

2024 | Vol 1 | Issue 1

WWRT

Wireless World Research and Trends Magazine



Official Magazine of the
WIRELESS WORLD
RESEARCH FORUM


River Publishers

Editor-in-Chief

Sudhir Dixit, Basic Internet Foundation, USA

Editorial Board

Angeliki Alexiou, University of Piraeus, Greece
 Ebtesam Almazrouei, AIE3, UAE
 Hendrik Berndt, WWRF, Germany
 Bharat B. Bhatia, ITU-APT Foundation of India (IAFI), India
 Mérouane Debbah, Khalifa University, UAE
 Nigel Jefferies, Montreal Consulting Ltd and Huawei Technologies
 Vinod Kumar, (Retd) Alcatel-Lucent Bell Labs, France
 Seshadri Mohan, University of Arkansas, USA
 Christos Politis, Kingston University London, UK
 Steffen Ring, RING Advocacy LLC, Denmark
 Knud Erik Skouby, Aalborg University, Denmark
 Vino Vinodrai, Vinodrai & Associates Inc, Canada

How to Reach Us

We welcome letters to the editor; but, we reserve the right to edit for space, style, and clarity. Include address and daytime phone number with your correspondence.
 E-mail: sudhir.dixit@gmail.com

If you would like to submit an article, please visit our submission portal for details:
<https://wireless-magazine.com/index.php/WWRT/about/submissions>

Official Magazine of the
**WIRELESS WORLD
 RESEARCH FORUM**



River Publishers

Alsbjergvej 10, 9260 Gistrup, Denmark
www.riverpublishers.com

2024 | Vol 1 | Issue 1

WWRT

Wireless World Research and Trends Magazine

EDITORIALS

Foreword from the Chairman of the Wireless World Research Forum <i>Nigel Jefferies</i>	iii
Welcome to the Inaugural Issue of the WWRT <i>Sudhir Dixit</i>	iv

EXPERTS COLUMN

Making a Better World with Wireless Technology <i>Martin Cooper</i>	vii
Standards Update <i>Bharat Bhatia, Sudhir Dixit, Hyeon Woo Lee, Kobei Satoh, Bin Tan, Vino Vinodrai</i>	xi

PEER REVIEWED ARTICLES

Advancements in Deep-Learning-Based Object Detection in Challenging Environments <i>Abbisbek Mukhopadhyay and Pradipta Biswas</i>	1
Emotion Detection from Speech: A Comprehensive Approach Using Speech-to-Text Transcription and Ensemble Learning <i>Fatima M Inamdar, Sateesh Ambesange, Parikshit Mahalle, Nilesb P. Sable, Ritesh Bachhav, Chaitanya Ganjiwale, Shantanu Badmanji and Sarthak Agase</i>	7
Beam-Tracking Challenges in THz Communications <i>Giorgos Stratidakis, Sotiris Droulias and Angeliki Alexiou</i>	13
Coherence Bandwidth of Rice Channels for Millimeter Wave and Sub-Terahertz Applications <i>Werner Mohr</i>	21
Adversarial Machine-Learning-Enabled Anonymization of OpenWiFi Data <i>Sambita Kuili, Kareem Dabbour, Irtiza Hasan, Andrea Herscovich, Burak Kantarci, Marcel Chenier and Melike Erol-Kantarci</i>	33
On the Impact of CDL and TDL Augmentation for RF Fingerprinting Under Impaired Channels <i>Omer Melib Gul, Michel Kulhandjian, Burak Kantarci, Claude D'Amours, Azzedine Touazi and Cliff Ellement</i>	43

NEWS AND INFORMATION

Call for Papers for the Next WWRF Meeting #51	53
Brochure for the WWRF Huddle	55
Upcoming Conferences	57

Foreword from the Chairman of the Wireless World Research Forum

Dr Nigel Jefferies



So here it is at last, the first edition of WWRF's new magazine: Wireless World Research and Trends. It's intended to include news, opinion, comment and reviews, as well as the best of the many papers contributed to WWRF meetings, and the latest results from researchers worldwide. As the chair of the Wireless World Research Forum, I am pleased that the extensive discussions on whether and in what form the world needs a new publication on mobile and wireless communications has been definitively answered by what you have in your hands or on your screen.

The launch of this magazine is very timely. At present, we have seen some coalescing of the vision of the next generation, whether as 6G or IMT-2030, around overarching drivers such as sustainability, ubiquity and intelligence. The challenge now is to identify and develop technologies that can deliver this vision. This will require a strong collaboration between industry and academia, working together to meet these goals at a time when the industry is in transition, and the economic situation is challenging.

And this is where WWRF can play a role, as the world's leading independent industry/academia collaboration, where the

research community can come together to tackle the challenges and exciting opportunities. Our members include a wide range of leading research universities, institutes, industry large and small, regulators and independent experts. If you want to be part of our work, you are very welcome to join us. Through our meetings, both regular Forums and our world-renowned annual Huddle, our publications, including book series and our Outlook series of white papers, as well as this magazine of course, and the fantastic networking opportunities of being part of WWRF, we can help you play a part in developing the next generation.

I trust that you find the contents stimulating and informative, and encourage you strongly to consider contributing your own research results, views and news to future editions. The editor-in-chief and his team have worked extremely hard to get this first edition together, and we are pleased with and proud of the result. We await your feedback with interest!

A handwritten signature in blue ink that reads "Nigel Jefferies". The signature is written in a cursive, flowing style.

Welcome to the Inaugural Issue of the WWRT

Dr Sudhir Dixit
Editor-in-Chief



It gives me immense pleasure to launch this inaugural issue of the *Wireless World Research and Trends (WWRT)* magazine. At WWRF we endeavour to spot promising technologies and trends and accelerate them from the research phase to standardization to development and ultimately to deployment. Launching of the *Wireless World Research and Trends (WWRT)* Magazine is intended to disseminate that knowledge through published media. The editors of the magazine hope that the readers will find the content interesting and useful.

We are living in exciting times when ubiquitous connectivity at high speeds is transforming everything in our daily lives, be it personal or business. The technologies driving connectivity and applications thereupon are intertwined, and a new era of cross-domain research and innovation is thriving and will only accelerate in the future. Communications is not only about radio and transport, but encompasses materials, electronics, radio waves, computer science and even biological sciences. After a long and spirited internal deliberation, *Wireless World Research Forum (WWRF)* decided to launch a new type of a magazine in the open access space that is inter-disciplinary and has a broad scope of coverage of recent trends, avoids deep mathematical content, and is practical from feasibility perspective. The magazine also welcomes articles designed to drive the standards and trends pertaining to techno-economic and social aspects. While the magazine will evolve over time, we are starting off with articles from experts, standards updates from around the world, and announcements about the upcoming major academic and industry events. The magazine will be published twice a year: January/February and July/August of each year with no limit on the number of articles that would be published in each issue. The editorial board is constituted with some of the best-known experts in academia and industry. Each submitted article will undergo a rigorous peer review process by the reviewers. Proposals for special issues are most welcome from anyone who will assume the role of guest editors.

In this first issue, we are publishing six articles, one expert column and a standards update column. The expert's column is written by WWRF Fellow Martin Cooper, who in 1972–73 built the first mobile cell phone and made the first cell phone call. He is widely regarded to be the father of the mobile cell phone. In a story-telling genre, Cooper describes the history of mobile cell

phone and provides a glimpse of all the wonderful innovations that are yet to come. The standards update column summarizes some of the key developments from the different regions of the world by those based in those regions. The six carefully chosen articles cover a number of today's hot topics.

Since unstructured data is becoming quite dominant with the increased use of speech and images/video in applications, we include two articles in this domain: one on object detection utilizing deep learning techniques in difficult image-based environments, such as in autonomous vehicle use cases, and the second one on emotion detection from speech and text. In the first article, Mukhopadhyay and Biswas consider difficult environments where they focus on identifying objects in outdoor scenes with varying lighting, weather conditions, and background clutter. Applications such as autonomous driving, surveillance, and industrial production inspection fall under these scenarios. The paper proposes novel solutions to improve accuracy of identification by calling for the need of relevant datasets to validate models under unconstrained scenarios.

While there is a large body of research in AI and ML (machine learning)-based natural language processing (NLP) in speech-to-text and text-to-speech processing, identifying the emotional nuances has been challenging. Inamdar, et al., present an overview of the state-of-the-art of text-to-emotion analysis, and present a framework for designing emotionally realistic chatbots, aiming to enhance their believability. The proposed framework leverages AIML (Artificial Intelligence Markup Language) and information states while infusing chatbots with emotional expression and responsiveness.

High bandwidth demand calls for the use of spectrum in the mmWave and THz range, which, in instances of non-line-of-sight (nLoS) links, causes increased attenuation levels reducing the quality of wireless link. To counter this increased pathloss, the use of directional antennas is widely recommended which have to be properly aligned. In the third article, Stratidakis, et al., describe the challenges of beam tracking in communications in the THz range and propose a number of solutions. In the fourth article, Coherence Bandwidth of Rice Channels for Millimeter Wave and Sub-Terahertz Applications, Mohr investigates in detail the coherence bandwidth and time for Rice channels based on


an approximate approach for the fading statistics. This analysis is particularly useful for communications in the Terahertz range, where operations in such high frequency range pose the challenge of multipath propagation under shadowing conditions which affects radio propagation, where multipath propagation results in frequency-selective fading. In these frequency ranges shadowing can be overcome by additional means in the network deployment such as reflectors, RIS arrays or repeaters.

Data protection and privacy are becoming important requirements in the present day data driven society. In the next article, Kuili, et al., propose utilizing artificial Intelligence to anonymize data in open Wi-Fi networks to prevent data leakage and provide security and privacy before the network operators or data owners forward the data.

There is a growing interest in Cyber-Physical Systems (CPS) in numerous industry verticals due to automation and increased efficiencies. Unfortunately, this introduces the potential for cyber

attacks and compromised security, particularly if wireless technologies are being utilized. Significant research is ongoing on Radio frequency (RF) fingerprinting as a promising approach to thwart the unwanted attacks. Gul, et al., investigate the impact of decoupling tapped delay line and clustered delay line (TDL+CDL) augmentation-driven deep learning (DL) on transmitter-specific fingerprints to isolate malicious users from legitimate ones, and compare the results with only one type of augmentation or none at all with the former (joint technique) delivering much better performance.

