related works are CANet [2], INDRA [3], and the single TCN (S-TCN) anomaly detector architecture from [4]. To avoid confusion, from now on, we will refer to the Single TCN method (S-TCN), and refer to our proposed solution described in Section VI-B as Correlation-based TCN (C-TCN).

The INDRA framework has been shown to outperform other relevant unsupervised approaches including CANet regarding false positives and detection accuracy. Moreover, according to numerical experiments on two datasets, the SynCAN dataset [2] and Dataset-1, the S-TCN approach has larger accuracy with a significantly lower false positive rate than INDRA. Therefore, it is sufficient to show that our solution outperforms the S-TCN approach, because it has demonstrated superior performance compared to CANet and INDRA [4].

To properly compare the two results, we adapt the S-TCN approach by training one TCN model per signal but keeping the rest of the process, i.e., the data pre-processing, the same as our C-TCN solution. As expected, this adapted approach can reconstruct the expected behavior of CAN signals individually.

D. Evaluation Metrics

We evaluate both the baseline S-TCN and our proposed C-TCN method using standard performance metrics: accuracy, false positive rate, precision, and recall.

Precision and recall are particularly important metrics in this context, since the *testing* dataset is often imbalanced; attacks on the CAN bus are often short, which means that the number of benign instances significantly exceeds the number of attack instances.

In addition, we also measure the time it takes to detect attacks (denoted by T_D), and the fraction of attacked traces that are successfully detected (denoted by R_D):

$$T_D = \frac{\sum_{n=1}^{N_t} (t_{detection} - t_{attack})}{N_t} \tag{2}$$

$$R_D = \frac{\sum_{n=1}^{N_t} \mathbb{1}_{\{\text{trace } n \text{ is detected as anomalous}\}}}{N_t} \tag{3}$$

where N_t is the number of attacked traces, $t_{detection}$ is the time of detection (time of the first message whose signal values trigger anomaly), t_{attack} is the starting time of the attack (time of first attacked message) and $\mathbb{1}$ is the indicator function. Note that, while recall measures the detection performance on individual messages, detection rate measures the recall with respect to the traces. Indeed, both datasets used for evaluation includes short driving scenarios affected by various types of attacks, as described in Section VI-A, and an attacked trace is successfully detected if at least one message belonging to the attacked section of the trace triggers detection.

E. Results

All experiments were done using the TCN implementation in Keras [19].

Table I shows the accuracy and false positive rate for benign and malicious test sets, as well as the precision, recall, detection rate, and detection delay for attacked traces for both message modification and message injection attacks. These metrics are calculated across multiple traces and averaged to provide the overall results shown in the table.

To investigate the use of only one IDS system in a vehicle, we also tested our solution against message injection attacks. Although we do not focus on detecting these attacks, we demonstrate that the solution can be applied to detect message injections as well.

After experimenting, we conclude that correlation-based C-TCN can effectively detect attacks on CAN bus data. Our major findings are as follows:

1) Grouping of CAN signals based on correlation improves the detection performance from 68% to 95% which means that our proposed C-TCN method can detect 95% of all the modification attack scenarios. These attacks are detected with a delay of 0.38 seconds on average.
2) Correlation-based C-TCN significantly outperforms S-TCN on all evaluated metrics, especially regarding precision and recall, where C-TCN achieves 80-83% average performance.
3) In addition to modification attacks, C-TCN also effectively identifies injection attacks, allowing the identification of both types of attacks by a single algorithm.

As Figure 3 shows, S-TCN fails to detect the stealthier ADD-DECR attack, which slowly modifies the original signal message-by-message. It is only detected when the attack abruptly stops, and the signal returns to its original value. In contrast, our C-TCN model can detect the attack earlier when the modification induces a detectable change in the cumulative prediction error.

VII. CONCLUSION

This paper presented a novel approach to intrusion detection on the CAN bus. We mainly aimed at detecting message

TABLE I
COMPARING OVERALL RESULTS OF EVALUATING THE BASELINE S-TCN AND THE PROPOSED CORRELATION-BASED C-TCN ON BENIGN AND MALICIOUS TEST TRACES FROM BOTH DATASET.

	Benign		Message Modification		Message Injection	
	S-TCN	C-TCN	S-TCN	C-TCN	S-TCN	C-TCN
Acc.	0.98	**0.99**	0.93	**0.98**	0.96	**0.99**
FPR	0.03	**0.02**	0.05	**0.04**	0.01	0.01
Prec.	-	-	0.30	**0.80**	0.67	**0.88**
Recall	-	-	0.24	**0.83**	0.28	**0.70**
R_D	-	-	0.68	**0.95**	0.79	**0.94**

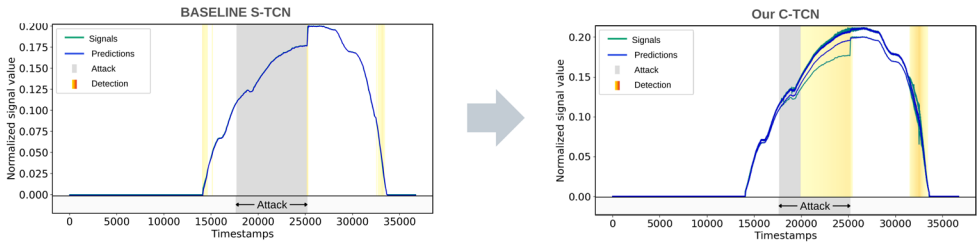


Fig. 3: Comparative evaluation of S-TCN vs. C-TCN on the same attacked trace, containing an ADD-DECR attack. The figure shows the attacked region marked by grey vertical lines and detections marked by yellow to red vertical lines, with the magnitude of the cumulative prediction error indicated by the darkness of the color.

modification attacks, the most complex attack type possible on the CAN bus. We showed that a correlation-based TCN model can efficiently predict the subsequent values of the vehicle signals, which can be used for anomaly detection. Finally, we also presented measurements demonstrating that our approach outperforms the state-of-the-art.

Our main contribution is to combine correlation analysis with time-series forecasting to improve detection accuracy. By grouping signals first based on their correlation, we create models that can predict future values with a high accuracy. During an attack, the forecasting of a group of correlated signals is significantly less accurate, allowing the detection of the anomaly. Furthermore, by grouping the signals, we can use fewer models resulting in a smaller footprint, which is an important factor for embedded systems.

In case an attacker knows which signals are clustered together and understands how the signals usually behave, it may be able to modify all the signals in the group without being detected. This requires maintaining the normal signal behavior including the inter-dependencies between different signals. However, it is unlikely that the attacker have all these capabilities in practice, especially if the groups are sufficiently large and the device running our integrated solution is adequately protected.

In our future work, we plan to analyze correlations in different traffic situations to improve our solution.

VIII. Acknowledgement

The research presented here have been supported by the NRDI Office, Ministry of Innovation and Technology, Hungary, within the framework of the Artificial Intelligence National Laboratory Programme, and the NRDI Fund based on the charter of bolster issued by the NRDI Office.

References

[1] M. Bozdal, M. Samie, S. Aslam, and I. Jennions, "Evaluation of can bus security challenges," *Sensors*, vol. 20, no. 8, 2020. [Online]. Available: https://www.mdpi.com/1424-8220/20/8/2364

[2] M. Hanselmann, T. Strauss, K. Dormann, and H. Ulmer, "Canet: An unsupervised intrusion detection system for high dimensional can bus data," *IEEE Access*, vol. 8, pp. 58 194–58 205, 2020.

[3] V. K. Kukkala, S. V. Thiruloga, and S. Pasricha, "Indra: Intrusion detection using recurrent autoencoders in automotive embedded systems," 2020.

[4] I. Chiscop, A. Gazdag, J. Bosman, and G. Biczók, "Detecting message modification attacks on the CAN bus with temporal convolutional networks," in *Proceedings of the 7th International Conference on Vehicle Technology and Intelligent Transport Systems*, 2021. [Online]. Available: DOI: 10.5220/0010445504880496

[5] A. Gazdag, G. Lupták, and L. Buttyán, "Correlation-based anomaly detection for the can bus," in *Euro-CYBERSEC*, Nice, France, 2021.

[6] P. Moriano, R. A. Bridges, and M. D. Iannacone, "Detecting can masquerade attacks with signal clustering similarity," *ArXiv*, vol. abs/2201.02665, 2022. [Online]. Available: https://api.semanticscholar.org/CorpusID:245836760

[7] H. M. Song, H. R. Kim, and H. K. Kim, "Intrusion detection system based on the analysis of time intervals of can messages for in-vehicle network," *2016 International Conference on Information Networking (ICOIN)*, pp. 63–68, 2016. [Online]. Available: https://api.semanticscholar.org/CorpusID:9333718

[8] A. Gazdag, D. Neubrandt, L. Buttyán, and Z. Szalay, "Detection of injection attacks in compressed can traffic logs," in *International Workshop on Cyber Security for Intelligent Transportation Systems, Held in Conjunction with ESORICS 2018*. Springer, 2018.

[9] S. Lee, H. J. Jo, A. Cho, D. H. Lee, and W. Choi, "Ttids: Transmission-resuming time-based intrusion detection system for controller area network (can)," *IEEE Access*, vol. 10, pp. 52 139–52 153, 2022.

[10] S. Checkoway, D. McCoy, B. Kantor, D. Anderson, H. Shacham, S. Savage, K. Koscher, A. Czeskis, F. Roesner, and T. Kohno, "Comprehensive experimental analyses of automotive attack surfaces," in *20th USENIX Security Symposium (USENIX Security 11)*. San Francisco, CA: USENIX Association, aug 2011. [On-line]. Available: https://www.usenix.org/conference/usenix-security-11/comprehensive-experimental-analyses-automotive-attack-surfaces

[11] M. E. Verma, R. A. Bridges, J. J. Sosnowski, S. C. Hollifield, and M. D. Iannacone, "Can-d: A modular four-step pipeline for comprehensively decoding controller area network data," *IEEE Transactions on Vehicular Technology*, vol. 70, pp. 9685–9700, 10 2021.

[12] K.-T. Cho and K. G. Shin, "Fingerprinting electronic control units for vehicle intrusion detection," in *Proceedings of the 25th USENIX Security Symposium*. *USENIX Association*, 2016, pp. 911–927. [Online]. Available: https://www.usenix.org/system/files/conference/usenixsecurity16/sec16_paper_cho.pdf

[13] E. F. El-Hashash and R. H. A. Shiekh, "A comparison of the pearson, spearman rank and kendall tau correlation coefficients using quantitative variables," *Asian Journal of Probability and Statistics*, vol. 20, no. 3, p. 36–48, Oct. 2022. [Online]. Available: https://journalajpas.com/index.php/AJPAS/article/view/425

[14] M. Z. Rodriguez, C. H. Comin, D. Casanova, O. M. Bruno, D. R. Amancio, L. d. F. Costa, and F. A. Rodrigues, "Clustering algorithms: A comparative approach," *PLOS ONE*, vol. 14, pp. 1–34, 01 2019. [Online]. Available: DOI: 10.1371/journal.pone.0210236

[15] B. J. Frey and D. Dueck, "Clustering by passing messages between data points," *Science*, vol. 315, no. 5814, pp. 972–976, 2007. [Online]. Available: https://www.science.org/doi/abs/10.1126/science.1136800

[16] M. A. Carreira-Perpinan, "A review of mean-shift algorithms for clustering," 2015.

[17] M.-C. Dani, F.-X. Jollois, M. Nadif, and C. Freixo, "Adaptive threshold for anomaly detection using time series segmentation," in *Neural Information Processing*, S. Arik, T. Huang, W. K. Lai, and Q. Liu, Eds. Cham: Springer International Publishing, 2015, pp. 82–89.

[18] A. Gazdag, R. Ferenc, and L. Buttyán, "Crysys dataset of can traffic logs containing fabrication and masquerade attacks," *Scientific Data*, 2023.

[19] P. Remy, "Temporal convolutional networks for keras," https://github.com/philipperemy/keras-tcn, 2020.

Beatrix Koltai is an MSc student at the Budapest University of Technology and Economics (BME) in Hungary. Her general research interests include the applicability of sequential data in security, and attack detection using machine learning methods.

András Gazdag assistant professor at Budapest University of Technology and Economics (BME), in Hungary. His research interests are in embedded systems security (with a special focus on vehicle security) and embedded system forensics.

Gergely Ács is an associate professor at Budapest University of Technology and Economics (BME), in Hungary. Before that, he was a post-doc and then research engineer at INRIA, in France. His general research interests include data privacy and security, as well as machine learning in this context.

Detection strategies for post-pandemic DDoS profiles

Péter Orosz, Balázs Nagy, and Pál Varga

Abstract—**The global pandemic lockdowns fostered the digital transition of companies worldwide since most of their employees worked from home using public or private cloud services. Accordingly, these services became the primary targets of the latest generation DDoS threats. While some features of current DDoS attack profiles appeared before the pandemic period, they became significant and reached their current complexity in the recent period. Besides applying novel methods and tools, the attacks' frequency, extent, and complexity also increased significantly. The combination of various attack vectors opened the way for multi-vector attacks incorporating a unique blend of L3-L7 attacking profiles. Unifying the hit-and-run method and the multi-vector approach contributed to the remarkable rise in success rate.**

The current paper has two focal points. First, it discusses the profiles of the latest DDoS attacks discovered in real data center infrastructures. To demonstrate and emphasize the changes in attack profile, we reference attack samples recently collected in various data center networks. Second, it provides a comprehensive survey of the state-of-the-art detection methods related to recent attacks. The paper especially focuses on the accuracy and speed of these, mostly networking-related detection approaches. Furthermore, we define features and quantitative and qualitative requirements to support detection methods handling the latest threat profiles.

Index Terms—**Intrusion detection and prevention, DDoS, Network security, Machine learning.**

I. INTRODUCTION

The pandemic lockdown fostered the ongoing digital transformation of society in many ways. Remote working and distance learning opened the way for new forms of group interactions. Online sale platforms were out for a more personal shopping experience for their customers. All of these transient shifts were supported by highly centralized cloud infrastructures that became the primary target of Distributed Denial of Service attacks. To improve the success rate, post-pandemic threats involve new methods and tools. A new set of protection methods should be developed and deployed to effectively improve the security level against high-complexity and high-intensity DDoS attacks. Our survey targets the presentation of post-pandemic DDoS attack profiles and their detection strategies, and goes beyond previous studies in highlighting technical depth. Moreover, we collected attack samples in a real data center and made them openly available (see related references in Section IV).

Department of Telecommunications and Media Informatics Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics, Budapest, Hungary
E-mail: orosz, bnagy, pvarga {@tmit.bme.hu}

Many industrial stakeholders predict that DDoS attacks are becoming "bigger" and more frequent in the coming years (according to Cisco [1], and Akamai [2]). Some recent DDoS attacks in 2020 already reached 2.3Tbps (AWS), and then 2.5Tbps (Google), which are much larger than the Mirai botnet attack against the DNS provider Dyn, estimated to reach as high as 1.5Tbps in 2016. A more sophisticated, multi-vector Mirai botnet variant attack reaching almost 2Tbps has also been captured at the end of 2021 by Cloudflare. These incidents dominated the most worrying global news. However, there are countless attack cases that may not hit the front page, although their *relative* impact on the given (less widely used) service or (less known) company could be much more pronounced.

How can we keep up with the adversaries? It is not only a matter of more machinery in the defense: detection methods need to be faster and more precise.

The exact methods to be used depend on the attack type; but detection time is a critical factors of success. Within the three main attack types – i.e., volumetric, protocol-based, and application-specific – the somewhat traditional attacker approach is brute force. However, the new breed of DDoS attacks has two typical types: massive volume amplification and/or volatile presence (Fig. 1).

There are numerous survey papers on the topic, although this current study of ours goes beyond their target in terms of timely presentation of new-generation DDoS attacks, as well as in technical depth. We focus on the accuracy and speed of threat detection. Among the many overviews, some of the suggested survey papers on the topic are the following. Peng, Leckie, and Ramamohanarao [3] surveyed "network-based" defense mechanisms against DDoS attacks in 2007. Their paper already included many of the terms, architectures and mechanisms we use today as a basic reference point. Zargar, Joshi, and Tipper [4] provided one of the earliest comprehensive surveys on modern defense mechanisms against DDoS flooding attacks in 2013. Masdari and Jalali [5] provided a comprehensive-at-the-time taxonomy of DDoS attack types in 2016, extending the focus to cloud infrastructures as well. In the same year, Yan et. al [6] described DDoS attacks from the perspective of Software-Defined Networking (SDN) and highlighted research issues and challenges, some of which are still open to this day. In 2018-2019 the challenges described by Yan were still not solved, but many significant steps were taken to harden SDN against DDoS [7], [8], [9].

DOI: 10.36244/ICJ.2023.4.4

Volumetric attacks have become more significant and use a broader set of methods than ever, especially for mixing various strategies. Devices and botnets have become rental objects; hence the group of users has also grown. These changes motivated the current article to go beyond previous overviews of the topic.

The contributions of the current paper are the following:

1) First, we define the main terms around DDoS analysis,
2) We provide a condensed comparison of the new breeds of DDoS attacks and discuss the related detection and mitigation methods,
3) We provide real-life captured DDoS traffic traces and analyze them in Section IV to help general comprehension.

The structure of the paper is the following. Section II provides basic definitions for the standard terms in the DDoS domain. Section III describes the new breed of DDoS attacks and the challenges raised by their existence, whereas Section IV provides a comprehensive and structured survey on the related detection methods. Section V surveys the modern methods for DDoS detection, including those based on artificial intelligence – especially machine learning – techniques. Finally, Section VI gives an outlook on DDoS trends in the future, and Section VII concludes the paper.

II. DEFINITIONS

This section provides brief definitions of terms that are commonly used in the domain of DDoS attacks, their detection and mitigation.

(D)DoS: The (Distributed) Denial of Service attack is a cyber threat that targets network segments or online services to deny access to certain resources and/or services. (D)DoS can be classified as an attack against the base of the CIA triad (availability). Since DDoS attacks are a lot more widespread now than DoS, we commonly refer to (D)DoS attacks as DDoS in this article.

IDS: The Intrusion Detection System monitors the network traffic for suspicious activity and issues alerts when such action is discovered. Intrusion detection systems are not designed to block attacks but to monitor the network and send alerts to system administrators if a potential threat is detected.

IPS: The Intrusion Prevention System supervises the access to an IT network and protect it from abuses and attacks. These systems are designed to monitor system data and take the necessary action to prevent an attack from developing.

IP spoofing: IP spoofing is the process of creating Internet Protocol (IP) packets that have a modified source address to either hide the identity of the sender, impersonate another computer system, or both. In theory, IP spoofing should not exist because ISPs are advised to implement source IP egress filtering. Still, in reality, many ISPs do not implement these filters. Spoofing is still very common in 2023.

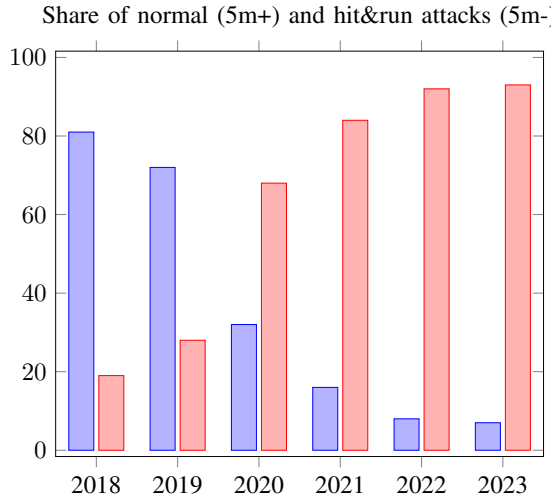


Fig. 1. Comparison of number of normal (longer than 5 minutes, blue) attacks and hit&run (shorter than 5 minutes, red) through 2018-2022 at the networks protected by AITIA SGA-NEDD

Volumetric DDoS attack (or Layer-3,-4 attack, flood attack): A DDoS attack that uses the sheer number (volume) of forged packets to achieve denial of service. Volumetric DDoS attacks primarily target network segments such as switches, routers, network processors, and data-links. This method of DDoS is by far the most popular among DDoS types because: a) the Internet is littered with poorly secured machines, IoT devices mainly, which can be organized into powerful botnets, and b) one botnet can be used to mount an effective attack against all targets.

Reflection DDoS attack: A volumetric DDoS attack that uses intermediary services of the Internet to amplify its attack throughput. It requires vulnerable Internet services – such as the NTP protocol – and the ability to inject packets into the network with spoofed source IP addresses. Reflection is a very effective and popular attack method: multi 100 Gbps attacks can be achieved with ease, and according to Akamai Inc., gives more than 50% of all DDoS attacks. The reasons are mainly the following: a) even 10000x amplification can be achieved, b) the attack's origin is obfuscated, and c) there is an abundance of widely used vulnerable services on the Internet.

Amplification DDoS attack: A DDoS attack that exploits a vulnerability related to asymmetric request-response volumes, where the response takes significantly more effort or contains considerably more data than the request. It is often used together with a reflection method. Hence, the attacker issues a "tiny" request (in effort or volume) to the reflection nodes, which reflects its (relatively) massive amount of data response to the victim node (instead of addressing the attacker). It is implemented using IP spoofing.

Application layer DDoS attack (or Layer 7 attack): A DDoS attack that uses application vulnerabilities to achieve denial of service. Application layer DDoS attack primarily targets computational resources like server processors and memory.

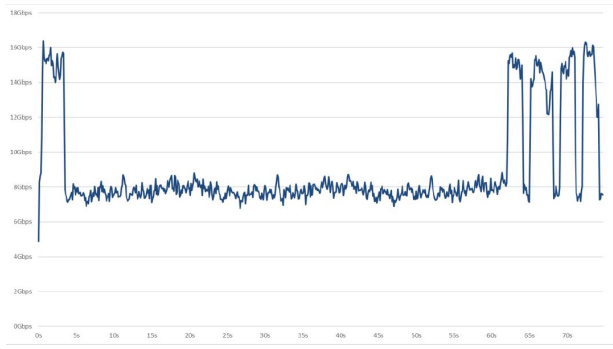


Fig. 2. An example of hit-and-run attack. This attack was captured by the authors of this paper in the network of KIFÜ (Hungarian Governmental Agency for IT Development). The x-axis represents time, and the y-axis shows measured throughput.

These attacks are tailored for their target, so each attack has a limited number of targets. This quality makes this kind of attack quite rare. These attacks are extremely different from volumetric attacks in the method of attack and the mitigation. This paper does not focus on application layer attacks; it instead aims to provide a detailed overview of volumetric attacks.

Hit-and-Run Attack: Volumetric DDoS attack that uses short bursts of attacking traffic to achieve its goals (Fig. 2). This attack is becoming increasingly popular because i) IDSs have problems detecting these kinds of attacks, and ii) it can achieve a lasting impact on the network through congestion control mechanics, such as TCP congestion control.

False positive detection rate: The portion of traffic identified falsely as a DDoS attack – although that was genuinely legitimate traffic. It is calculated as the ”number of packets falsely identified as belonging to the DDoS attack” divided by the ”number of all packets” that arrived in the time period.

False negative detection rate: The portion of traffic identified falsely as legitimate – although that was truly a DDoS attack. It is calculated as the ”number of packets falsely identified as legitimate” divided by the ”number of all packets” that arrived in the time period.

Detection time of a DDoS attack: The time span between the arrival time of the first packet of the attack and the decision at the IDS. Commonly also referred to as detection lag.

North-South attack: An attack where malicious traffic originates from outside the data-center hosting the under-attack service.

East-West attack: An attack where the malicious traffic originates from the data-center hosting the targeted service. Method to circumvent the main-defensive lines of the DCN, east-west (internal) routes are almost always less protected than north-south routes.