Finally, yet importantly, we would like to receive proposals for special issues, original articles, and love to hear your feedback about our magazine: what you like, what you don't, and suggestions on how to improve the magazine. Happy reading!



Making a Better World with Wireless Technology

Martin Cooper



The Wireless World Research Forum could not have selected a better time to launch a journal that seeks to influence the future of wireless science and technology. You may find this notion counterintuitive. After all, research into electromagnetic waves has been happening for nearly 150 years. The explosive impact of life-changing wireless applications, starting with Marconi's commercialization of radio communications around 1900, has been repeated with broadcast radio, television, cellular telephony, and more recently, the merger of personal communications with Internet access and artificial intelligence.

Wireless might be viewed as mature technology, aged, grey, and boring. I see it as exciting and magical. Can it be a coincidence that this combination of three personal technologies, universal wireless connectivity, unlimited processing capability, and accessibility to all the recorded knowledge of mankind be a mere coincidence? Or are we about to witness the greatest impact of technology upon humanity in the history of mankind?

I think it's the latter.

Before I try to persuade you that wireless technology is on the brink of an explosion of creativity and innovation, allow me to briefly review its history.

In the late 1890's Heinrich Hertz, a brilliant German physicist demonstrated to a group of his colleagues and students the ability of electromagnetic energy to act at a distance, a phenomenon James Clerk Maxwell had predicted theoretically twenty years earlier.

On Hertz's bench lay a wooden chest topped by two small coils of wire spaced about ten centimeters apart and topped by silvered spheres. A conducting loop with a gap of about a centimeter was several meters away. He closed the shades, turned the oil lamps off, and connected a voltaic pile to contacts on the box. A spark formed between the spheres. A smaller spark simultaneously appeared at the loop across the room.

Hertz said, "In this demonstration we have transmitted electromagnetic energy at the speed of light without the use of wires." He continues, "Of course, this device has no practical application, but we have now validated Maxwells theory of electromagnetism."

How wrong could he be! A few years later, Guglielmo Marconi created a practical application by scaling up Hertz's apparatus to a huge generator putting out kilowatts of power to wires strung across telephone poles hundreds of meters apart. The equivalent of a lightning bolt leapt between a group of telephone poles in

Newfoundland. The energy traveled across the Atlantic Ocean and was received in the British Isles. Barely, but successfully. Within a few years, Marconi commercialized his demonstration and placed Marconi transmitters and receivers on ships crossing the Atlantic Ocean. A radio operator on one of these ships was responsible for saving over 700 lives when the Titanic sank, pretty much disqualifying Hertz as a forecaster of practical application!

Within 30 years radio and television broadcasting became major industries.

In 1947 Rae Young and Doug Ring at Bell Laboratories devised a method to increase the capacity of mobile radio systems by deploying transmitting stations in repeating patterns and providing for the hand-off of their radio channels to different frequencies as mobile receivers moved to different areas within these patterns. This reuse of frequencies would, they said, increase the number of car telephones served in a typical city, then a few hundred, by a hundred times or more. Doug Ring's technical memo declaring that "the concept is obvious" and that there has no immediate application for this approach, was relegated to the Bell Labs archives.

Twenty-one years later the Bell System introduced the concept of "cellular" telephony revisiting the Young/Ring concept but picturing the "cells" as hexagons in place of the circles envisioned by Ring. Bell engaged the consulting firm McKinsey to predict the market for car telephones. McKinsey concluded that the maximum market, worldwide, would never exceed 900,000 car phones (which, remarkably turned out to be true: the maximum number of car phones in the world never did exceed 900,000.) Bell requested the FCC in the U.S. to allocate it the exclusive block of 900 MHz spectrum, because the ultimate market as predicted by McKinsey was clearly not large enough to support competitive carriers. Also, because the limited market forecast would not, alone, justify Bell's investment, Bell asked the FCC for exclusive authority to serve the air-to-ground market and to compete in the land-mobile market, then robustly competitive.

My colleagues and I at Motorola disputed The Bell prediction, stating that technology was now available for people to talk wirelessly and directly to other people, that wireless personal communications was now practical, and that the market for *hand-held* phones would be huge. In testimony before the FCC, I predicted, "Someday, a person was born, he or she would be assigned a phone number. If she didn't answer, she had died."

There are now *more* cell phones in the world than people. More than 6 billion people use cell phones every day. The wired phone is fading into obsolescence.

I hope I've made my point that predictions of the future can be amazingly unreliable. Here's another example: In 1980 I witnessed a presentation by Bill Howard, chief technologist of Motorola's semiconductor division. "We are reaching the end of the road in the density of semiconductor integrated circuits," he said, "we have achieved a pitch (the closest together that two conducting lines can be fabricated on a chip) of one micron. We can now almost count the atoms of gold in the width of a line on our chips."

Semiconductor fabs are now producing Integrated circuits with a pitch of a few nanometers, increasing the density of devices on a chip by almost a million times more than in 1980. There are billions of transistors on a single chip in every one of today's

smartphones. Bill was a brilliant engineer manager but another poor forecaster.

In full knowledge of the foolhardiness of technology forecasting, here is my view of the future impact of wireless technology.

Despite this enormous increase in processing power, with a concomitant increase in speed; despite the integration of internet access, availability of millions of applications, and of digital cameras, we are still in early days in our understanding of the power of connectivity and digital access to knowledge. Technologists have become captivated by the idea of the "internet of things" when they haven't come close to exploiting the "internet of *people*".

We are on the brink of a world truly empowered by wireless technology applied to the solving the highest priorities of the problems of humanity. The following problems, I believe, deserve higher priority than the internet-of-things. This prioritization does not presume that the internet-of-things should be neglected. The objective of us technologists should be to achieve a balance between serving the direct interests of people and the indirect benefit to humanity by deploying *things* efficiently.

Elimination of the Digital Divide

In today's digital age, access to the internet is an indispensable, an essential tool for acquiring knowledge and skills. And yet millions of students struggle with digital learning because they don't have full-time access to the internet on devices that allow for immersion in the wealth of knowledge the internet offers. Investing in connectivity and technology, digital literacy, and digital learning solutions help to connect today's learners with success – in and out of the classroom. Having full-time, mobile access to the internet – at school, on the way to and from school, and at home – has become an essential ingredient for a comprehensive education.

The greatest opportunity and the most severe challenge to society, and to the wireless community, is addressing the revolution in education that is happening in our most advanced schools – a revolution still in its infancy. Those K-12 students (under 18 year-olds) who have full-time access to the internet and availability of an enlightened educational process are far more capable of dealing with technological social and political complexities of the modern world. The traditional educational process involving lecturing for extended times on narrow subjects is woefully inadequate. Students who have full-time access to the internet, to artificial intelligence tools, and to informed guidance by their teachers will be better thinkers and superior problem solvers than those who don't. The role of teachers will be more focused on the use of the new tools, on discriminating between information and disinformation, on ethics, collaboration, and principles.

Students with full-time access to the Internet have more information at their fingertips than their teachers can possibly provide in lectures. And "full-time" means *always*. Wi-Fi and cable can be valuable, yet no adult would be satisfied with those services alone. Why should we expect that our students should be. Less than 50 percent of K-12 students have such access today, partially because of incomplete coverage but also because of the lack of affordability. Both deficiencies are solvable by technology optimized for education. Systems embracing solutions to the educational process, affordability, and coverage are being explored and implemented in many communities. The use of smart antennas (MIMO) and mesh

networks is already being explored for addressing the economic and coverage needs of education.

Our educators need to teach their constituency – their students – that the internet is a gold mine, a wonderful treasure. They need to make the educational experience more engaging, more interesting, and more enjoyable than the games their students play. We spend billions on these games. Isn't educating our future citizens as important as entertaining them? Why not do both at the same time?

And finally, we need to recognize that if we don't take solving the educational digital divide seriously, we will find ourselves facing the specter of a new kind of division. Of a generation of smarter, more engaged citizens who have been educated the modern way in contrast with those who haven't had that privilege. That digital divide is unacceptable and unnecessary in modern society.

Elimination of Disease

Our health care system is broken. We spend extravagantly on curing disease and woefully less on *preventing* disease. The intelligent smartphone combined with newly developed sensors promises to become the basis of a revolution in health care by eliminating diseases before they happen.

Smart phones, smart watches and electronic finger rings already measure many of the vital signs that characterize a person's health. Metrics like heart rate, blood pressure, body temperature, respiration rate, and even quality of sleep, are in widespread use. Researchers are getting close to being able to measure glucose levels in human beings without the traditional blood prick. A sensor has been demonstrated that will anticipate heart failure and provide sufficient warning to prevent it. Inexpensive add-ons to a cell phone use its camera and processor to remotely examine a patient's eyes, and even make a sonogram of an unborn baby.

The concept of the annual physical examination has had disappointing results. Diseases happen unscheduled, and people don't maintain exam schedules. But imagine a physical exam happening every minute and processed by supercomputers continuously. Imagine sensors distributed on or imbedded in our bodies tailored to our vulnerabilities by DNA measurement. Within a few generations, it's possible that sensors will be capable of detecting the loss of control of one's immune system that we call cancer, providing us with the ability to destroy the cancer cells before they threaten life.

Known technology can, today, anticipate and warn us about the onset of some diseases. It may be a dream, but it is a realizable dream that, in the future, *all disease will be anticipated and eliminated*.

The Elimination of Poverty

The cell phone has already improved productivity in many ways. That's according to Jeffrey Sachs, an economist and author of the book *The End of Poverty*. The cell phone has been one of the biggest contributors to falling poverty across the world. In 1981, 44 percent of the world's population lived in extreme poverty; that was down from 72 percent in 1950. Today, just 10 percent are classified as living in extreme poverty. The Economist estimates that a billion

people in Africa have moved out of severe poverty in the past 25 years.

Despite the distractions introduced into our lives by our cell phones, we can't live without them. *Cell phones make our lives more efficient*, much more so the lives of the poorest citizens of the world. The many ways that this efficiency manifests itself cannot be abbreviated here, but involve innovations like mobile money and bringing market information to the poorest entrepreneurs.

Where will the Radio Spectrum Come From?

The predictions of a more educated, healthier, and more productive world that I made above are subject to the same naivety expressed by Heinrich Hertz, Bell System engineer Doug Ring, and Motorola executive Bill Howard. But these predictions are merely extensions of existing technological advances. These advances are already happening. That includes a prediction that there will be an enormous increase in usage of the radio frequency spectrum necessary to implement the intranet-of-people, the internet-of-things, and numerous other spectrum consuming applications. If you believe the doomsayers, we are about to run out of usable radio frequency spectrum. They say that spectrum is like beach-front property: once it's occupied by someone it can't be used by someone else.

Nothing could be further from the truth.

The capacity of the existing radio spectrum has increased by billions of times since Marconi put it to commercial use. Yet, we are still extraordinarily inefficient in our use of the spectrum even after all that progress. Although the methods for increasing spectrum capacity are not as obvious as the examples of education, healthcare, and poverty, there is enough happening to support an optimistic view. Almost all the energy transmitted by modern communications devices is wasted. We are just starting to learn how to efficiently share the spectrum rather than providing exclusive use to individual entities or services. Our methods of allocation encourage this waste.

These challenges, and numerous others, are what make the future of wireless technology so exciting.

The combination of the ability to send information at the speed of light for long distances, the network that allows us to connect any person or device to any other person or device, the accessibility of all the information in the world and the ability to process that information, is extraordinarily powerful. Artificial intelligence magnifies that power.

The concept of disease will become obsolete. No one will go hungry. Future generations will be a lot smarter than we are, perhaps smart enough to eliminate the blight of war. It's going to take more than cell phones with artificial intelligence to solve all the problems in today's world but there is the potential of a more educated populous in a world that has no poverty, no disease and no wars.

I know you think I'm being overly optimistic; I'm not ashamed of that. I won't be around to see much of that future, but my readers the wireless technologists who are reading this today—will help create that wonderful future. And I hope you'll remember that Marty told you it was all possible.

Standards Update

Each issue of the magazine will include a section on updates from the major standards organizations around the world. This issue includes brief introductions and key updates from some of them. We welcome readers to contribute to this section if they have any news items or major updates/releases from their regions as far as the standards are concerned.

Third Generation Partnership Project (3GPP)

3GPP™ is a partnership project bringing together national Standards Development Organizations (SDOs) from around the globe. It partners with seven SDOs: ETSI (European Telecommunications Standards Institute and rest of world), ATIS (USA), ARIB and TTC (Japan), TTA (South Korea), CCSA (China) and TSDSI (India). Thus, 3GPP is an umbrella organization for a number of SDOs which develops protocols and specifications for mobile communications. It organizes its work into three different broad areas: Radio Access Networks, Services and Systems Aspects, and Core Network and Terminals. Recently, its work has focused on 5G and 5G-Advanced. 3GPP has also planned work on 6G, which is expected to be commercially launched around 2030. ETSI Standards are typically adopted by the European Commission as the technical base for directives or regulations. ETSI's mission is to produce globally applicable standards for ICT-enabled systems, applications and services for deployment across all sectors of industry and society, and it is by far the most dominant SDO in Europe.

Japan

Beyond 5G Promotion Consortium (B5GPC), Japan released “Beyond 5G White Paper Version 2.0” on 13th (https://b5g.jp/doc/whitepaper_en_2-0.pdf). The White Paper Subcommittee of B5GPC has considered Beyond 5G, which is expected to be commercialized around 2030, assuming future daily life of users and a wide range of usage scenarios that span many industries. Through this study with various perspectives, B5GPC examined the requirements that a communication system as a social infrastructure should have. B5GPC also considered Beyond 5G key concepts, network requirements, capabilities, architectures, and key technologies based on usage scenarios and future technology. Based on these studies, this white paper summarizes issues and expectations from each industry for Beyond 5G, capabilities required for Beyond 5G, and technological trends, etc.

Japan is also active in contributing to the following important documents in ITU-R Working Party 5D: Recommendation ITU-R M.2160 on “Framework and overall objectives of the future development of IMT for 2030 and beyond,” and a draft new Report ITU-R M. [IMT.ABOVE 100 GHz].

Korea

In Korea, as of 1st half 2023, 5G network is already deployed nationwide by 3 mobile operators using 3.5 GHz band. 5G penetration is higher than 40% and average monthly user data consumption by 5G is around 26 GB which is more than 3 times of 4G LTE

data consumption. 5G geographical coverage is more than 75% of territory and population coverage is more than 95%. In addition to public 5G services focusing on eMMB, private 5G networks, so called eUM-5G, targeting URLLC and mMTC services are under deployment by numerous players using 4.7 GHz and 28 GHz.

Regarding 6G preparation, Korean government has already funded 150 million USD for 6G core technologies R&D for 5 years from 2021 and recently decided to invest an additional 300 million USD for 5 years from 2024 for 6G system technology R&D. If both programs are successful, Korean consortium may demonstrate pre 6G system technologies in 2026 before official draft standard is available by 3GPP and ITU-R around 2028–2029.

India

In India, a number of agencies and government organisations are involved in telecom standards development activities. Among these, the main are Telecommunications Standards Development Society, India (TSDSI), the Bureau of Indian Standards (BIS), Wireless Planning and Coordination wing (WPC) and the Telecommunication Engineering Centre (TEC). These are briefly described below.

Telecommunications Standards Development Society, India (TSDSI)

TSDSI plays a pivotal role in the development of telecommunications standards in India. It focuses on tailoring global standards to suit the specific requirements and challenges of the Indian market, encompassing a wide range of technologies and services, including 5G, IoT. TSDSI actively collaborates with international standardisation bodies while ensuring that standards are accessible and affordable for all stakeholders, promoting innovation and interoperability in India's rapidly evolving telecommunications ecosystem. TSDSI is also the organisational partner of 3GPP, which is the global SDO for mobile telecommunications. TSDSI was key in the development of the Indian 5G standard called 5Gi which was later merged with 3GPP 5G.

Wireless Planning and Coordination wing (WPC)

WPC is the government organisation responsible for spectrum management and licensing of wireless frequencies in India. It also develops and prescribes standards for wireless equipment in India, in particular standards for unlicensed wireless equipment. WPC is responsible for planning, engineering, regulating, managing and monitoring of Radio Frequency (RF) spectrum and satellite orbits, including geo-stationary (GSO) and non-GSO orbit.

Telecommunication Engineering Centre (TEC)

TEC is the technical arm of the Department of Telecommunications (DoT), of the Government of India and is responsible for formulating and updating technical standards and specifications for telecommunications equipment and services in India. It is integral to ensuring the quality and reliability of telecommunication networks and infrastructure across the country. TEC's standards activities encompass areas such as network operation, maintenance, and security. TEC is also responsible for certification of

telecom products to comply with essential technical requirements, fostering a secure and robust telecommunications environment in India. TEC also runs the Mandatory Testing and certification of Telecom equipment (MTCTE) programme of the DoT.

The Bureau of Indian Standards (BIS)

BIS plays a significant role in the development and enforcement of standards for IT (Information Technology) in India. Its role includes Standard Development, Certification and Marking and Consumer Protection. BIS formulates and revises standards and specifications for a wide range of IT and consumer telecom equipment to ensure their quality, safety, and performance. These standards cover aspects such as electromagnetic compatibility, electrical safety, product quality, and interoperability. BIS mandates that certain categories of IT and consumer telecom equipment conform to Indian standards and undergo certification. Products that meet these standards receive the BIS certification mark, signifying compliance with safety and quality requirements. BIS often aligns its standards with international standards to promote interoperability and ease of trade for IT and consumer telecom equipment in the global market.

China

CCSA (China Communication Standardization Association) is a national and unified communications standardization organization in China. It was established on December 18, 2002 with the approval of the former Ministry of Information Industry, the Standardization Administration of the People's Republic of China, and the Ministry of Civil Affairs, in order to be in line with international standards and China's national conditions.

It is a non-profit organization composed of telecommunications operators, equipment manufacturers, research institutions, universities, and other relevant organizations.

The main functions of CCSA include:

1. Formulating and revising technical standards for telecommunications
2. Promoting the implementation and application of telecommunications standards
3. Providing technical support and consultation services for telecommunications development
4. Participating in international standardization activities and promoting the internationalization of Chinese telecommunications standards
5. Establishing and maintaining a platform for communication and cooperation among members.

CCSA has played a crucial role in promoting the development of China's telecommunications industry by formulating and revising a large number of important technical standards, and has also actively participated in international standardization activities.

United States

IETF

There are several activities in IETF related to 5G and B5G networks. Some important ones are: (1) Methods and settings for using IPv6 to communicate among IPv6 nodes within the communication range of one another over 5G V2X links (draft-jeong-6man-ipv6-over-5g-v2x-02). Support for these methods and

settings require minimal changes to the existing IPv6 protocol stack, including limitations associated with using these methods, (2) IETF draft-ietf-raw-use-cases-11 describes Reliable and Available Wireless (RAW) Use Cases. This document presents wireless use cases (such as aeronautical communications, amusement parks, industrial applications, pro audio and video, gaming, Unmanned Aerial Vehicle (UAV) and vehicle-to-vehicle (V2V) control, edge robotics, and emergency vehicles), demanding reliable and available behavior. There has been a nice blog by Jari Arkko and Jeff Tantsura (both IAB Members) on 5G and Internet technologies on how the former affects the latter, and have categorized the interactions between the two technologies in several categories.

5G Americas

5G Americas is an industry trade organization composed of leading telecommunications service providers and manufacturers. The organization's mission is to facilitate and advocate for the advancement of 5G and beyond throughout the Americas. During the last year, they published a number of white papers, such as 5G Use Cases, Open RAN Update, Energy Efficiency and Sustainability in Mobile Communications Networks, Evolving Devices for 5G Adoption, Transport Networks for 5G. In December 2023, they held a 6G Evolution Summit.

Alliance for Telecommunications Industry Solutions (ATIS)

ATIS' mission is to advance ICT industry transformation with a focus on 6G by (i) Taking a fast-forward approach to the 6G future through their new initiative, called Next G Alliance, (ii) Incorporating environmental sustainability goals into future networks, and (iii) Architecting the industry-led solutions and standards that create a more competitive and secure landscape for technology-dependent goods and services. Next G Alliance provides the foundation for North America to develop a 6G national roadmap and set of research priorities. It is a broad initiative that addresses the full wireless technology lifecycle from research to commercialization, and engages with diverse ecosystem of stakeholders. The two organizations have published roadmaps for 6G and related technologies, organized several events, including Webinars, while driving the beyond 5G strategy and North American global leadership over the 5G evolutionary path and 6G early development.