III. New profiles of DDoS attacks: methods, tools, and challenges

As a generic definition for Denial-of-Service (DoS) attack, it is a particular type of malicious traffic that attempts to make an online service unavailable for normal service users. Its distributed version (Distributed-DoS) enhances the threat’s effectiveness by concurrently generating malicious traffic from many contributing sources (usually many thousands or even more) to a single target. The traffic distribution enables a much larger traffic volume (nowadays, it may well exceed the Terabit order) to be developed and directed toward the targeted host or service. In the last decade, we could face a new wave of DDoS methods and attacks that have become the most common threats on the Internet due to their relatively easy and automated execution (see Table I). DDoS attempts usually target the resources of service and cloud providers. The new breed of DDoS threats may involve key novelties: i) vulnerable IoT devices as their security suites often miss even the basic protecting tools, ii) shorty living and pulsating volatile traffic patterns to be under the radar even for state-of-the-arts IDS/IPS systems, iii) very high volume of cumulative traffic generated by various amplification techniques, and iv) composite malicious traffic by combining various DDoS types (so-called vectors) to construct a multi-vector attack.

Typically, we distinguish three main categories of DDoS attacks: volumetric, protocol-based, and application-specific. While volumetric attacks focus on saturating bandwidth on the server’s local network, protocol-based variants target the exhaustion of server-side hardware resources, i.e., system memory, CPU, and IO bus. From the complexity perspective, application-specific attacks have significantly more sophisticated operations, specifically targeting a web service or other application.

Here, we provide reasons and arguments for the appearance of multi-vector attacks during the pandemic.

A. New methods and tools

The post-pandemic DDoS threats’ major novelty over the more conventional DDoS operational patterns is the development and amalgamation of two previously existing techniques: massive volume amplification and volatile presence. Moreover, applying this blend of techniques in multiple attack vectors challenges the security systems of data centers and cloud services and calls for a new generation of DDoS detection methods and implementations. Using the latest techniques, an attacker does not even require to access large-scale botnet resources and gain control over them to achieve a substantial attack volume. Instead, new attack techniques make one or many public service hosts send a response message to a spoofed destination address, i.e., to the targeted server host’s address. An alternative way to amplify malicious traffic is to send a small-sized request message to the targeted host with a spoofed source address, which triggers a large response message to that address. This asymmetry between request and response messages results in low resource utilization on the attacker-side and may sink all resources on the server-side.

Amplification/reflection: By sending spoofed requests, the attacker triggers responses from a group of open DNS or NTP servers back to the victim’s address (Fig. 3). Since the reply is typically more extensive than the request, the cumulative traffic

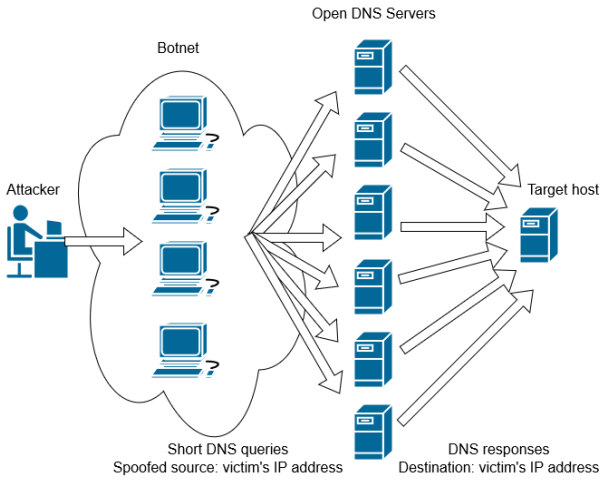


Fig. 3. Attack amplification/reflection mechanism

of the targeted response messages can saturate the network path between the attacked host and the Internet.

Volatile (hit-and-run) attack: In contrast to conventional DDoS threats, volatile attacks apply a periodic on/off strategy for controlling their presence on the network (see Fig. 2). In this case, the ON period is concise, typically lasting from milliseconds to minutes only, followed by an extended OFF period. This behavior is often successful since most IDS/IPS systems today have a detection time in the second range. Thus, these malicious traffic transients can reach the target host under the detection radar.

Multi-vector attack: It combines multiple methods and techniques to over-consume the resources of the target system in various ways. Mitigating these attacks can be challenging and often requires a multi-layer mitigation strategy. An efficient way to make an attack successful is to generate a complex traffic pattern that is easy to blend with regular traffic. Thus, multi-vector attacks may increase the probability of false positive detection that can block out an indefinite portion of user traffic along with the malicious one. The most popular component vectors are DNS reflection/amplification, TCP-Syn, TCP-Ack, TCP-Syn/Ack, TCP-Rst, and ICMP flood.

B. New generation botnets

The primary sources of DDoS attacks are botnets of various scales and feature sets. While a typical botnet is based on desktop computers, the security suites (including firewalls, virus, and intrusion detection systems) designed for desktop computers have evolved dynamically in the last decade. Accordingly, it became more challenging for hackers to infect a large number of computers with malicious codes. State-of-the-art desktop security suites typically incorporate a broad spectrum of protection features: anti-virus, web, email, user data, anti-hacking, and payment protections. Additionally, the processing power of the popular desktop processors enables to run of these detection features in real-time. Subsequently, there is a shift in the target of hackers towards alternative equipment with a lower security level, i.e., home, mobile, and IoT devices. In the last couple of years, numerous volumetric DDoS attacks approached or even exceeded the terabit-scale and originated from IoT botnets (Mirai-based botnets, as recent examples).

1) IoT-based botnets: The security protection of IoT devices is often overlooked by their developers due to strict delivery deadlines, lack of technical security background, or hardware cost. Moreover, the operating system of these devices is typically a stripped-down Linux distribution, omitting even the basic security subsystem. In addition, the generic Linux-based runtime environment enables attackers to effectively compile their malware codes to a broad spectrum of IoT devices. Considering IoT security, we should also focus on network-level defense beyond device-level security. From the networking perspective, IoT nodes like CCTV cameras routinely access the Internet with no rate limiting, which is an appealing feature for attackers. Since the IoT development life-cycle is relatively short, developers may reuse firmware codes or even web certificates and SSH keys. On the user-side, IoT equipment requires low maintenance, and they are considered deploy-and-forget devices. Thus, access passwords are often unchanged from the factory-default. These device-level shortcomings can be eliminated by setting up a strict network-level security and password policy specifically tailored to the deployed IoT device pool.

Ali et al. in [10] ”Systematic Literature Review on IoT-Based Botnet Attack” performed a systematic literature review including the state-of-the-art of IoT-based botnet attacks. This review paper revealed that research in this domain is gaining momentum, particularly in the last 3 years.

N. Koroniotis et al. in [11] ”Forensics and Deep Learning Mechanisms for Botnets in Internet of Things: A Survey of Challenges and Solutions” discusses the origin of botnets, overview the network forensic methods and focus on deep learning mechanisms and their roles in network forensics. Forensics of DDoS attacks is still a widely researched subject today; no standard method has been found, and most stakeholders are not interested in it. The main criterion of forensic research usefulness is how easy it is to deploy the system over the current Internet. Shi et al. and Ding et al. [12], [13] give a good overview of the current challenges and state-of-the-art.

T. S. Gopal et al. in [14] ”Mitigating Mirai Malware Spreading in IoT Environment” analyzed the Mirai malware in detail and presented its exploitation techniques. They proposed a white-listing method to prevent an IoT-based botnet from spreading.

H. -V. Le and Q. -D. Ngo in [15] ”V-Sandbox for Dynamic Analysis IoT Botnet” discuss the importance of sandbox environments in collecting behavior data from botnets in a secure way. They overview the limitations of the existing sandbox solutions and introduces the V-sandbox method for a dynamic analysis of IoT botnets. This proposal enables IoT botnet samples to reveal all of their malicious properties.

W. Li et al. in [16] ”Analysis of Botnet Domain Names for IoT Cybersecurity” discusses the role of the global DNS

service in supporting botnets to connect bots to C&C servers. To avoid tracking the C&C through the DNS information, botnets use sophisticated schemes such as fast-flux. Authors performed an in-depth analysis of the activities of Rustock botnet domain names, which use the fast-flux as the connection method between bots and C&C server.

R. Vinayakumar et al. in [17] "A Visualized Botnet Detection System Based Deep Learning for the Internet of Things Networks of Smart Cities" proposes a botnet detection system based on a two-level deep learning framework for semantically discriminating botnets and legitimate behaviors at the application layer of the domain name system (DNS) services. In the first level of the framework, the similarity measures of DNS queries are estimated using siamese networks based on a predefined threshold for selecting the most frequent DNS information across Ethernet connections. In the second level of the framework, a domain generation algorithm based on deep learning architectures is suggested for categorizing normal and abnormal domain names.

Y. Jia et al. in [18] "FlowGuard: An Intelligent Edge Defense Mechanism Against IoT DDoS Attacks" propose an edge-centric IoT defense scheme called FlowGuard for the detection, identification, classification, and mitigation of IoT DDoS attacks. They present a new DDoS attack detection algorithm based on traffic variations and design two machine learning models for DDoS identification and classification.

N. Ravi et al. in [19] "Learning-Driven Detection and Mitigation of DDoS Attack in IoT via SDN-Cloud Architecture" present a security scheme that leverages the cloud and software-defined network (SDN) paradigm to mitigate DDoS attacks on IoT servers. They have proposed a novel mechanism named learning-driven detection mitigation (LEDEM) that identifies DDoS using a semi-supervised machine-learning algorithm and mitigates DDoS. Authors tested LEDEM in the testbed, emulated topology, and compared the results with state-of-the-art solutions. They achieved an improved accuracy rate of 96.28% in detecting DDoS attacks.

2) Mobile-based botnets: Smart mobile phones can be considered as handheld computers with ever-increasing processing power and network bandwidth. An LTE or 5G mobile network enables the transmission of multiple 100 Mbps of data from a single mobile device. M. Eslahi et al. in [20] "MoBots: A new generation of botnets on mobile devices and networks" present an overview of mobile botnets, including studies on the new command and control mechanisms, actual examples, and malicious activities. N. Hoque et al. in [21] "Botnet in DDoS Attacks: Trends and Challenges" present a comprehensive overview of DDoS attacks, their causes, types with a taxonomy, and technical details of various attack launching tools. Authors give a detailed discussion of several botnet architectures and tools developed using botnet architectures. Moreover, the dominant Android mobile operating system has an approx. 72% market-share worldwide. The combination of high processing and networking capacities and a single highly prevalent OS platform made a mobile device an appealing target for hackers for many malicious purposes. Primarily, due to the less sophisticated security suites, attackers can remotely install malware codes to the mobile device. Mobile botnets are a group of unrelated mobile devices infected by a common botnet malware. The operational scheme is similar to that of the desktop-based variant; the botnet master remotely manages the botnet by a command and control mechanism to initiate a DDoS attack towards a target victim. Z. Lu et al. in [22] "On the Evolution and Impact of Mobile Botnets in Wireless Networks" adopt a stochastic approach to study the evolution and impact of mobile botnets. Authors find that node mobility can be a trigger to botnet propagation storms. They also reveal that mobile botnets can propagate at the fastest rate of quadratic growth in size, which is substantially slower than the exponential growth of Internet botnets. A. A. Santos et al. in [23] "A Stochastic Adaptive Model to Explore Mobile Botnet Dynamics" propose a stochastic adaptive model for the dynamics and the self-organized and self-adaptive behavior of mobile botnets to perform DDoS attacks.

Beyond the legacy command and control protocols (e.g., IRC, HTTP, and P2P), mobile-specific control mechanisms such as SMS-, MMS-, or Bluetooth-based variants have also emerged. The most challenging mobile botnet is the SMS- and P2P-based architecture in terms of detection complexity. E. Johnson and I. Traore in [24] "SMS Botnet Detection for Android Devices through Intent Capture and Modeling" investigated mobile botnets focusing on the Android operating system. Authors discuss a short messaging service (SMS) botnet structure and investigate a new detection model using the concept of intents. They show that transparent control can be achieved by a remote endpoint yet also detected by the proposed intent detection model.

C. Today's challenges

Increasing traffic volume requires ever more protective network resources. Volumetric attacks can quickly exhaust even the most considerable amount of Internet access capacity.

Shared botnets (many available for hiring): Hiring a botnet is a viable business option for botnet masters. In this model, hired resources are often accounted and paid for on a time basis. A major economic challenge here is a significant asymmetry in the expense of the attack and the defense. Renting botnet resources for a 10-minute attack costs as low as 35 cents [25].