Institute of Electrical and Electronics Engineers Standards Association (IEEE SA)

IEEE SA develops global standards in a broad range of industries, including: power and energy, artificial intelligence systems, internet of things, consumer technology and consumer electronics, biomedical and health care, learning technology, information technology and robotics, telecommunication, automotive, transportation, home automation, nanotechnology, information assurance, emerging technologies, and many more. IEEE is well known for having developed standards IEEE 802.1 through 802.24 related to Ethernet, WiFi, optical and LAN/MAN technologies. The Future Networks Technical Community (FNTC) is dedicated to advancing technologies related to 5G, B5G and 6G technologies.

ITU

International Telecommunications Union's (ITU's) World Radio Communication Conference 2023 (WRC-23), held in November

2023 in Dubai, UAE, was the biggest Radio communication event of the year. Decisions made at this conference laid the foundation for the future of the wireless world for the next several years. WRC-23 conference was attended by over 3,900 delegates from 163 Member States, including 88 ministerial-level participants.

Another landmark achievement of the WRC-23 was the 22% representation of women among delegates, marking a 4-percentage-point climb since 2019 in female participation, marking a crucial step towards gender parity in the critical arena of radio communication.

WRCs are held every three to four years. The next one will be in 2027 and one after that in 2031. It is the role of WRC to review, and, if necessary, revise the Radio Regulations, the international treaty governing the use of the radio-frequency spectrum and the geostationary-satellite and non-geostationary-satellite orbits. Revisions to the Radio Regulations are made on the basis of an agenda determined by the ITU Council, which takes into account recommendations made by previous world WRCs. The general scope of the agenda of WRCs is established four to six years in advance, with the final agenda set by the ITU Council two years before the conference, with the concurrence of a majority of Member States. Agenda for WRC 2023 was set in 2019.

On the basis of the contributions from administrations, and common proposals from Regional Groups such as CITELE (Organization of American States), CEPT (representing the European countries), APT (representing the Asian and Pacific countries), RCC (representing the Russian bloc countries), ATU (representing the African countries), ASMG (representing the Arabic countries), and the studies carried out by the ITU study groups during the last four years, WRCs make decisions on various regulatory, technical, operational and procedural matters relating to all wireless services. It must be noted that WRCs are treaty conferences and binding on member states/administrations which ratify the outcome of these conferences. At the end of a WRC, each member state signs the treaty. WRCs are one of the most important conferences within the ITU as this is where access and future spectrum decisions are made. The WRC usually is attended by many of the 193 member states and the delegations of many states are led by the senior government officials and ministers.

Major outcomes of the WRC 2023 are:

1. WRC-23 identified spectrum for International Mobile Telecommunications (IMT), which will be crucial for expanding broadband connectivity and developing IMT mobile services, also known as 4G, 5G and, in the future, 6G. That new spectrum includes the 3,300-3,400 megahertz (MHz), 3,600-3,800 MHz, 4,800-4,990 MHz and 6,425-7,125 MHz frequency bands in various countries and regions. The WRC also recognised the use of 6,425 to 7,125 MHz for wireless access and RLANS in the Radio Regulation table of frequency allocations for the first time.
2. WRC-23 also identified several IMT bands below 2.7 GHz for using high-altitude platform stations as IMT base stations (HIBS) and established regulations for their operations. This technology offers a new platform to provide mobile broadband with minimal infrastructure using the same frequencies and devices as IMT mobile networks. HIBS can contribute to bridging the digital divide in remote and rural areas and maintain connectivity during disasters.
3. WRC-23 also approved new studies for identification of IMT in 4.4-4.8 GHz, 7-8 GHz and 15 GHz for additional 2 GHz mid-band spectrum for 4G, 5G and 6G at WRC-27.
4. WRC-23 also approved new studies for use of LEO and GEO satellites for direct connectivity to mobile phones using IMT technology.
5. For non-geostationary fixed-satellite service Earth Stations in Motion (ESIMs), WRC-23 identified new frequencies to deliver high-speed broadband on-board aircraft, vessels, trains, and vehicles. These satellite services are also critical during disasters where local communication infrastructure is damaged or destroyed.
6. To support the modernization of the Global Maritime Distress and Safety System (GMDSS), WRC-23 took regulatory actions including the implementation of e-navigation systems to enhance distress and safety communications at sea.
7. WRC-23 also decided to protect ship and aircraft mobile service stations located in international airspace and waters from other stations within national territories.
8. WRC-23 allocated new frequencies to the aviation industry for aeronautical mobile satellite services (117.975-137 MHz).
9. WRC-23 agreed to allocate additional frequencies for passive Earth exploration satellite services to enable advanced ice cloud measurements for better weather forecasting and climate monitoring.
10. WRC-23 agreed to allocate frequency bands 15.41-15.7 GHz and 22-22.2 GHz in Radio Regulations Region 1 and some Region 3 countries for non-safety aeronautical applications. This will enable aircraft, helicopters, and drones to carry sophisticated aeronautical digital equipment for purposes such as surveillance, monitoring, mapping, and filming, and have the capacity to transfer large data from these applications using wideband radio links.
11. WRC-23 agreed for the importance of space weather observation in a new Resolution and a new Article in the Radio Regulations to recognize the operation of space weather sensors as part of the meteorological aid service to observe space weather phenomena including solar flares, solar radiation and geomagnetic storms which can interfere with radio communication services including satellites, mobile phone services and navigation systems.
12. WRC-23 agreed for the development of regulatory measures to limit the unauthorized operations of non-geostationary-satellite orbit (non-GSO) earth stations in the fixed-satellite service (FSS) and mobile-satellite service (MSS). It was also agreed to develop technical and regulatory measures for fixed satellite systems (FSS) while taking into account the specific needs of developing countries including the need for equitable access to the relevant frequency bands.
13. WRC-23 approved the recommendation of the Radio Regulations Board to allow 41 countries to acquire new and usable orbital resources for satellite broadcasting. The countries were unable to use their assigned orbital slots in recent years due to factors such as lack of coordination and interference from other satellite networks.
14. It was also agreed for allocations to a possible new or modified space research service (space-to-space) for future development of communications on the lunar surface, and between lunar orbit and the lunar surface.

In all, WRC-23 approved 43 new resolutions, revised 56 existing ones, and suppressed 33 resolutions and also approved the agenda items for the next World Radio Communication Conference (WRC-27) and the provisional agenda for WRC-31. In addition to the above, many more decisions were taken during WRC-23. The conference proved to be the most successful event, as it brought out significant results that will contribute to the advancement of numerous radio services, serving the interests of countries, societies, and humanity at large.

In addition, following recommendations were adopted:

- Agreement on “IMT-2030” as the technical reference for the 6th generation of mobile systems.
- Revision of ITU-R Resolution 65, paving the way for studies on the compatibility of current regulations with potential

6th generation mobile system radio interface technologies for 2030 and beyond.

- Adoption of the new Recommendation ITU-R M.2160 on the “IMT-2030 Framework,” setting the basis for the development of IMT-2030. The next phase will be the definition of relevant requirements and evaluation criteria for potential radio interface technologies (RIT).
- Adoption of a new resolution on the use of IMT technologies for fixed wireless broadband.

Standards column contributed by Bharat Bhatia, Sudhir Dixit, HyeonWoo Lee, Kohei Satoh, Bin Tan, Vino Vinodrai

Advancements in Deep-Learning-Based Object Detection in Challenging Environments

Abhishek Mukhopadhyay and Pradipta Biswas*

Abstract: This article focuses on object detection in challenging environments, where objects of interest need to be detected in images captured under unconstrained conditions. These environments can include outdoor scenes with varying lighting, weather conditions, and background clutter. Object detection in such scenarios is crucial for applications like autonomous driving, surveillance, and industrial production inspection. The paper explores existing techniques for object detection in challenging environments and proposes novel solutions to enhance its performance. One key aspect emphasized in the paper is the need for relevant datasets to validate models under unconstrained scenarios. While several datasets already exist, the research identifies gaps in specific situations and addresses them by proposing new datasets, such as the Indian Lane Dataset for autonomous vehicles. The proposed object detection framework leverages spatial information to detect objects in challenging environments, and its performance is evaluated against benchmark datasets and the newly proposed datasets. A real-world application demonstrates significant improvements in object detection performance in natural environments.

Keywords: Object detection, convolutional neural network, lane detection, novel traffic participants, automatic taxi.

1. Introduction

Object detection is fundamental research in computer vision to perceive the environment and localize objects. Modern object detection applications have been prevalent since the early 1960s. Its first applications were in office automation-related tasks, such as character pattern recognition systems and assembly & verification processes in the semiconductor industry. These applications directly contributed to European countries' economic development [1, 2]. Despite advancements, object detection remains challenging among all computer visual tasks. Analysing a scene and recognizing all the object's constituents is a daunting task [3]. Object detection is challenging for several reasons, like variations in object appearance, lighting conditions,

scale, occlusions, and cluttered backgrounds. Zou et al. [4] provide a comprehensive overview of the history, challenges, and recent advances in object detection. The authors discuss the limitations of traditional object detection approaches, such as sliding window and feature-based methods, and the emergence of deep-learning-based methods. Liu et al. [5] provide an in-depth review of deep-learning-based object detection methods. The authors discuss the challenges of object detection, such as scale variation, occlusion, and multi-object detection, and how deep-learning-based methods address these challenges. Objects can be partially or fully occluded by other objects, making them difficult to detect. This is particularly true in industrial production or road participants detection, where checking the orientation of components becomes challenging due to illuminations or small objects being difficult to detect due to the high density of road participants. Variability in shape, size, and texture makes it challenging to detect them using a fixed set of criteria. The cluttered background is another challenging situation. Objects can be camouflaged by the background, making them difficult to distinguish. The requirement of large-scale datasets, robustness, adaptability to different environments and situations, and optimization towards speed and accuracy are the criterion for acceptance of any object detection model for real-time use. Traditional object detection techniques have several limitations, including limited flexibility, difficulty handling complex scenarios, and computational cost. Traditional object detection techniques can be computationally expensive, especially when processing large images or video streams. The performance of the traditional object detection model reached saturation point in 2010 [4]. The rebirth of the convolutional neural network changed the scenario [6]. Deep learning approaches can learn robust and high-level features from the image specific to the object class being detected, allowing them to adapt to new object classes or variations in object appearance. The progress in the computational system makes them faster and more efficient than traditional models. Overall, all these models broadly can be categorized into four classes, (i) Two-stage models, (ii) Single stage models, (iii) Transformer based models, and (iv) Segmentation based models. Table 1 explains the advantages and disadvantages of all four types of models. In recent time, researchers [4] suggested that semantic segmentation can improve object detection because of its more precise object localization, better feature representation, improved context modelling and so on.

This article describes one such novel hybrid semantic segmentation models for addressing object detection in challenging environments. It shows the application in two areas important for

Center for Product Design and Manufacturing, Indian Institute of Science, Bangalore, India

E-mail: abhishek mukh@iisc.ac.in; pradipta@iisc.ac.in

*Corresponding Author

Manuscript received 08 August 2023, accepted 02 January 2024, and ready for publication 21 March 2024.

© 2024 River Publishers

autonomous vehicle. Finally, it describes one such case study to show how the model can be useful in real-world applications.

2. Architecture of Semantic Segmentation Model

A model is developed combining a dilated convolution branch in parallel to the encoder-decoder branch, inspired by the work of Badrinarayanan et al. [7]. The encoder part of the model utilizes the first three convolutional blocks of the Visual Geometry Group (VGG) 16 network [8] to extract image features. In the decoder part, each layer upsamples the feature maps corresponding to its encoder counterpart using memorized max pooling indices. These sparse feature maps are then convolved with decoder filters to produce dense feature maps. However, while predicting lane markings, segmentation methods based on the encoder-decoder architecture may struggle to preserve global context, smoothness, and continuity in the presence of occlusions and other road objects [9]. To address this, dilated convolutional layers are incorporated in parallel to the encoder-decoder branch. These layers enrich the feature map by leveraging the low-level shape features of lanes. The proposed dilated convolutional network consists of 5 convolutional layers that apply 3×3 convolutions with different dilation rates. A 1×1 convolution layer is used at the end to ensure the same number of channels as the input. Ultimately, a hybrid structure is constructed by employing a weighted summation of the outputs from the encoder-decoder branch and the dilated convolution network. The weights α and β are used to obtain the weighted sum of the two branches, representing the confidence scores for each model's prediction. Another branch, comprising convolutional and fully connected layers, is trained separately to predict α and β based on the input image. After feature fusion, a fully convolutional layer is employed for classifying objects of interest and background pixels. The block diagram of proposed architecture is depicted in Figure 1. In the following two sections, we have explained working of this

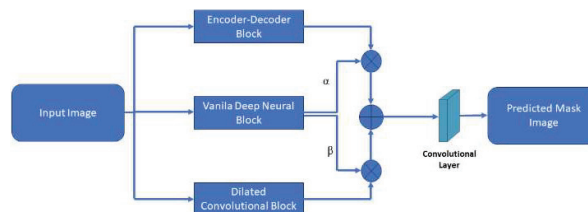


Figure 1.

The proposed semantic segmentation model.

model for two solve two different problems, i.e., lane detection from unconstrained road environments and irregular road object detection.

2.1. Lane Detection from Unstructured Road Conditions

Lane detection is apparently a solved problem. However, there are a plethora of challenges in implementing it for real-life applications. The challenges in this problem include handling different road conditions, such as curved and straight roads, shadows, occlusions, and varying lighting conditions. These difficulties are addressed by creating an unstructured, challenging lane dataset. In this context, unstructured road scenarios exhibit ill-defined lane markings, ambiguous road boundaries, a diverse range of traffic participants, and variations in ambient conditions. The India Driving Dataset (IDD) [10] fulfils the requirements for an unstructured road scenario. The Indian Lane Dataset (ILD) was prepared from the India Driving Dataset (IDD) Segmentation (IDD 20k Part II) dataset [10] by making modifications. Computer Vision Annotation Tools (CVAT) [11] were employed to annotate the lanes manually. Lane markings were labeled and assigned indices that increased

Table 1.

Comparing performance of different object detection models on IDD dataset				
Models	Single-stage Object Detection	Two-stage Object Detection	Transformer-based Object Detection	Semantic Segmentation-based Object Detection
Advantages	<ul style="list-style-type: none"> – Simpler architecture – Faster inference speed – Fewer hyperparameters – Better for real-time applications 	<ul style="list-style-type: none"> – Higher accuracy – Robust to occlusions – Handling small objects – Strong performance on large-scale datasets 	<ul style="list-style-type: none"> – Better handling of global context – Adaptive receptive fields – Enhanced attention mechanisms – Can capture long-range dependencies 	<ul style="list-style-type: none"> – Precise object boundaries – Accurate pixel-level labelling – Good for highly textured objects – Good for segmenting instances of same class
Disadvantages	<ul style="list-style-type: none"> – Lower accuracy – Limited handling of occlusions – Difficulty with small objects – May miss objects at different scales 	<ul style="list-style-type: none"> – Slower inference speed – More complex architecture – More hyperparameters – Sensitivity to initialization 	<ul style="list-style-type: none"> – Higher computational cost – Longer training time – Memory-intensive – Large model size 	<ul style="list-style-type: none"> – Segmentation-based methods involve dense prediction, leading to slower inference – Challenging for objects with fine details – Semantic segmentation may struggle to detect small or heavily occluded objects – Requires additional memory to store per-pixel labels

Table 2.

Comparing performance of different lane detection models on ILD dataset by using IoU score						
Model	Low Light	Shadow	Curve	Highway	Normal	mIoU
DLF model	0.046	0.062	0.082	0.117	0.063	0.072
LaneNet	0.028	0.075	0.098	0.136	0.077	0.082
SCNN	0.022	0.130	0.153	0.187	0.126	0.124
RESA	0.050	0.192	0.269	0.332	0.228	0.214
Our model	0.044	0.248	0.433	0.535	0.281	0.308

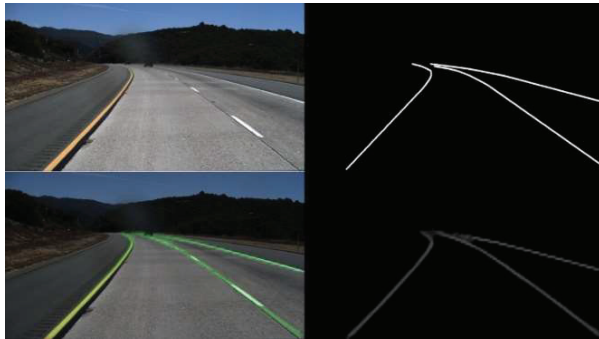


Figure 2.

Performance of the model in detecting lane. Clockwise: Input image, ground truth label, predicted mask image, superimposed predicted mask on original image.

from left to right. In cases where lane markings were not visible or absent, road boundaries and road dividers were annotated as lanes. The IDD dataset was analyzed and categorized into subgroups based on lighting and road conditions, such as no lane, low light, shadow, highway, curve, and normal. Finally, a labelled dataset of 614 images were created and named as Indian Lane Dataset (ILD). The proposed segmentation model is structured for single class segmentation, i.e., 1 for lane and 0 for non-lane pixels.

The proposed model was evaluated individually on five categories (low light, shadow, curve, highway, and normal) in ILD test images and was compared with four other models, including Discriminative Loss Function (DLF) based model [12], RESA [9], SCNN [13] and LaneNet [17]. A detailed comparison chart explaining performance between the proposed model and other models is reported in Table 2. It was found that irrespective of the road scenario, the proposed model achieved 42% improvement over RESA [9]. However, Intersection over Union (IOU) score was lower than RESA by 1% in low-light conditions. In this context, IoU measures the overlap between predicted and ground truth bounding boxes or segmentation masks.

2.2. Detection of Non-Conventional Traffic Participants

Another case study explores the detection of novel road participants, including pedestrians, autorickshaw, different type of trucks, motorcycles and riders in the challenging Indian road

Table 3.

Comparing performance of different object detection models on IDD dataset			
Models	F1 Score	Mean IoU	Latency (FPS)
YOLOv3	0.538	0.356	46.380
YOLOv4	0.403	0.267	30.616
YOLOv5	0.5341	0.355	29.858
YOLOv6	0.468	0.325	45.601
YOLOv7	0.557	0.393	60.259
RetinaNet	0.410	0.290	6.180
Mask RCNN	0.291	0.226	0.602
DeTR	0.356	0.234	0.355
UNet	0.134	0.070	6.250
I-ROD	0.478	0.363	11.638

environment. The unique characteristics of Indian roads, such as congestion, a diverse range of vehicles and transportation methods, and variations in object appearances, pose challenges for accurate detection. Initially, three state-of-the-art object detection models (Mask R-CNN, RetinaNet, YOLOv3) were compared, with YOLOv3 demonstrating superior performance in terms of both accuracy and latency [15]. The proposed segmentation model is modified to address the multi-class semantic segmentation problem.

The proposed model aims to detect a diverse range of road users in unconstrained road environments like in India. It should be trained and evaluated in this environment to make the model robust and accurate. India Driving Dataset [10] fulfills all these criteria as it covers a diversity of vehicles and pedestrians, ambient conditions, and so on. The proposed model is evaluated against nine other models, including two-stage models (Mask RCNN), single-stage models (YOLOv3, YOLOv4, YOLOv5, YOLOv6, YOLOv7, Retinanet), a transformer-based model (DETR), and a segmentation-based model (UNet). The evaluation metric used for comparison is the average Intersection Over Union (IOU) score. Table 3 provides a summary of the results, including mean IOU, F1 score, and latency. All models are tested on the same system (NVIDIA GeForce RTX 2070 GPU) to ensure consistent reporting of accuracy and latency. While YOLOv7 demonstrated better accuracy compared to proposed model, it is worth noting that researchers [14] argue in favor of pixel-wise predictions over bounding box-based predictions. Pixel-wise predictions assess the probability of an object's presence at each pixel, enabling more precise localization and segmentation, even in dense and cluttered scenes. They are also more effective at handling non-rigid objects and complex object shapes. In light of these arguments, this study aims to compare YOLOv7 and proposed model on a pixel-wise basis. To obtain pixel-wise segmented results from YOLOv7, the predicted bounding boxes (indicated by the green box in the YOLOv7 output image in Figure 3) are considered and converted into a segmented image. These segmented images are then compared with ground truth label images. The study reveals that the accuracy (IoU) drops from 0.55 to 0.39 when compared IoU of 0.45 proposed segmentation model. Figure 3 provides a detailed explanation of the comparison.