Linux-based DDoS malware: The latest Windows versions enable running a complete Linux run-time environment on a Windows-based laptop or desktop computer. This feature opened the possibility for malware authors to cross-compile botnet code to run on both Windows and Linux systems. This option raises crucial challenges in the defense strategy: i) a high number of IoT devices with common security vulnerabilities run a Linux-based operating system, ii) Linux-based datacenter servers possess a high amount of computational and bandwidth resources to execute a heavy-hitter DDoS attack.

Launching attacks by non-technical users: Volumetric attacks can be initiated with dedicated control programs and scripts available on the darknet or offered to the attacker by

the bot master of the rented botnet. These tools are easy to use; therefore, even a non-technical user can initiate and control a powerful attack.

Attack from mobile and IoT equipment: The increasing computational power of handheld devices and the transmission capacity of 4G and 5G networks open the way to deliver a wide range of botnet malware to mobile devices. Moreover, mobile security suites typically have a lower level of defense against malware deployment. Thus, handheld devices may become the next target of the bot master (a person who owns the botnet). Reputation-based detection is inefficient for identifying infected mobile devices since user equipment's IP addresses frequently change in mobile communication networks.

Hit-and-run and multi-vector attacks continue to evolve. H-a-R is still popular due to its low cost and ease of deployment. At the same time, multi-vector variants are very effective in bypassing traditional mitigation strategies. Recently, we have seen a significant rise in the popularity of multi-vector attacks incorporating 15 or more vectors. Combining the hit-and-run and multi-vector strategies resulted in a shorter attack duration with an increased success rate. Since attackers often rent a shared botnet to execute the DDoS attack, their ambition to reduce the duration is reasonable. Moreover, the shortened attack has a higher probability of bypassing security systems with a larger detection window.

Browser-based bot attacks: Websites are attractive platforms to deliver malware to a high number of user devices via popular web browsers. Javascript-based codes do not depend on the operating systems and exploit the web browsers' vulnerabilities. While these codes stop running as the user quits the browser application, they are re-downloaded and re-initialized as one re-visits the compromised web page.

Emerging encrypted attacks. TLS- and ESP-based attacks have two key advantages: i) they consume extra CPU resources to perform encryption and decryption, ii) many DDoS detection systems do not support the inspection of TLS- and ESP-encrypted traffic.

Distributed targets: From the infrastructural perspective, popular cloud-based services are distributed across many physical servers, and many of them are often located in dedicated IP subnets. Instead of attacking a single IP node, this type of DDoS threat increases the success rate by targeting an entire IP subnet incorporating a set of servicing nodes.

Application-specific attacks: The majority of application-specific attacks target a specific service and not a service type in general, e.g., developed to attack a specific streaming service. It means that no attack is capable of targeting streaming services universally. Meanwhile, a recent method called mimicked user browsing is very effective for a large-scale of web applications. It is a web-based application-specific attack type developed to imitate the behavior of real user interaction with the service provider nodes. The major challenge is its low false rate detection since its traffic pattern is identical to that of a real user. Due to the similarity property, it can easily maintain its success rate even using a large number of participating botnet nodes.

D. Lessons learned in DDoS challenges

Recent research works propose several methods and tools for effectively detecting the new-generation DDoS attack types (see Section IV). However, a new breed of attack techniques (especially the combination of hit-and-run and multi-vector attacks) still challenges protection systems with a more sophisticated traffic pattern combined with a large traffic volume within a very short time period. Besides the new types of network layer attacks, the mimicked user browsing attack targets a specific service with a high success rate. Moreover, shared low-cost botnets create a high resource and economic imbalance between the expense of the attack and the defense. In Section IV, we discuss the major scientific works for detecting the presented threat types.

IV. DETECTION OF NEW GENERATION DDOS THREATS

A. Hit-and-run

The so-called hit-and-run (or shrew) DDoS attacks are attacks that operate with multiple high throughput short bursts, [26], [27]. These attacks are dangerous because: i.) They cause significant quality of service degradation through TCP congestion control, ii.) many DDoS detection engines cannot identify them, iii.) even if they are detected if there is a human operator in the decision-loop, for her, the number of signals can be overwhelming. i.) Network equipment has relatively small intermediary buffers that can be saturated in less than an ms. Saturated buffers imply packet loss. After the initial packet loss(es), the TCP connection's congestion control throttles the connection speed. After this event, the TCP connection will need seconds to recover to the pre-loss throughput. This kind of QoS drop was very hard to quantify in the past, so operators ignored unconventional hit-and-run attacks. Aleksandar Kuzmanovic and Edward W. Knightly did the first research on this subject; they published their results in "Low-Rate TCP-Targeted Denial of Service Attacks: The Shrew vs. the Mice and Elephants" [28]. In this paper, they proved that a DDoS attack that delivers its payload in short bursts significantly affects the throughput of TCP flows. Kuzmanovic's method was relatively complex and had a high margin of error; thus, it wasn't used much. Since then, this kind of DDoS have become the most researched subject in the field because it lacks the throughput footprint of regular DDoS [29], [30], [31]. In "A Way to Estimate TCP Throughput under Low-Rate DDoS Attacks: One TCP Flow" [32] Kieu et al. propose a precise and straightforward method to quantify the damage caused by unconventional low-throughput or hit-and-run attacks. Kieu proves that their method is accurate using the NS-2 simulator. ii.) The detection time of the DDoS detector is in the range of seconds, which is longer than the time required to disrupt TCP flows. If the IDS doesn't have detection times in the ms range, it will always lag after the effect of the attack. iii.) The human-in-the-loop is a multiple way inadequate to deal with hit-and-run attacks. The time needed to make a human decision is multiple orders of magnitude longer than the duration of a DDoS burst that can successfully disrupt the TCP

TABLE I
OVERVIEW: THE NEW BREED OF DDOS ATTACKS

Attack type	Key characteristics	Special features
Amplification/Reflection	Spoofed request to a set of public servers triggers responses toward the targeted system	Cumulative response traffic can saturate the network path of the vicim host
Hit-and-Run (Volatile)	Periodic On/Off strategy	Short On period (form milliseconds to minutes) followed by an extended Off period
Multi-vector	Combines multiple methods and attack types (most popular are: DNS reflection, TCP smart attacks and ICMP flood) into a single attack	Challenging mitigation: complex traffic pattern blended easily to normal traffic
Linux-based botnet	Cross compiled botnet code to run in Linux systems	Infection targets: IoT devices and data center servers
Mobile botnet	With 72% of market share, Android devices are in the focus of botnet malware. Key method is remotely install the malware code.	Increasing processing power and network bandwidth of mobile devices
Browser-based	Exploits vulnerability of web browsers to deploy malware JavaScript codes	Code stops running as user quit the browser application and therefore it requires a re-visit of the compromised site.
Distributed targets	Attack targets a set of server nodes within a subnet	Physical servers behind a cloud-base service are typically located in a dedicated data center subnet. Attacking a set of servers increases the success rate.
Mimicked user browsing	An application-specific attack aiming to replicate the behavior of real user interaction with the service provider nodes	Detection is challenging even when a large number of botnet nodes are participating in the attack

flows. If the attack changes some parameters (IP/port/protocol) between bursts, each burst will generate a discrete detection signal. Authors identified attacks that operated with changing parameters by generating more than 2000 signals per day for weeks. Such a high number of signals cannot be efficiently processed and validated by a human operator.

To successfully mitigate the effect of hit-and-run attacks, the network has to be changed, or the IDS has to: a.) detect and mitigate attacks within milliseconds, b.) have an acceptably low false detection rate to work without human validation.

There is a relatively small number of research results on this subject.

In "Low-rate TCP DDoS Attack Model in the Southbound Channel of Software Defined Networks" [33], Balarezo et al. showcase how low-rate DDoS attacks can exploit TCP congestion control to cause significant QoS drop in SDN networks. They propose a method to model the attacks and their effects in SDN.

In "On a Mathematical Model for Low-Rate Shrew DDoS" [34], Luo et al. present a new, more accurate analytical method to model the effect of a wide variety of hit-and-run attack patterns. This method aims to be significantly more accurate than current state-of-the-art methods. It reduces the average margin of error from 69% to 10% for most network environments and attack patterns. By making accurate models and understanding how network environments and attack patterns determine the effect of attacks, they managed to build a novel defense method against hit-and-run attacks. The proposal significantly reduces the impact of the attack.

In "Stability of TCP/AQM Networks Under DDoS Attacks With Design" [35], Tan et al. propose a method to tweak TCP active queue management to mitigate the effect of hit-and-run attacks on TCP congestion control. The results of this research are promising since they prove that TCP throughput can be stabilized at an acceptable level during an attack without sacrificing anything else or adding new network components.

In "A new network flow grouping method for preventing periodic shrew DDoS attacks in cloud computing" [36], Liu et al. propose a new method to extend the usability of the BIRTH algorithm in detecting shrew (hit-and-run) attacks. The primary deficiency of BIRTH is its long detection time, which makes it hardly usable against hit-and-run attacks. By clustering and re-merging traffic using flow-level frequency domain characteristics, this method appears to significantly improve the detection time of the BIRTH algorithm.

In "An optimized design of reconfigurable PSD accelerator for online shrew DDoS attacks detection" [37] Chen et al. propose the idea of abandoning the time-domain approach in favor of frequency domain analysis. It is a logical step because the main difficulty of detecting hit-and-run threats is the short length of the attacks, e.g., the detection window in the time-domain is short. Meanwhile, in the frequency domain, the energy of the attack is unmaskable. They use FPGA hardware to implement a DFT (Discrete Fourier Transformation) algorithm complemented by their auto-correlation algorithm. This approach proves to be significantly more efficient than the regular approach.

In "Low-Rate DoS Attack Detection Using PSD Based Entropy and Machine Learning" [38], Zhang et al. propose a novel method using supervised learning on frequency domain data. This appears to be an efficient approach because the time-domain problem can be eliminated completely, and ML provides a robust framework to detect a wide variety of attacks.

The future of detecting hit-and-run attacks seems to be in the frequency domain. The need for ms range detection is eliminated by performing frequency domain analysis.

B. Distributed targets- carpet bombing attacks

There is a new trend of attacking distributed cloud services by not just attacking the public-facing IP but all the physical servers of the service. This makes attacks more challenging to mitigate since multiple attacks have to be handled in the same time. Apart from this difference, the mitigation of these attacks is identical to the mitigation of regular attacks. There is no research on this subject because the mentioned difference is not a scientific but an engineering challenge, and still only relatively few services are out there worth being attacked distributively.

C. QUIC-based DDoS

QUIC-based DDoS attacks represent a significant evolution in the landscape of distributed denial-of-service threats, leveraging the unique characteristics of the QUIC (Quick UDP Internet Connections) protocol. Developed as an alternative to the traditional TCP/IP model, QUIC offers faster and more secure data transmission over the Internet. In a QUIC-based DDoS attack, attackers exploit these properties to overwhelm target servers with a high volume of encrypted requests, making detection and mitigation challenging. The encryption masks the malicious traffic, blending it with legitimate requests. One can examine the currently publicly available QUIC-based attack tools at [39]. The effects and prevention of QUIC DDoS are not yet a very well-researched subject. There are only researches studying the difference between QUIC and TCP/TLS-based attacks [40], [41], [42].

D. Application

The detection (and mitigation) of application-based attacks mostly rely on the application developer instead of a universal security solution provider. The main reason for this is that application-based attacks do not follow any universal rules, which can be observed through a wide variety of attack types because application attacks exploit very specific application-related vulnerabilities. The detailed problems, challenges, and solutions of app DDoS are described in this survey [43]. There are proposed universal methods detecting these attacks [44], [45]. These methods mainly utilize machine learning or entropy-based methods. With the help of security professionals, these vulnerabilities can be eliminated or at least mitigated by the app developer. The mimicked user browsing attack is the newest ”widespread” (still very uncommon compared to universal volumetric attack types) application-based attack. For mimicked user browsing, a victim can take two routes: i) use universal anti-bot services like Google capcha or ii) use machine learning to profile their traffic and identify irregular attack traffic. Recently, a novel attack type called DNS Water Torture [46] appeared in the toolkit of adversaries. It is an application layer attack that overloads the targeted DNS servers with a high volume of fraudulent domain request messages. Often, DNS water torture attacks are combined with more common DDoS attacks. By overshadowing the application layer attack, mitigating with first-line security defense is more challenging.