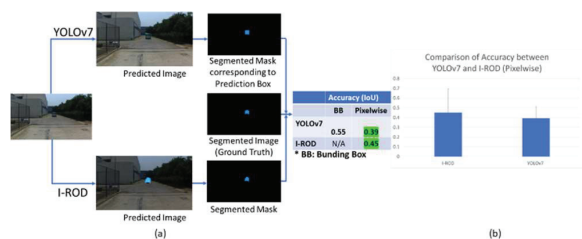


Figure 3.

A Comparison study between YOLOv7 and proposed model to show superiority of semantic segmentation models. This figure is best viewed in electronic form.

3. Application: Automatic Taxiing of Aircraft

The proposed lane and object detection model was successfully applied in the development of a fully autonomous system for aircraft taxiing. This system relied on the lane detection model as a backbone for lane navigation, while the object detection model played a crucial role in collision avoidance. Liu et al. [18] pursued a methodology akin to ours, framing the issue within a controlled simulation environment. This system use a lane and obstacle detection algorithm benchmarked against state-of-the-art models before implementation in our system. The experimental arrangement features a model replicating a taxiway, incorporating a light and lane indicator, as shown in supplementary video.¹ The green light guides the path toward the runway or hangar, while the red light serves as a stopping boundary for the aircraft. For this study, the TurtleBot3 serves as a substitute for the actual aircraft. This programmable Robot Operating System (ROS)-based mobile robot was chosen due to its ease of control via ROS and its capacity for sensor attachment, such as a camera. A Microsoft Livecam, affixed to the TurtleBot3, provides a vertical field of view (FOV) of 37.73 degrees, a horizontal FOV of 57.88 degrees, and covers distances up to 35 cm in the frontal direction.

For lane navigation, the system employed lane detection followed by a lane navigation algorithm to ensure that the aircraft stays on the center line and halts before entering the runway. The lane detection model provided points that were used in the control generation algorithm to generate inputs for autonomous navigation of the aircraft. The algorithm utilized an image mask

Table 4.

Accuracy analysis of the object detection model for Automatic taxiing (class wise)			
Classes	IoU	Precision	F1 Score
Airplane	0.306	0.694	0.455
Bus	0.356	0.427	0.422
Car	0.277	0.539	0.405
Other Vehicles	0.078	0.413	0.156
Person	0.238	0.515	0.399
Truck	0.270	0.934	0.436

¹ https://www.youtube.com/watch?v=A_qG_a5w7lc.

Table 5.

Accuracy analysis of the object detection model for Automatic taxiing (overall)	
Observed Units	Hybrid Models
mIoU	0.262
Precision	0.587
Recall	0.298
F1 Score	0.395
Latency (in milliseconds)	114.348

containing only the detected lane points obtained from the lane detection algorithm. The 'steerBias' parameter indicated the distance between the camera's center and the middle lane in the x direction. If the center lane was positioned to the left of the camera center, the error was considered positive; otherwise, it was negative.

To prevent collisions, the object detection model was fine-tuned using an airport dataset before integrating it into the proposed system. The dataset consisted of six classes of objects, including airplanes, buses, cars, other vehicles, persons, and trucks. The "other vehicles" category encompassed dollies, pushback tugs, and tractors. The overall accuracy (mean IoU) of the object detection model was reported as 0.262, with a processing speed of 8.742 frames per second. Tables 4 and 5 provide a summary of the model's performance in the system.

4. Discussions and Conclusions

This article makes significant contributions to the field of object detection research, focusing on both novelty and practical applications. A novel approach for lane detection in unconstrained environments is proposed, demonstrating real-time performance and outperforming other models. A new lane dataset is introduced and compared with existing models, showcasing the proposed model's superior accuracy and robustness across unseen environments, making it suitable for various applications. Further details of this work can be found in [16]. The article also addresses the limitations of state-of-the-art bounding box-based models in precise object localization, particularly in dense and crowded environments. To overcome this challenge, an in-depth study is conducted, leading to the proposal of a new object detection model with pixel-wise localization. This model exhibits improved performance in critical scenarios where precision is crucial. Moreover, the article presents a novel automated system for aircraft taxiing that integrates the lane and object detection algorithms to provide collision avoidance and real-time assistance. The navigation and collision avoidance system demonstrate efficacy in different lighting conditions and complex scenarios.² To summarize, this article contributes novel approaches to lane detection and object detection, along with the development of an automated system for taxiing aircraft, showcasing the practical applications of the proposed algorithms in real-time situations.

² https://www.youtube.com/watch?v=A_qG_a5w7lc.

References

- [1] L. G. Roberts. Pattern recognition with an adaptive network. In in: Proc. IRE International Convention Record, pages 66–70, 1960.
- [2] James T Tippett, David A Borkowitz, Lewis C Clapp, Charles J Koester, and Alexander Vanderburgh Jr. Optical and electro-optical information processing. Technical report, Massachusetts Inst of Tech Cambridge, 1965.
- [3] Richard Szeliski. Computer vision: algorithms and applications. Springer Nature, 2022.
- [4] Zhengxia Zou, Keyan Chen, Zhenwei Shi, Yuhong Guo, and Jieping Ye. Object detection in 20 years: A survey. *Proceedings of the IEEE*, 2023.
- [5] Li Liu, Wanli Ouyang, Xiaogang Wang, Paul Fieguth, Jie Chen, Xinwang Liu, and Matti Pietikäinen. Deep learning for generic object detection: A survey. *International journal of computer vision*, 128:261–318, 2020.
- [6] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6):84–90, 2017.
- [7] Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 39(12):2481–2495, 2017.
- [8] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.
- [9] Tu Zheng, Hao Fang, Yi Zhang, Wenjian Tang, Zheng Yang, Haifeng Liu, and Deng Cai. Resa: Recurrent feature-shift aggregator for lane detection. *arXiv preprint arXiv:2008.13719*, 2020.
- [10] Girish Varma, Anbumani Subramanian, Anoop Nambodiri, Manmohan Chandraker, and CV Jawahar. Idd: A dataset for exploring problems of autonomous navigation in unconstrained environments. In *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 1743–1751. IEEE, 2019.
- [11] Intel Community. Computer vision annotation tool (cvat). <https://github.com/openvinotoolkit/cvat/>.
- [12] Bert De Brabandere, Davy Neven, and Luc Van Gool. Semantic instance segmentation with a discriminative loss function. *arXiv preprint arXiv:1708.02551*, 2017.
- [13] Xingang Pan, Jianping Shi, Ping Luo, Xiaogang Wang, and Xiaoou Tang. Spatial as deep: Spatial cnn for traffic scene understanding. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2018.
- [14] Ali Farhadi and Joseph Redmon. Yolov3: An incremental improvement. In *Computer vision and pattern recognition*, volume 1804, pages 1–6. Springer Berlin/Heidelberg, Germany, 2018.
- [15] Abhishek Mukhopadhyay, Imon Mukherjee, and Pradipta Biswas. Comparing cnns for non-conventional traffic participants. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications: Adjunct Proceedings*, pages 171–175, 2019.
- [16] Abhishek Mukhopadhyay, L. R. D. Murthy, Imon Mukherjee, and Pradipta Biswas. “A hybrid lane detection model for wild road conditions.” *IEEE Transactions on Artificial Intelligence*, 1–10, 2022.
- [17] Z. Wang, W. Ren, and Q. Qiu, “Lanenet: Real-time lane detection networks for autonomous driving,” *arXiv preprint arXiv:1807.01726*, 2018.
- [18] C. Liu and S. Ferrari, “Vision-guided planning and control for autonomous taxiing via convolutional neural networks,” in *AIAA Scitech 2019 Forum*, p. 0928, 2019.

Biographies



Abhishek Mukhopadhyay is working as Post Doctoral Researcher in I3D lab at Centre for Product Design and Manufacturing, Indian Institute of Science Bangalore. He completed his PhD in Computer Science and Engineering from IIT Kalyani. He works in the area of Computer Vision, Object Detection, Machine Learning, Virtual Reality, HCI. He worked with Robert Bosch, Wipro, Faurecia, British Telecom in his research tenure. He achieved one best paper award (47th WWRF meeting, Bristol, 2022), ACM global fellowship to present his work in AutomotiveUI 2019 Doctoral Colloquium in the Netherlands, student travel grant in WWRF 48 in Abu Dhabi. Before joining PhD curriculum, he worked as assistant professor for 6 years in multiple Institutions in Kolkata, India.



Pradipta Biswas is an Associate Professor at the Centre for Product Design and Manufacturing and associate faculty at the Robert Bosch Centre for Cyber Physical Systems of Indian Institute of Science. He is a vice chairman of ITU Study Group 9 and also a Co-Chair of the IRG AVA and Focus Group on Smart TV at International Telecommunication Union. His research focuses on user modelling and multimodal human-machine interaction for aviation and automotive environments and for assistive technology. Earlier, he was a Senior Research Associate at Engineering Department, Research Fellow at Wolfson College and Research Associate at Trinity Hall of University of Cambridge. He completed Ph.D. in Computer Science at the Rainbow Group of University of Cambridge Computer Laboratory and Trinity College in 2010 and was awarded a Gates-Cambridge Scholarship in 2006.

Emotion Detection from Speech: A Comprehensive Approach Using Speech-to-Text Transcription and Ensemble Learning

Fatima M Inamdar^{1,}, Sateesh Ambesange², Parikshit Mahalle¹, Nilesh P. Sable¹, Ritesh Bachhav¹, Chaitanya Ganjiwale¹, Shantanu Badmanji¹ and Sarthak Agase¹*

Abstract: The field of text-to-emotion analysis is investigated in this study, which uses an interactive methodology to reveal subtle emotional insights in textual data. The research explores the complex relationship between language and emotion using sophisticated methods without focusing on any particular frontend or backend technology. The research attempts to improve our understanding of how literary information transmits emotional subtleties by emphasizing a broad but methodical examination. The lack of mentions of particular libraries and backends indicates an emphasis on the general ideas and techniques used in text-to-emotion analysis.

The findings demonstrate the possibility of deriving significant emotional context from text, opening doors for applications in a variety of fields where user sentiment analysis is essential. This study adds to the body of knowledge on emotional intelligence in computational linguistics and lays the groundwork for future developments in text analysis techniques.