E. Browser-based bot attacks

The dominant browser-related threat is a malware packed into a browser extension [47]. The main benefit of using the extension framework to execute malicious codes is twofold: i) one-time download, ii) JavaScript-based portable code. In contrast to malware downloaded via compromised websites, extension-based variants reload to the system memory each time the user starts the browser. These benefits make the browsers an appealing target for criminals. Often, while the malware function runs silently in the background, the extension also provides a valuable function to the users. The primary concern is that an extension may have a privilege to access and manipulate the DOM (Document Object Model) of a web page, user’s browsing history, bookmarks, or even files on the local storage system. Meanwhile, browser developers have a constant effort to make extension APIs stricter and more secure.

F. Multi-vector

A multi-vector attack combines multiple techniques shown in Fig. 4 to increase its success rate. Furthermore, the incorporating vectors may have unique timing properties to switch on and off or rise and fall the traffic volume. This feature enables to construct a wide variety of attack scenarios that are easy to re-organize by the attackers. The benefit of applying multiple vectors are two-fold [48]: i) the traffic volume of the individual vectors is additive, ii) the generated traffic pattern can reach higher complexity and, thus, is more effectively blended to the regular user traffic. The most popular vectors are volumetric type, i.e., DNS/NTP amplification, UDP flood, Chargen, and SSDP. Often, TCP Syn or application-specific vectors are also added to the vector mix, [49], [50], [51], [52]. Besides being automated, the most sophisticated attacks dynamically adjust the parameters of the individual vectors in response to the applied mitigation strategy. The intelligent control of the vectors allows attackers to tailor attacks to be shorter (typically in the 10-minute order) and yet more effective.

G. Lessons learned in DDoS detection

The emergence of new DDoS threats provides new challenges for both researchers and solution providers. The most concerning new trend is the emergence of encrypted attacks, including QUIC-based threats, which are very hard to detect. Application-based attacks became very sophisticated; the detection and mitigation of these attacks relied mostly on payload inspection, which becomes impossible with encryption. Likely, aggregated metadata inspection will become more prevalent in this field. Multi-vector and volatile attacks are not as novel as encrypted attacks. Still, their maturity and real-world share are very concerning.

For attacks that apply the hit-and-run method, detection algorithms should focus on short-time high-intensity bursts combined with an on-off traffic pattern. The novel time-domain behavior claims for algorithms with low false rate without human validation. The scope of potential botnet sources is

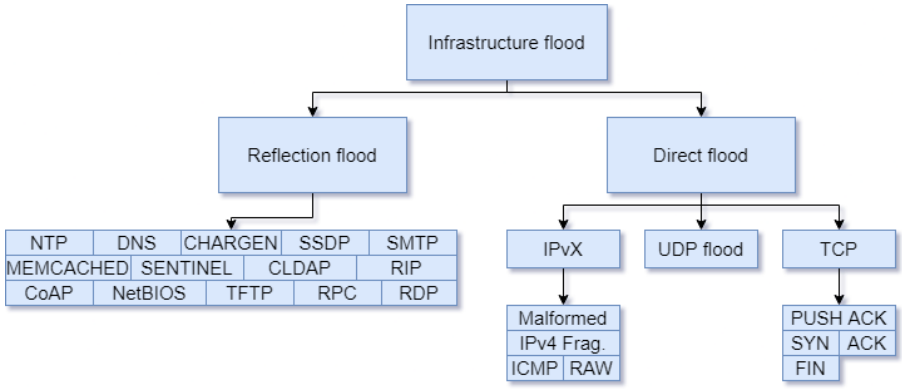


Fig. 4. The taxonomy of flood-based DDoS attacks

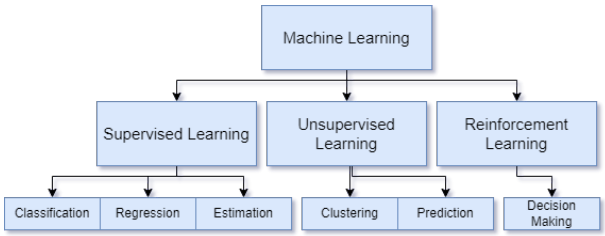


Fig. 5. The high level breakdown of machine learning methods

also significantly extended with a large number of IoT and mobile devices. These new types of botnets incorporate novel infection and attack strategies as well. The latest research works focus on the potential of deep learning to adapt to the new attack patterns (see Table II). Furthermore, in the last couple of years, we have seen a significant rise in the popularity of multi-vector attacks incorporating 10+ vectors. The primary protection challenge is that each incorporating vector should be individually detected and mitigated.

V. SUPPORTING DDOS DETECTION WITH ML

The detection of DDoS attacks is traditionally done by rule-based or heuristic software running on CPUs. Machine Learning (ML) is quite a broad subject; it is categorized into three main categories (see Fig. 5): supervised learning, where the machine is taught with inputs for which the correct output is known, unsupervised learning, where there is only input data, no information about the expected outcomes, reinforcement learning, where reward function is known. ML-supported DDoS detection became a viral subject for researchers in the past decade because ML has the potential to solve two major research-development gaps that are hard-to-impossible to solve using rule-based or heuristic detection methods: 1.) Detecting novel zero-day attacks automatically, 2.) Detect non-malicious anomalous network events (not-scope of this paper). While machine learning holds the promise to build a universal so-called ”Silver Bullet” system, there are significant challenges. The major drawback of machine learning is false positive detection. False positive detection is a serious issue because blocking the traffic of paying customers has more severe consequences than letting an attack pass through. For this reason, ML-based detection has not achieved major industrial success yet.

DDoS detection was most studied from the ML perspective in the past three years (see Table III). There is a plethora of research from this perspective. This section only draws a broad picture of how ML accelerates DDoS detection and mitigation while focusing on the two mentioned research gaps.

A. Detecting novel attacks with ML

Scaranti et al., in ”Artificial Immune Systems and Fuzzy Logic to Detect Flooding Attacks in Software-Defined Networks” [53], propose a novel AIS-based defense architecture for SDN systems. This system can detect and mitigate multiple types of DDoS attacks with minimal false detection (less than 1%). Scaranti et al. concept and results are imposing because they solved the issue of a high false detection rate while being able to detect previously unknown attacks, and they verified their system on publicly available datasets.

Poongodi et al., in ”DDoS Detection Mechanism Using Trust-Based Evaluation System in VANET” [54], propose a novel method to segregate DDoS attackers. This method is based on trust and clustering. This method has two main benefits: 1.) It is resource efficient, 2.) It can be scaled very well. Their system is benchmarked against the AODV protocol and Firecol technique. The method developed by Poongodi et al. is significantly better in the achieved goodput, latency, and energy consumption than the other two state-of-the-art methods.

Nezhad et al. in ”A Novel DoS and DDoS Attacks Detection Algorithm Using ARIMA Time Series Model and Chaotic System in Computer Networks” [55] propose a novel method (TNA) to amend the main backdraw of ARIMA (auto-regression). They combine multiple previously known methods, including Box-Cox, Lyapunov, and chaotic error detection, to increase the detection rate. They successfully enhance the detection rate on large data sets to 99.5%, which is 1.1% higher than the previous best-known algorithm.

TABLE II
CURATED OVERVIEW OF ARTICLES ON THE TOPIC OF NEW GENERATION DDOS THREATS

Author(s)	Reference	Threat type	Novelty	Results
Ali et. al.	[10] in III-B1	IoT-based botnet attack	Systematic literature review	Focusing on research works of the recent years
N. Koroniotis et. al.	[11] in III-B1	IoT-based botnets	A survey of forensics and deep learning mechanisms	Overviews deep learning-based network forensic methods
M. Eslahi et. al.	[20] in III-B2	Mobile botnet	Overview of the novel command and control mechanisms and their malicious activities	Reviews the limitations of botnet detection in mobile environment
N. Hoque et. al.	[21] in III-B2	Mobile botnet	A survey of various botnet architectures and tools	Pros and cons analysis
Z. Lu et. al.	[22] in III-B2	Mobile botnet	Impact of mobile botnets on wireless networks	Node mobility can trigger a botnet propagation storm. Mobility range over a threshold enables the botnet to growth quadratically (theoretical maximum). Comparing to the exponentially expanding Internet botnets, it is significantly slower mechanism.
Kuzmanovic et. al.	[28] in IV-A	Hit&Run	Analysis of Hit&Run	Demonstration of distruptiveness of hit&run attacks
Kieu et. al.	[32] in IV-A	Hit&Run	Analysis and simulation of Hit&Run	Demonstration of distruptiveness of hit&run attacks
Balarezo et. al.	[33] in IV-A	Hit&Run	Congestion analysis	Demonstration of distruptiveness of hit&run attacks
Luo et. al.	[34] in IV-A	Hit&Run	Mathematical model of Hit&Run attacks	Very high precision model for wide variety of attacks
Tan et. al.	[35] in IV-A	Hit&Run	TCP congestion control algorithm	Resilient congestion control algorithm
Teyssier et. al.	[40] in IV-C	QUIC	Attack evaluation	Attack effectiveness against QUIC evaluation method
Balaji et. al.	[41] in IV-C	QUIC	Attack method	Showcase of novel QUIC-based attack
Wang et. al.	[44] in IV-D	Application	Detection method	Novel universal entropy-based L7 detection method
Yadev et. al.	[45] in IV-D	Application	Detection method	Novel universal ML-based L7 detection method
Perotta et. al.	[47] in IV-E	Browser based	Case study	Study of the detection challenges of attacks originating from browsers
Dimolianis et. al.	[48] in IV-F	Multi-vector	Mitigation method	Novel method to mitigate and detect multi-vector attacks

Simpson et al., in ”Per-Host DDoS Mitigation by Direct-Control Reinforcement Learning” [56], propose a new mitigation method based on reinforced machine learning (RL). Regular machine learning has a hard time keeping up with the constantly changing patterns of DDoS attacks. By monitoring the result of the mitigation and using it to reinforce the per-flow decision-making, they achieve increased goodput compared to the state-of-the-art.

These four papers illustrate well how ML can be used to detect previously unknown attacks. This is the single most significant achievement of ML in this field from the industrial point of view.

B. ML detection on a small footprint

The scope of this paper is to discuss the DDoS attacks threatening DCN and ISP networks. Still, there is relevant research outside the DCN scope that can and should be applied to this subject as well. One of the main problems of ML-based DDoS detection is its relatively high resource utilization. This problem becomes a vital issue, even in DCNs, when detection is extended to east-west routes. In this subsection, we will showcase ML methods from resource-sensitive fields (IoT, VANET) where these methods have been implemented with a minimal footprint.

Kim et al. in ”Intelligent Application Protection Mechanism for Transportation in V2C Environment” [57] proposes a novel image-based system resource monitoring AI for DDoS detection in Vechile-to-cloud (V2C) systems. V2C systems are not safety-critical, but there has been no previous research on the safety of these systems. This kind of AI can be a great fit for IoT or distributed systems because this AI does not sample the traffic but the system’s resource utilization. This approach is extremely resource-efficient, but a significant detection lag exists. By combining the memory, CPU, and network utilization, they managed to achieve a 7.36% false detection rate.

Gao et al. in ”A Distributed Network Intrusion Detection System for Distributed Denial of Service Attacks in Vehicular Ad Hoc Network” [59] propose a novel massively-scalable DDoS detection system for vehicular ad-hoc networks (VANET). The system proposed by Gao is partitioned into subsystems: 1.) Real-time traffic collection subsystem, 2.) Spark-based attack detection subsystem. The detection system was moved into the cloud to access computing resources and aggregate the traffic of detected attacks. This approach solves the cost-sensitive nature of VANET nodes (Vehicles), and by using Big-data resources, it can approach very low (0.05%-1%) false detection ratios.

Yang et al. in ”Adaptive Measurements Using One Elastic Sketch” [58] propose a novel method, called Elastic Sketch, to measure the network during attacks. The main advantage of using Elastic Sketch is that it can adapt very well to rapidly changing network conditions. Elastic sketch has a 50 times shorter measuring speed than the current state-of-the-art sketch and a much lower error rate.

Xiao et al. in ”IoT security techniques based on machine learning: How do IoT devices use AI to enhance security?” [60] identify IoT attack models and propose defense methods

TABLE III
OVERVIEW OF ARTICLES ON THE TOPIC OF ML ACCELERATION

Author(s)	Reference	Accelerator	Novelty	Results
Scaranti et. al.	[53]	ML	AIS-based method	Less-than 1% false detection, ability to detect zero-day attacks
Kim et. al.	[57]	ML	Image-based method	Extremly low resource usage, on a large data-set
Nezhad et. al.	[55]	ML	Auto-regression-based method	Very-low false detection
Yang et. al.	[58]	ML	New algorithm	50 times shorter detection time, than SOTA sketch

against them. They showcase how ML must meet unique challenges if applied in the IoT security scene.

C. Lessons learned

There is a growing demand from users to integrate ML into DDoS protection systems, which has not been done so far by most industrial solution providers. The reason for this is three-fold i.) validating security systems is a very resource-intensive task (why should a provider spend immense resources on a method that is not proven to be effective on an industrial scale), ii.) establishing causality links between decisions and data is crucial for any security system not solved on the research level (eXplainable AI, root cause analysis, etc.), iii.) patching false detections is not an easy task in an ML pipeline.

The authors of this paper believe that the integration of ML into industrial DDoS detection will not be quick but inevitable. We predict that the first ML-based DDoS detection will be utilized to detect zero-day attacks, and eventually, more and more traditional algorithms will be superseded by ML-based methods.

VI. THE FUTURE OF DDOS

This section summarizes what we, as researchers and industrial solution providers, experience about the latest DDoS trends and the solutions to these new challenges, and we try to predict the future direction of this topic. We also looked into what global security solution providers see and forecast for the future [61] [62]. We found that our observations and predictions match their reported trends. In previous sections, we demonstrated that DDoS attacks are evolving at an unprecedented speed, with the following main characteristics:

TABLE IV
THE MOST PREVALENT CURRENT DDOS TRENDS AND THEIR EXPECTED FUTURE RELEVANCE. THE CURRENT RELEVANCE WAS MEASURED BY THE AUTHORS IN PROTECTED DCNS. A TWO-YEAR LONG CONTINUOUS MEASUREMENT PERIOD IS THE BASIS OF OUR ESTIMATION FOR THE FUTURE.

Trend nr.	Share 2022	Expected Share 2024	Challenge
I.	60%	90%	Engineering
II.	75%	85%	Engineering
III.	3%	20%	Engineering
IV.	3%	15%	Research
V.	1%	10%	Research

I.) The attacks' duration and ramp-up period become shorter and shorter while the peak throughput of the same attacks increases. **Solution:** The mitigation process has to be fully automated. Per-packet analysis has to be used for detection. The human reaction time is not fast enough to mitigate DDoS attacks reliably in under two minutes. Solution providers must provide highly reliable solutions that can be trusted as active devices. NetFlow and other flow aggregation-based DDoS detection methods have an aggregation period of a few minutes, which induces an intolerable mitigation lag. In contrast, per-packet traffic analysis can provide highly detailed attack insight in ms-s.

II.) Multi-vector attacks became the new norm. **Solution:** Multi-vector attacks can be mitigated with black-hole routing. Suppose we want a little bit more sophisticated mitigation, which can protect the user as well as the network. In that case, every attack vector must be analyzed and mitigated separately. So we need algorithms that not only detect attacks, but classify them on a vector-level resolution.

III.) One of the scientifically most exciting frontiers of DDoS research is the protection of IoT networks. The adoption of 5G networks boosts the number and significance of IoT devices; critical infrastructures adopt the IoT approach, like vehicular networks, thus making the protection of IoT more critical than ever. Meanwhile, IoT devices still do not have the resources necessary for straightforward DDoS protection. Currently, researchers are working on two tracks, developing alternative methods like [63], [64], [65], or extending protection at the 5G packet gateway, like [66], [67].

IV.) User/application mimicking DDoS attacks became a measurable (1-5%) share of all attacks. **Solution:** New methods of attack detection have to be developed by researchers, which can detect attacks and use the historical context of the end-points' regular traffic to detect these new smart attacks. In this field, per-endpoint-based unsupervised learning shows the greatest promise, but no industrial-grade solutions have been provided.

V.) The East-West attacks became a measurable (0.5-1%) share of all attacks. Most DDoS detection solutions monitor only the north-south links of the DCN. The current most common application architectures can not be scaled to cover every possible route between tenants. A very conservative estimate for the protecting cost of a 1000MW DCN on every east-west route with an industry-standard active inline DDoS mitigation device would be 1-2 billion USD annually. **Solution:** New data-collection schemes have to be devised by researchers to collect data, which could be used to feed the next generation of DDoS detection systems.

After reading this, one could ask themselves: What does the future hold in the next few years? Our guess is, according to Table IV, that Trend III and Trend IV will become much more prevalent (5-15%), making these challenges unavoidable

by the SotA providers. Trend I and Trend II became the normal method of DDoS attacks making up 80+% of all attacks.

VII. CONCLUSIONS

In this survey, we discussed how DDoS attacks prove to be continuously evolving prevalent threats. First, we defined the basic terminology to navigate the landscape of post-pandemic DDoS. We presented how little the DDoS attack scene of the post-Covid world resembles the attacks of 5 years ago, providing examples of novel attacks, methods, and tools. We provided a survey of attack profiles prevalent in the network of Hungarian ISPs. After this, we discussed state-of-the-art research considering the detection of DDoS attacks. As part of this, we summarized the research considering acceleration schemes and discussed the rich literature on machine learning-based methods, their benefits, and their challenges.

The key lesson to be learned through this paper is that DDoS attacks might be considered old brute-force methods, but plenty of new threats are worthy of research. With the increasingly widespread QoS-sensitive applications, the multivector *ephemeral* – short and high volume – attacks became the new norm, making human-in-the-loop systems nearly obsolete.

The research has been partially funded by the Hungarian government's National Security Cooperative PhD program.

REFERENCES

[1] Cisco White Paper, "Cisco annual internet report (2018–2023) white paper," https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html, 2020, accessed: 10-01-2023.

[2] Akamai, "2021: Volumetric ddos attacks rising fast," https://www.akamai.com/blog/security/2021-volumetric-ddos-attacks-rising-fast, 2021, accessed: 10-01-2023.

[3] T. Peng, C. Leckie, and K. Ramamohanarao, "Survey of network-based defense mechanisms countering the dos and ddos problems," vol. 39, no. 1, 2007. [Online]. Available: DOI: 10.1145/1216370.1216373

[4] S. T. Zargar, J. Joshi, and D. Tipper, "A survey of defense mechanisms against distributed denial of service (ddos) flooding attacks," *IEEE Communications Surveys Tutorials*, vol. 15, No. 4, pp. 2046–2069, 2013. [Online]. Available: DOI: 10.1109/surv.2013.031413.00127

[5] M. Masdari and M. Jalali, "A survey and taxonomy of dos attacks in cloud computing," *Security and Communication Networks*, vol. 9, no. 16, pp. 3724–3751, 2016. [Online]. Available: DOI: 10.1002/sec.1539

[6] Q. Yan, F. R. Yu, Q. Gong, and J. Li, "Software-defined networking (sdn) and distributed denial of service (ddos) attacks in cloud computing environments: A survey, some research issues, and challenges," *IEEE Communications Surveys Tutorials*, vol. 18, no. 1, pp. 602–622, 2016. [Online]. Available: DOI: 10.1109/comst.2015.2487361

[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. [Online]. Available: DOI: 10.1109/JSAC.2018.2869997

[8] S. Dong, K. Abbas, and R. Jain, "A survey on distributed denial of service (ddos) attacks in sdn and cloud computing environments," *IEEE Access*, vol. 7, pp. 80 813–80 828, 2019. [Online]. Available: DOI: 10.1109/ACCESS.2019.2922196

[9] J. Zheng, Q. Li, G. Gu, J. Cao, D. K. Y. Yau, and J. Wu, "Realtime ddos defense using cots sdn switches via adaptive correlation analysis," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 7, pp. 1838–1853, 2018. [Online]. Available: DOI: 10.1109/TIFS.2018.2805600

[10] I. Ali, A. I. A. Ahmed, A. Almogren, M. A. Raza, S. A. Shah, A. Khan, and A. Gani, "Systematic literature review on iot-based botnet attack," *IEEE Access*, vol. 8, pp. 212 220–212 232, 2020. [Online]. Available: DOI: 10.1109/ACCESS.2020.3039985

[11] N. Koroniotis, N. Moustafa, and E. Sitnikova, "Forensics and deep learning mechanisms for botnets in internet of things: A survey of challenges and solutions," *IEEE Access*, vol. 7, pp. 61 764–61 785, 2019. [Online]. Available: DOI: 10.1109/ACCESS.2019.2916717

[12] L. Shi, J. Li, M. Zhang, and P. Reiher, "On capturing ddos traffic footprints on the internet," *IEEE Transactions on Dependable and Secure Computing*, vol. 19, no. 4, pp. 2755–2770, 2022. [Online]. Available: DOI: 10.1109/TDSC.2021.3074086

[13] D. Ding, M. Savi, and D. Siracusa, "Tracking normalized network traffic entropy to detect ddos attacks in p4," *IEEE Transactions on Dependable and Secure Computing*, vol. 19, no. 6, pp. 4019–4031, 2022. [Online]. Available: DOI: 10.1109/TDSC.2021.3116345

[14] T. S. Gopal, M. Meerolla, G. Jyostna, P. Reddy Lakshmi Eswari, and E. Magesh, "Mitigating mirai malware spreading in iot environment," in *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 2018, pp. 2226–2230. [Online]. Available: DOI: 10.1109/ICACCI.2018.8554643

[15] H.-V. Le and Q.-D. Ngo, "V-sandbox for dynamic analysis iot botnet," *IEEE Access*, vol. 8, pp. 145 768–145 786, 2020. [Online]. Available: DOI: 10.1109/ACCESS.2020.3014891

[16] W. Li, J. Jin, and J.-H. Lee, "Analysis of botnet domain names for iot cybersecurity," *IEEE Access*, vol. 7, pp. 94 658–94 665, 2019. [Online]. Available: DOI: 10.1109/ACCESS.2019.2927355

[17] R. Vinayakumar, M. Alazab, S. Srinivasan, Q.-V. Pham, S. K. Padannayil, and K. Simran, "A visualized botnet detection system based deep learning for the internet of things networks of smart cities," *IEEE Transactions on Industry Applications*, vol. 56, no. 4, pp. 4436–4456, 2020. [Online]. Available: DOI: 10.1109/TIA.2020.2971952

[18] Y. Jia, F. Zhong, A. Alrawais, B. Gong, and X. Cheng, "Flowguard: An intelligent edge defense mechanism against iot ddos attacks," *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 9552–9562, 2020. [Online]. Available: DOI: 10.1109/JIOT.2020.2993782

[19] N. Ravi and S. M. Shalinie, "Learning-driven detection and mitigation of ddos attack in iot via sdn-cloud architecture," *IEEE Internet of Things Journal*, vol. 7, no. 4, pp. 3559–3570, 2020. [Online]. Available: DOI: 10.1109/JIOT.2020.2973176

[20] M. Eslahi, R. Salleh, and N. B. Anuar, "Mobots: A new generation of botnets on mobile devices and networks," in *2012 International Symposium on Computer Applications and Industrial Electronics (ISCAIE)*, 2012, pp. 262–266. [Online]. Available: DOI: 10.1109/ISCAIE.2012.6482109

[21] N. Hoque, D. K. Bhattacharyya, and J. K. Kalita, "Botnet in ddos attacks: Trends and challenges," *IEEE Communications Surveys Tutorials*, vol. 17, no. 4, pp. 2242–2270, 2015. [Online]. Available: DOI: 10.1109/COMST.2015.2457491

[22] Z. Lu, W. Wang, and C. Wang, "On the evolution and impact of mobile botnets in wireless networks," *IEEE Transactions on Mobile Computing*, vol. 15, no. 9, pp. 2304–2316, 2016. [Online]. Available: DOI: 10.1109/TMC.2015.2492545

[23] A. A. Santos, M. Nogueira, and J. M. F. Moura, "A stochastic adaptive model to explore mobile botnet dynamics," *IEEE Communications Letters*, vol. 21, no. 4, pp. 753–756, 2017. [Online]. Available: DOI: 10.1109/LCOMM.2016.2637367

[24] E. Johnson and I. Traore, "Sms botnet detection for android devices through intent capture and modeling," in *2015 IEEE 34th Symposium on Reliable Distributed Systems Workshop (SRDSW)*, 2015, pp. 36–41. [Online]. Available: DOI: 10.1109/SRDSW.2015.21

[25] NETSCOUT, "Netscout threat intelligence report," vol. 5, 2020. [Online]. Available: https://www.netscout.com/

[26] B. Nagy, "Chargen udpfrag tcp syn multivector ddos attack," Oct 2021. https://zenodo.org/record/5578700.