Keywords: Text-to-Emotion analysis, natural language processing, sentiment analysis, emotion recognition, computational linguistics, emotional insights, textual sentiment, emotional nuances, user sentiment, language and emotion.

1. Introduction

In the ever-evolving landscape of computational linguistics, the intersection of natural language processing and emotional intelligence has emerged as a focal point of exploration. This research embarks on an insightful journey into the realm of text-to-motion analysis, seeking to unravel the intricate connections between language and sentiment. As the digital age transforms communication paradigms, understanding and interpreting the emotional nuances embedded within textual data becomes paramount. Our

investigation delves into the methods and approaches employed in real-time emotion analysis, employing a lens that avoids strict delineation of backend and frontend technologies, fostering a holistic exploration of the field.

The surge in digital content consumption, especially in the context of news and information dissemination, underscores the need for personalized and engaging experiences. Text-to-Emotion Analysis offers a gateway to comprehend not only the explicit meaning but also the underlying sentiments woven into language. This research aims to contribute to the broader discourse on emotional intelligence in computational linguistics by examining how textual content resonates emotionally with users. By acknowledging the dynamic nature of emotional responses, we aspire to pave the way for enhanced user experiences and applications across diverse domains. In the following sections, we delve into the methodologies, insights, and implications drawn from our exploration of text-to-motion analysis.

2. Literature Survey

This research introduces the Avatar Affordances Framework, a sophisticated tool designed to map design trends and affordances in character creation interfaces. The framework aims to enhance the understanding of interactive character design across diverse applications. It meticulously examines the evolving landscape of character design interfaces, providing valuable insights for designers, developers, and researchers. By mapping affordances, the authors contribute to the field's theoretical foundations, offering a systematic approach for analyzing and optimizing character creation interfaces. The paper emphasizes the importance of character design in interactive systems and lays the groundwork for future advancements in this dynamic and evolving field [1].

The paper, "Face Detection and Recognition Using OpenCV" by Ramadan TH. Hasan and Amira Bibo Sallow aim to investigate the crucial role of OpenCV in computer vision, specifically focusing on face detection and recognition. The research explores various algorithms, modules, and applications of OpenCV, offering a comprehensive assessment of recent literature. The objective is to emphasize OpenCV's diverse applications and its statistical effectiveness in enhancing human life through image and video processing. The discussed technologies include VGGFace, Attitude Tracking Algorithm (EATA), YCbCr Color Model, and LBPH (Local Binary Pattern Histogram). The paper provides a thorough examination of OpenCV's applications,

¹Vishwakarma Institute of Information Technology, Pune, Maharashtra, India

²Pragyan Smartai Technology Llp, Bangalore, Karnataka, India
E-mail: fatima.inamdar@viit.ac.in; sateesh.ambesange@gmail.com; parikshit.mahalle@viit.ac.in; riteshbachhav251@gmail.com; chaitanyaganjiwale812@gmail.com; shantanubadmanji1912@gmail.com; sarthak.agase@gmail.com

* Corresponding Author

Manuscript received 07 November 2023, accepted 02 February 2024, and ready for publication 21 March 2024.

© 2024 River Publishers

shedding light on face detection, recognition, and real-world implementations. However, it falls short of providing an in-depth exploration of certain algorithms, and the literature reviews are confined to specific timeframes, potentially missing recent advancements in the field [2].

This research introduces a groundbreaking model of a humanoid robot equipped with facial expression capabilities and natural language processing (NLP) skills. The authors showcase the practical implementation of this novel model, emphasizing the robot's ability to recognize objects and convey emotions through facial expressions. The integration of NLP suggests a sophisticated level of interaction, allowing the robot to understand and respond to human language. The implementation of such a robot holds promise for advancements in human-robot interaction, particularly in scenarios where robots are expected to comprehend and express emotions, enhancing their utility in various contexts, from customer service to educational settings [3].

Focusing on human-computer interactions, this research presents an emotionally responsive avatar with dynamic facial expressions. The authors delve into the intricacies of designing an avatar capable of conveying emotions responsively. By incorporating dynamic facial expressions, the avatar enhances the interactive experience, allowing for more nuanced and emotionally resonant digital interactions. This work contributes to the growing field of affective computing, where the goal is to imbue computers and digital interfaces with the ability to recognize and respond to human emotions. The advancements presented in this paper have implications for the design of emotionally engaging avatars in applications ranging from virtual assistants to gaming [4].

Addressing challenges in 3D avatar generation, this research proposes a novel method guided by text and image prompts. Leveraging the Contrastive Language Image Pre-training (CLIP) model and a pre-trained 3D Generative Adversarial Network (GAN), the approach enables efficient and high-level manipulations of shape and texture in 3D avatars. The use of CLIP allows for contextual understanding from textual and visual prompts, facilitating more intuitive and user-friendly avatar manipulation. This work contributes to the field of computer graphics and virtual environments, providing a pathway for creating personalized and expressive 3D avatars for various applications, including virtual worlds, gaming, and digital communication [5].

This research focuses on Persian texts, presenting a speech act classifier and its application in identifying rumors. The authors address the challenge of classifying speech acts in Persian language texts, contributing to natural language processing tasks specific to this linguistic context. The application in rumor identification highlights the broader societal implications of the research, as it provides a tool for automatically discerning linguistic expressions associated with rumors. This work contributes not only to the field of natural language processing but also to misinformation detection, offering a valuable resource for analyzing and mitigating the spread of rumors in Persian-language digital communication [6].

The authors systematically review emotion recognition and detection methods in this comprehensive survey. Covering a wide range of approaches and techniques, the survey serves as a valuable resource for researchers in the field of affective computing. The paper categorizes and analyzes various methods employed for recognizing and detecting emotions, considering modalities such as facial expressions, speech, and physiological signals. This work

not only provides an overview of the current state of the art but also identifies trends and challenges in emotion recognition, paving the way for future research and development in this dynamic and evolving field [7].

This research delves into linguistic features to explore the automatic identification of mental states from language text. The authors propose a method to read "mindprints" by analyzing deeper linguistic features, providing insights into the cognitive and emotional states reflected in language. The work contributes to understanding the intricate relationship between language and cognition, offering a nuanced perspective on how linguistic expressions can serve as indicators of mental states. By leveraging deeper linguistic analysis, the research expands the toolkit for studying and interpreting mental states through textual communication, opening avenues for further exploration in cognitive science and natural language processing [8].

This research presents a framework for designing emotionally realistic chatbots, aiming to enhance their believability. Leveraging AIML (Artificial Intelligence Markup Language) and information states, the framework focuses on imbuing chatbots with emotional expression and responsiveness. By incorporating elements of emotion into chatbot interactions, the authors seek to create a more engaging and believable conversational experience. The framework's potential applications range from customer service bots to virtual companions, where emotional intelligence is increasingly recognized as a valuable aspect of human-computer interaction. The work contributes to the development of more sophisticated and emotionally resonant conversational agents in the field of artificial intelligence [9].

3. Existing System and Algorithm

Different approaches are used in the field of text-to-emotion analysis to interpret the emotional undertones included in textual data. Lexicon-based methods use sentiment lexicons or dictionaries to identify the overall emotional tone of a text by giving words sentiment ratings and adding them together. Though basic, these approaches could have trouble with subtleties and expressions that depend on context. However, machine learning models which include deep learning architectures and conventional classifiers use labeled datasets to identify complex patterns and forecast the emotional content of text. Although they need a large amount of training data, these models are excellent at capturing context-dependent emotions and intricate interactions.

Natural Language Processing (NLP) techniques contribute significantly to text-to-emotion analysis by extracting meaningful features from the text. Techniques such as part-of-speech tagging and syntactic analysis aid in understanding the structural and grammatical elements that contribute to the emotional context of the text. Additionally, word embeddings, such as Word2Vec and GloVe, provide continuous vector representations of words, capturing semantic relationships and enriching the understanding of emotional nuances.

Furthermore, rule-based systems explicitly define linguistic rules for identifying emotional content, offering transparency and interpretability. However, these systems may face challenges in handling diverse expressions and linguistic variability. Hybrid approaches, combining rule-based components with machine learning models, aim to harness the interpretability of rule-based

systems and the predictive power of machine learning, offering a potential middle ground for accurate and robust emotion analysis. The choice of methodology depends on the specific requirements, nuances, and context of the text data under consideration.

4. Proposed Methodology

This research paper proposes a comprehensive methodology for emotion detection from speech, leveraging the capabilities of speech-to-text transcription and ensemble learning techniques. The initial phase involves the use of open-source speech-to-text models, particularly the OpenAI Whisper model, to accurately transcribe spoken words into textual representations. These transcriptions undergo preprocessing to address challenges such as noise and accents. Subsequently, the text is processed using natural language processing techniques, including tokenization, normalization, and the implementation of a bag-of-words model to capture emotional content effectively.

In the next step, machine learning classification algorithms are employed to categorize emotions within the transcriptions. Selected algorithms, such as Support Vector Machines, Random Forest, and Naive Bayes, are trained and evaluated using a divided dataset for training and testing purposes. Hyperparameter tuning and optimization are performed through cross-validation to enhance model performance. To further improve accuracy, Ensemble Methods for Classification are explored, combining predictions from multiple base models using techniques like Voting, Bagging, and Boosting.

The evaluation of the proposed methodology relies on standard metrics such as accuracy, precision, recall, and F1 score. Comparative analyses are conducted between individual machine learning models and the ensemble model to discern the impact of ensemble methods on overall performance. The results obtained from these experiments, along with discussions on the strengths and limitations of the methodology, form the core of the findings. Practical implications and potential applications of the proposed approach are deliberated, concluding with suggestions for future research directions in the domain of emotion detection from speech. The paper is substantiated by references to relevant literature, previous studies, and resources that informed the research methodology and analysis.

5. Results

In the pursuit of emotion detection from speech, the research employed a methodology encompassing speech-to-text transcription using the OpenAI Whisper model, subsequent text preprocessing with Natural Language Processing (NLP) techniques facilitated by NLTK tools, and the transformation of text into a bag-of-words model for training machine learning classification models. The dataset chosen for evaluation was the “Emotions dataset for NLP,” consisting of 23,000 sentences, each associated with one of the six emotions: anger, fear, happiness, love, sadness, and surprise.