[27] ——, "Encrypted traffic with esp and tls," Oct 2021. https://zenodo.org/record/5578676.

[28] A. Kuzmanovic and E. W. Knightly, "Low-rate tcp-targeted denial of service attacks: The shrew vs. the mice and elephants," in *Proceedings of the 2003 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*, ser. SIGCOMM '03. New York, NY, USA: Association for Computing Machinery, 2003, pp. 75–86. [Online]. Available: DOI: 10.1145/863955.863966

[29] K. Hong, Y. Kim, H. Choi, and J. Park, "Sdn-assisted slow http ddos attack defense method," *IEEE Communications Letters*, vol. 22, no. 4, pp. 688–691, 2018. [Online]. Available: DOI: 10.1109/LCOMM.2017.2766636

[30] J. A. Pérez-Díaz, I. A. Valdovinos, K.-K. R. Choo, and D. Zhu, "A flexible sdn-based architecture for identifying and mitigating low-rate ddos attacks using machine learning," *IEEE Access*, vol. 8, pp. 155 859–155 872, 2020. [Online]. Available: DOI: 10.1109/ACCESS.2020.3019330

[31] A. Praseed and P. S. Thilagam, "Multiplexed asymmetric attacks: Next- generation ddos on http/2 servers," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 1790–1800, 2020. [Online]. Available: DOI: 10.1109/TIFS.2019.2950121

[32] M. V. Kieu, D. T. Nguyen, and T. T. Nguyen, "A way to estimate tcp throughput under low-rate ddos attacks: One tcp flow," in *2020 RIVF International Conference on Computing and Communication Technologies (RIVF)*, 2020, pp. 1–8. [Online]. Available: DOI: 10.1109/RIVF48685.2020.9140777

[33] J. F. Balarezo, S. Wang, K. G. Chavez, A. Al-Hourani, J. Fu, and K. Sithamparanathan, "Low-rate tcp ddos attack model in the southbound channel of software defined networks," in *2020 14th International Conference on Signal Processing and Communication Systems (ICSPCS)*, 2020, pp. 1–10. [Online]. Available: DOI: 10.1109/ICSPCS50536.2020.9310040

[34] J. Luo, X. Yang, J. Wang, J. Xu, J. Sun, and K. Long, "On a mathematical model for low-rate shrew ddos," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 7, pp. 1069–1083, 2014. [Online]. Available: DOI: 10.1109/TIFS.2014.2321034

[35] L. Tan, K. Huang, G. Peng, and G. Chen, "Stability of tcp/aqm networks under ddos attacks with design," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 4, pp. 3042–3056, 2020. [Online]. Available: DOI: 10.1109/TNSE.2020.3012002

[36] Z. Liu, X. Yin, and H. J. Lee, "A new network flow grouping method for preventing periodic shrew ddos attacks in cloud computing," in *2016 18th International Conference on Advanced Communication Technology (ICACT)*, 2016, pp. 66–69. [Online]. Available: DOI: 10.1109/ICACT.2016.7423276

[37] H. Chen, Y. Chen, D. H. Summerville, and Z. Su, "An optimized design of reconfigurable psd accelerator for online shrew ddos attacks detection," in *2013 Proceedings IEEE INFOCOM*, 2013, pp. 1780–1787. [Online]. Available: DOI: 10.1109/INFCOM.2013.6566976

[38] N. Zhang, F. Jaafar, and Y. Malik, "Low-rate dos attack detection using psd based entropy and machine learning," in *2019 6th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)/ 2019 5th IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom)*, 2019, pp. 59–62. [Online]. Available: DOI: 10.1109/CSCloud/EdgeCom.2019.00020

[39] @efchatz, "Quic-attacks," https://github.com/efchatz/QUIC-attacks, 2023.

[40] B. Teyssier, Y. A. Joarder, and C. Fung, "An empirical approach to evaluate the resilience of quic protocol against handshake flood attacks," in *2023 19th International Conference on Network and Service Management (CNSM)*, 2023, pp. 1–9. [Online]. Available: DOI: 10.23919/CNSM59352.2023.10327907

[41] A. S. Balaji, V. Anil Kumar, P. P. Amritha, and M. Sethumadhavan, "Quicloris: A slow denial-of-service attack on the quic protocol," in *Advanced IoT Sensors, Networks and Systems*, A. K. Dubey, V. Sugumaran, and P. H. J. Chong, Eds. Singapore: Springer Nature Singapore, 2023, pp. 85–94. [Online]. Available: DOI: 10.1007/978-981-99-1312-1_7

[42] V. V. Tong, S. Souihi, H.-A. Tran, and A. Mellouk, State of the Art on Network Troubleshooting, 2023, pp. 1–23. [Online]. Available: DOI: 10.1002/9781394236664.ch1

[43] A. Praseed and P. S. Thilagam, "Ddos attacks at the application layer: Challenges and research perspectives for safeguarding web applications," *IEEE Communications Surveys Tutorials*, vol. 21, no. 1, pp. 661–685, 2019. [Online]. Available: DOI: 10.1109/COMST.2018.2870658

[44] J. Wang, X. Yang, and K. Long, "A new relative entropy based app-ddos detection method," in *The IEEE symposium on Computers and Communications*, 2010, pp. 966–968. [Online]. Available: DOI: 10.1109/ISCC.2010.5546587

[45] S. Yadav and S. Subramanian, "Detection of application layer ddos attack by feature learning using stacked autoencoder," in *2016 International Conference on Computational Techniques in Information and Communication Technologies (ICCTICT)*, 2016, pp. 361–366. [Online]. Available: DOI: 10.1109/ICCTICT.2016.7514608

[46] K. Hasegawa, D. Kondo, and H. Tode, "Fqdn-based whitelist filter on a dns cache server against the dns water torture attack," in *2021 IFIP/IEEE International Symposium on Integrated Network Management (IM)*, 2021, pp. 628–632. [Online]. Available: DOI: 10.1109/tnsm.2023.3277880

[47] R. Perrotta and F. Hao, "Botnet in the browser: Understanding threats caused by malicious browser extensions," *IEEE Security Privacy*, vol. 16, no. 4, pp. 66–81, 2018. [Online]. Available: DOI: 10.1109/MSP.2018.3111249

[48] M. Dimolianis, A. Pavlidis, D. Kalogeras, and V. Maglaris, "Mitigation of multi-vector network attacks via orchestration of distributed rule placement," in *2019 IFIP/IEEE Symposium on Integrated Network and Service Management (IM)*, 2019, pp. 162–170.

[49] B. Nagy, "Cldap dns multivector ddos attack," Oct 2021. https://zenodo.org/record/5572097.

[50] ——, "Udp flood attack sample with short payload," Oct 2021. https://zenodo.org/record/5578727.

[51] ——, "Udp flood attack sample," Oct 2021. https://zenodo.org/record/5578712.

[52] ——, "Icmp, udp, tcp syn multivector ddos attack," Oct 2021. https://zenodo.org/record/5578703.

[53] G. F. Scaranti, L. F. Carvalho, S. Barbon, and M. L. Proença, "Artificial immune systems and fuzzy logic to detect flooding attacks in software-defined networks," *IEEE Access*, vol. 8, no. 10.1109/access.2020.2997939, pp. 100 172–100 184, 2020. [Online]. Available: DOI: 10.1109/access.2020.2991273

[54] M. Poongodi, M. Hamdi, A. Sharma, M. Ma, and P. K. Singh, "Ddos detection mechanism using trust-based evaluation system in vanet," IEEE Access, vol. 7, pp. 183 532–183 544, 2019. [Online]. Available: DOI: 10.1109/access.2019.2960367

[55] S. M. Tabatabaie Nezhad, M. Nazari, and E. A. Gharavol, "A novel dos and ddos attacks detection algorithm using arima time series model and chaotic system in computer networks," *IEEE Communications Letters*, vol. 20, no. 4, pp. 700–703, 2016. [Online]. Available: DOI: 10.1109/lcomm.2016.2517622

[56] K. A. Simpson, S. Rogers, and D. P. Pezaros, "Per-host ddos mitigation by direct-control reinforcement learning," *IEEE Transactions on Network and Service Management*, vol. 17, no. 1, pp. 103–117, 2020. [Online]. Available: DOI: 10.1109/TNSM.2019.2960202

[57] H. Kim, S. Hong, J. Kim, and J. Ryou, "Intelligent application protection mechanism for transportation in v2c environment," *IEEE Access*, vol. 8, pp. 86 777–86 787, 2020. [Online]. Available: DOI: 10.1109/access.2020.2991273

[58] T. Yang, J. Jiang, P. Liu, Q. Huang, J. Gong, Y. Zhou, R. Miao, X. Li, and S. Uhlig, "Adaptive measurements using one elastic sketch," *IEEE/ACM Transactions on Networking*, vol. 27, no. 6, pp. 2236–2251, 2019. [Online]. Available: DOI: 10.1109/tnet.2019.2943939

[59] Y. Gao, H. Wu, B. Song, Y. Jin, X. Luo, and X. Zeng, "A distributed network intrusion detection system for distributed denial of service attacks in vehicular ad hoc network," *IEEE Access*, vol. 7, pp. 154 560–154 571, 2019. [Online]. Available: DOI: 10.1109/access.2019.2948382

[60] L. Xiao, X. Wan, X. Lu, Y. Zhang, and D. Wu, "Iot security techniques based on machine learning: How do iot devices use ai to enhance security?" *IEEE Signal Processing Magazine*, vol. 35, no. 5, pp. 41–49, 2018. [Online]. Available: DOI: 10.1109/msp.2018.2825478

[61] Corero, "Corero ddos threat intelligence report 2021," https://go.corero.com/ddos-threat-intelligence-report-2021-ty, 2021, accessed: 10-01-2023.

[62] Netscout, "Netscout threat intelligence report 2021 q2," https://www.netscout.com/sites/default/files/2022-03/ThreatReport_2H2021_WEB.pdf, 2021, accessed: 10-01-2023.

[63] T.-C. Huang, C.-Y. Huang, and Y.-C. Chen, "Real-time ddos detection and alleviation in software-defined in-vehicle networks," *IEEE Sensors Letters*, vol. 6, no. 9, pp. 1–4, 2022. [Online]. Available: DOI: 10.1109/LSENS.2022.3202301

[64] Z. Li, Y. Kong, C. Wang, and C. Jiang, "Ddos mitigation based on space-time flow regularities in iov: A feature adaption reinforcement learning approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 3, pp. 2262–2278, 2022. [Online]. Available: DOI: 10.1109/TITS.2021.3066404

[65] X. Chen, L. Xiao, W. Feng, N. Ge, and X. Wang, "Ddos defense for iot: A stackelberg game model-enabled collaborative framework," *IEEE Internet of Things Journal*, vol. 9, no. 12, pp. 9659–9674, 2022. [Online]. Available: DOI: 10.1109/JIOT.2021.3138094

[66] X. Chen, Y. Chen, W. Feng, L. Xiao, X. Li, J. Zhang, and N. Ge, "Real-time ddos defense in 5g-enabled iot: A multidomain collaboration perspective," *IEEE Internet of Things Journal*, vol. 10, no. 5, pp. 4490–4505, 2023. [Online]. Available: DOI: 10.1109/JIOT.2022.3218728

[67] Y. Liu, K.-F. Tsang, C. K. Wu, Y. Wei, H. Wang, and H. Zhu, "Ieee p2668-compliant multi-layer iot-ddos defense system using deep reinforcement learning," *IEEE Transactions on Consumer Electronics*, vol. 69, no. 1, pp. 49–64, 2023. [Online]. Available: DOI: 10.1109/TCE.2022.3213872

Péter Orosz is an associate professor and the head of Smart Communications Laboratory at the Department of Telecommunications and Media Informatics, BME Hungary. He received his Computer Science master's degree in software engineering (2003) and Ph.D. in infocommunication systems (2010) at the University of Debrecen, Hungary. His research interests cover communication networks, network and service management, QoS-QoE managed networks, online QoE prediction for media services, and acceleration of network functions.

Balázs Nagy is a PhD student at the Smart Communications Laboratory (Department of Telecommunications and Media Informatics), BME Hungary. He received his electrical engineering master's degree in embedded systems (2019). Currently, he is working as a project manager at AITIA International. His research interests cover communication networks, network security, service management, and acceleration of network functions.

Pál Varga is the Head of Department of Telecommunications and Media Informatics at the Budapest University of Technology and Economics. His main research interests include communication systems, Cyber-Physical Systems and Industrial Internet of Things, network traffic analysis, end-to-end QoS and SLA issues – for which he is keen to apply hard-ware acceleration and artificial intelligence, machine learning techniques as well. Besides being a member of HTE, he is a senior member of IEEE, where he is active both in the IEEE ComSoc (Communication Society) and IEEE IES (Industrial Electronics Society) communities. He is Editorial Board member in many journals, Associate Editor in IEEE Transactions on Network and Service Management, and the Editor-in-Chief of the Infocommunications Journal.

Guidelines for our Authors

Format of the manuscripts

Original manuscripts and final versions of papers should be submitted in IEEE format according to the formatting instructions available on

https://journals.ieeeauthorcenter.ieee.org/
Then click: "IEEE Author Tools for Journals"
- "Article Templates"
- "Templates for Transactions".

Length of the manuscripts

The length of papers in the aforementioned format should be 6-8 journal pages.
Wherever appropriate, include 1-2 figures or tables per journal page.

Paper structure

Papers should follow the standard structure, consisting of *Introduction* (the part of paper numbered by "1"), and *Conclusion* (the last numbered part) and several *Sections* in between.
The Introduction should introduce the topic, tell why the subject of the paper is important, summarize the state of the art with references to existing works and underline the main innovative results of the paper. The Introduction should conclude with outlining the structure of the paper.

Accompanying parts

Papers should be accompanied by an *Abstract* and a few *Index Terms (Keywords)*. For the final version of accepted papers, please send the short cvs and *photos* of the authors as well.

Authors

In the title of the paper, authors are listed in the order given in the submitted manuscript. Their full affiliations and e-mail addresses will be given in a footnote on the first page as shown in the template. No degrees or other titles of the authors are given. Memberships of IEEE, HTE and other professional societies will be indicated so please supply this information. When submitting the manuscript, one of the authors should be indicated as corresponding author providing his/her postal address, fax number and telephone number for eventual correspondence and communication with the Editorial Board.

References

References should be listed at the end of the paper in the IEEE format, see below:

a) Last name of author or authors and first name or initials, or name of organization
b) Title of article in quotation marks
c) Title of periodical in full and set in italics
d) Volume, number, and, if available, part
e) First and last pages of article
f) Date of issue
g) Document Object Identifier (DOI)

[11] Boggs, S.A. and Fujimoto, N., "Techniques and instrumentation for measurement of transients in gas-insulated switchgear," IEEE Transactions on Electrical Installation, vol. ET-19, no. 2, pp.87–92, April 1984. DOI: 10.1109/TEI.1984.298778

Format of a book reference:

[26] Peck, R.B., Hanson, W.E., and Thornburn, T.H., Foundation Engineering, 2nd ed. New York: McGraw-Hill, 1972, pp.230–292.

All references should be referred by the corresponding numbers in the text.

Figures

Figures should be black-and-white, clear, and drawn by the authors. Do not use figures or pictures downloaded from the Internet. Figures and pictures should be submitted also as separate files. Captions are obligatory. Within the text, references should be made by figure numbers, e.g. "see Fig. 2."
When using figures from other printed materials, exact references and note on copyright should be included. Obtaining the copyright is the responsibility of authors.

Contact address

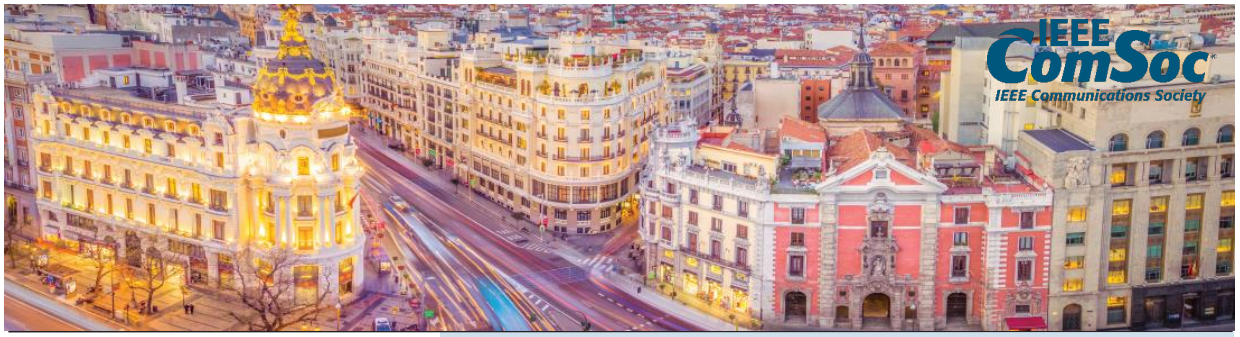
Authors are requested to submit their papers electronically via the following portal address:

https://www.ojs.hte.hu/infocommunications_journal/about/submissions

If you have any question about the journal or the submission process, please do not hesitate to contact us via e-mail:

Editor-in-Chief: Pál Varga – pvarga@tmit.bme.hu

Associate Editor-in-Chief:
József Bíró – biro@tmit.bme.hu
László Bacsárdi – bacsardi@hit.bme.hu



IEEE MeditCom

Organizing Committee

General Chair
Ana Garcia Armada, Universidad Carlos III de Madrid, Spain

TPC Co-Chairs
Besma Smida, University of Illinois at Chicago, USA
Daniel Benevides Da Costa, King Fahd University of Petroleum & Minerals, Saudi Arabia

Workshop Co-Chairs
Giovanni Geraci, Telefonica Research and Universitat Pompeu Fabra, Spain
Stefania Bartoletti, University of Rome Tor Vergata, Italy

Tutorial Co-Chairs
Ramón Agüero, Universidad de Cantabria, Spain
Pascal Lorenz, University of Haute Alsace, France

Panel Co-Chairs
Mona Ghassemian, Huawei, UK
Tuncer Baykas, Ofinno, USA
Richie (Ruiqi) Liu, ZTE, China

Keynote Chair
Angel Lozano, Universitat Pompeu Fabra, Spain

Finance Chair
Víctor P. Gil Jiménez, Universidad Carlos III de Madrid, Spain

Local Arrangements Chair
M. Julia Fernández-Getino García, Universidad Carlos III de Madrid, Spain

Publication Chair
Máximo Morales Cespedes, Universidad Carlos III de Madrid, Spain

Web & Social Media Chair
Kun Chen Hu, Aalborg University, Denmark

Travel Grants Co-Chairs
Virginia Pilloni, University of Cagliari, Italy
Karim Banawan, Alexandria University, Egypt

Publicity Co-Chairs
Daniel Corujo, Instituto de Telecomunicações and Universidade de Aveiro, Spain
Amr A. El-Sherif, Egypt University of Informatics, Egypt

www.ieee-meditcom.org

2024 IEEE International Mediterranean Conference on Communications and Networking

8-11 July 2024 // Madrid, Spain

CALL FOR PAPERS

IEEE MeditCom is the conference of the IEEE Communications Society serving the Mediterranean area and surrounding countries. It gathers visionary researchers in academia and industry from all over the world to the shores of the Mediterranean Sea. IEEE MeditCom features a comprehensive and timely technical program, that will address many of the outstanding challenges that exist in the areas of communications and networking. IEEE MeditCom 2024 solicits researchers in industry and academia to submit papers on a wide range of research subjects, encompassing theoretical and applied research. Original technical papers are sought, but are not limited, to the following areas:

- 5G/6G Systems and Networks
- Antennas, Propagation, and Channel Modeling
- Big Data and Machine Learning for Communications
- Cloud Communications and Data-Center Networks
- Coding/Decoding Theory and Techniques for Communications
- Cognitive Radio and Dynamic Spectrum Access
- Communication and Information Theory
- Edge Computing, Edge Intelligence, and Fog Networks
- Energy Efficient Communications and Computing
- Image, Speech, and Signal Processing for Communications
- Integrated Sensing and Communications
- Internet of Things, Smart Grids and Vehicular Networks
- Massive MIMO and Cell-Free Massive MIMO
- Millimeter-Wave, Sub-Terahertz, and Terahertz Communications
- Molecular and Nanoscale Communications
- Network Applications, Services, and Management
- Network Architecture, SDN, NFV
- Next-Generation Multiple Access Schemes
- Next-Generation Physical, Link, and Network Layers Techniques
- Optical Communications and Networks
- Performance Evaluation, Simulation, Testbeds and Prototypes
- QoE/QoS Support and Cross-Layer Design
- Quantum Communications and Computing
- Reconfigurable Intelligent Surfaces and Holographic Surfaces
- Satellite and Space Communications
- Security, Privacy, Trust and Blockchain
- Semantic and Goal-Oriented Communications
- Smart Grids and Energy Networks
- Underground and Underwater Communications

IMPORTANT DATES

Submissions Deadline:	3 March 2024
Acceptance Notification:	30 April 2024
Camera-Ready Submission:	20 May 2024

SCIENTIFIC ASSOCIATION FOR INFOCOMMUNICATIONS

Who we are

Founded in 1949, the Scientific Association for Infocommunications (formerly known as Scientific Society for Telecommunications) is a voluntary and autonomous professional society of engineers and economists, researchers and businessmen, managers and educational, regulatory and other professionals working in the fields of telecommunications, broadcasting, electronics, information and media technologies in Hungary.

Besides its 1000 individual members, the Scientific Association for Infocommunications (in Hungarian: Híradástechnikai és Informatikai Tudományos Egyesület, HTE) has more than 60 corporate members as well. Among them there are large companies and small-and-medium enterprises with industrial, trade, service-providing, research and development activities, as well as educational institutions and research centers.

HTE is a Sister Society of the Institute of Electrical and Electronics Engineers, Inc. (IEEE) and the IEEE Communications Society.

What we do

HTE has a broad range of activities that aim to promote the convergence of information and communication technologies and the deployment of synergic applications and services, to broaden the knowledge and skills of our members, to facilitate the exchange of ideas and experiences, as well as to integrate and harmonize the professional opinions and standpoints derived from various group interests and market dynamics.

To achieve these goals, we…

- contribute to the analysis of technical, economic, and social questions related to our field of competence, and forward the synthesized opinion of our experts to scientific, legislative, industrial and educational organizations and institutions;
- follow the national and international trends and results related to our field of competence, foster the professional and business relations between foreign and Hungarian companies and institutes;
- organize an extensive range of lectures, seminars, debates, conferences, exhibitions, company presentations, and club events in order to transfer and deploy scientific, technical and economic knowledge and skills;
- promote professional secondary and higher education and take active part in the development of professional education, teaching and training;
- establish and maintain relations with other domestic and foreign fellow associations, IEEE sister societies;
- award prizes for outstanding scientific, educational, managerial, commercial and/or societal activities and achievements in the fields of infocommunication.

Contact information

President: **FERENC VÁGUJHELYI** • *elnok@hte.hu*

Secretary-General: **GÁBOR KOLLÁTH** • *kollath.gabor@hte.hu*

Operations Director: **PÉTER NAGY** • *nagy.peter@hte.hu*

Address: H-1051 Budapest, Bajcsy-Zsilinszky str. 12, HUNGARY, Room: 502

Phone: +36 1 353 1027

E-mail: *info@hte.hu*, Web: *www.hte.hu*