To assess the individual performance of various machine learning models, accuracy scores were computed for each emotion detection task. The results are detailed as follows:

- Ensemble Classifier: 85.89%

- Neural Network: 80.82%
- GradientBoostingClassifier: 78.68%
- LGBM_classifier: 84.05%
- CatBoost: 84.87%
- XGB: 84.89%
- ExtraTreesClassifier: 85.06%
- Random Forest: 83.69%
- DecisionTree: 79.75%
- LogisticRegression: 84.61%
- Kernel SVM: 81.76%
- Linear SVM: 83.39%
- SVM: 81.76%
- KNeighbours: 53.85%
- Naive Bayes: 37.71%
- AdaBoostClassifier: 39.25%

Accuracy = 85.8947%				
Classification Report:				
	precision	recall	f1-score	support
anger	0.82	0.80	0.81	683
fear	0.85	0.83	0.84	520
happy	0.85	0.91	0.88	1579
love	0.83	0.74	0.78	303
sadness	0.90	0.88	0.89	1403
surprise	0.85	0.74	0.79	184
accuracy			0.86	4672
macro avg	0.85	0.82	0.83	4672
weighted avg	0.86	0.86	0.86	4672

Figure 1. Ensemble model classification report.

Furthermore, the Ensemble Model for Classification, designed to aggregate predictions from multiple base models, demonstrated an elevated accuracy score of 0.859. This result underscores the effectiveness of the ensemble approach in surpassing the performance of individual models, showcasing its potential for accurate and robust emotion detection from speech. The ensemble model’s ability to harmonize diverse classifiers and leverage their collective insights contributes significantly to the overall success of the proposed methodology.

These outcomes underscore the potential practical applications of the research in real-world contexts where nuanced emotion analysis from spoken language is essential. The detailed accuracy scores offer valuable insights into the comparative performance of various machine learning models, affirming the efficacy of the proposed methodology in enhancing emotion detection accuracy.

6. Future Work

The present research on emotion detection from speech lays the foundation for compelling avenues of future exploration. One prominent trajectory involves extending our current methodology to encompass AI-driven video emotion enhancement. This endeavor seeks to leverage identified emotions from speech to dynamically enhance emotional content within corresponding

video footage. Emphasis will be placed on mapping detected emotions to appropriate facial expressions and gestures, thereby elevating the emotional expressiveness of video content. Additionally, we envision integrating the developed emotion detection model with AI News Anchors, aiming to generate or edit videos dynamically. This endeavor involves rendering video content with appropriately matched emotional expressions and gestures based on the emotions detected from the speech or transcript of the news anchor.

Further directions include the exploration of real-time emotion detection and mapping for live speech scenarios, with potential applications in live broadcasts, interviews, and interactive video content. To enhance accuracy, we propose investigating advanced ensemble methods and their combination with deep learning approaches. This includes the exploration of neural network-based models within ensemble frameworks to optimize performance. The extension of our methodology to include multimodal features, incorporating both speech and facial expressions for a more comprehensive understanding of emotional states, is another avenue of interest.

Assessing the resilience of our approach in various linguistic and cultural situations is crucial, as we acknowledge. Increasing the dataset's diversity in terms of language and culture is essential to improving the model's capacity for generalization. To guide changes based on real-world user experiences, user feedback, and experience studies will be carried out to evaluate the perceived efficacy and accuracy of the emotion detection model in practical applications.

Furthermore, ethical considerations associated with emotion detection technology, such as privacy concerns and potential biases, will be addressed. Measures will be implemented to ensure fair and unbiased emotion analysis across various demographic groups. This collective future work aspires to extend the impact of our research into practical applications, revolutionizing the portrayal and perception of emotional content in multimedia formats, particularly in the emerging domain of AI News Anchors and video emotion enhancement.

7. Conclusion

In conclusion, this research endeavors to advance the field of emotion detection from speech by leveraging cutting-edge technologies and methodologies. The utilization of open-source models, particularly the Open AI Whisper model, for speech-to-text transcription, has proven to be a foundational step in transforming spoken words into a textual format. The subsequent processing of this text involves sophisticated Natural Language Processing (NLP) techniques, including the implementation of a bag-of-words model, to extract and categorize emotional states.

The integration of machine learning classification algorithms further enhances the accuracy of emotion detection. Our evaluation encompassed a diverse set of classifiers, demonstrating the efficacy of models such as CatBoost, XGBoost, and ExtraTreesClassifier in accurately discerning emotions from speech. To refine our approach and achieve even greater precision, Ensemble Methods for Classification were employed, showcasing the potential of combining multiple algorithms for improved performance.

Looking forward, the trajectory of our research extends beyond the realms of speech analysis into AI video emotion enhancement. Our vision encompasses the dynamic regeneration of video content based on detected emotions, a significant stride towards more emotionally resonant multimedia experiences. We anticipate transformative applications in content creation, storytelling, and media production by fusing the insights gained from speech analysis with AI-driven tools for video emotion enhancement.

In essence, this research not only contributes to the understanding of emotion detection from speech but also lays the groundwork for innovative applications in the multimedia domain. The fusion of speech-to-text transcription, NLP, machine learning, and the prospect of AI video emotion enhancement opens doors to a future where emotional content is discerned and dynamically translated into rich and expressive visual narratives. As we embark on the journey of implementing our methodology into practical video enhancement tools, we anticipate the continued evolution of emotion detection technologies and their profound impact on the way we experience and interact with multimedia content.

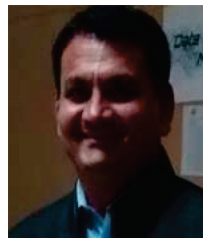
References

- [1] McArthur, V., Teather, R. J., and Jenson, J. (2015). The Avatar Affordances Framework: Mapping Affordances and Design Trends in Character Creation Interfaces. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play* (pp. 231–240). ACM. doi: 10.1145/2793107.2793121.
- [2] Hasan, R. Th., Sallow, A. B., and Hasan, A. B. (2021). Face Detection and Recognition Using OpenCV. *JSCDM*, 2(2). doi: 10.30880/jscdm.2021.02.02.008.
- [3] Budiharto, W., Andreas, V., and Gunawan, A. A. S. (2021). A Novel Model and Implementation of Humanoid Robot with Facial Expression and Natural Language Processing (NLP). *ICIC International 学会*. doi: 10.24507/icicelb.12.03.275.
- [4] Wang, H., Gaddy, V., Beveridge, J. R., and Ortega, F. R. (2021). Building an Emotionally Responsive Avatar with Dynamic Facial Expressions in Human-Computer Interactions. *MTI*, 5(3), 13. doi: 10.3390/mti5030013.
- [5] Canfes, Z., Atasoy, M. F., Dirik, A., and Yanardag, P. (2023). Text and Image Guided 3D Avatar Generation and Manipulation. In *2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)* (pp. 4410–4420). IEEE. doi: 10.1109/WACV56688.2023.00440.
- [6] Jahanbakhsh-Nagadeh, Z., Feizi-Derakhshi, M.-R., and Sharifi, A. (2019). A Speech Act Classifier for Persian Texts and its Application in Identifying Rumors. doi: 10.48550/ARXIV.1901.03904.
- [7] Saxena, A., Khanna, A., and Gupta, D. (2020). Emotion Recognition and Detection Methods: A Comprehensive Survey. *AIS*, 2(1), 53–79. doi: 10.33969/AIS.2020.21005.
- [8] Pearl, L. S., and Enverga, I. (2014). Can you read my mind print?: Automatically identifying mental states from language text using deeper linguistic features. *IS*, 15(3), 359–387. doi: 10.1075/is.15.3.01pea.
- [9] Sutoyo, R., Chowanda, A., Kurniati, A., and Wongso, R. (2019). Designing an Emotionally Realistic Chatbot Framework to Enhance Its Believability with AIML and Information States. *Procedia Computer Science*, 157, 621–628. doi: 10.1016/j.procs.2019.08.226.

- [10] Schuller, B., Steidl, S., Batliner, A., Vinciarelli, A., Scherer, K., Ringeval, F., ... and Marchi, E. (2013). The INTERSPEECH 2013 computational paralinguistics challenge: social signals, conflict, emotion, autism. In *Proceedings of Interspeech* (pp. 148–152). doi: 10.21437/Interspeech.2013-70.
- [11] Ayadi, M., Kamel, M. S., and Karray, F. (2011). Survey on speech emotion recognition: Features, classification schemes, and databases. *Pattern Recognition*, 44(3), 572–587. doi: 10.1016/j.patcog.2010.09.020.
- [12] Eyben, F., Scherer, K. R., Schuller, B. W., Sundberg, J., André, E., Busso, C., ... and Tavabi, L. (2015). The Geneva Minimalistic acoustic parameter set (GeMAPS) for voice research and affective computing. *IEEE Transactions on Affective Computing*, 7(2), 190–202. doi: 10.1109/TAFFC.2015.2457417.
- [13] Lee, C. M., and Narayanan, S. S. (2005). Toward detecting emotions in spoken dialogs. *IEEE Transactions on Speech and Audio Processing*, 13(2), 293–303. doi: 10.1109/TSA.2004.840609.
- [14] El Ayadi, M., and Kamel, M. S. (2007). A new framework for audio-visual speech emotion recognition. *Pattern Recognition*, 40(6), 1674–1687. doi: 10.1016/j.patcog.2006.11.029.
- [15] Busso, C., Lee, S., and Narayanan, S. S. (2009). Analysis of emotionally salient aspects of fundamental frequency for emotion detection. *IEEE Transactions on Audio, Speech, and Language Processing*, 17(4), 582–596. doi: 10.1109/TASL.2008.2010426.
- [16] Deng, J., Zhang, Z., Marchi, E., Schuller, B., and Wu, D. (2013). Sparse autoencoder-based feature transfer learning for speech emotion recognition. In *Proceedings of Interspeech* (pp. 236–240). doi: 10.21437/Interspeech.2013-79.
- [17] Koolagudi, S. G., and Rao, K. S. (2012). Emotion recognition from speech: A review. *International Journal of Speech Technology*, 15(2), 99–117. doi: 10.1007/s10772-012-9157-x.
- [18] Lotfian, R., and Saeidi, R. (2016). Speech emotion recognition using hidden Markov models. *Speech Communication*, 77, 1–17. doi: 10.1016/j.specom.2015.10.010.
- [19] Li, C., and Deng, J. (2014). Emotion recognition from speech signals using new harmony features. *IEEE Transactions on Multimedia*, 16(7), 1904–1916. doi: 10.1109/TMM.2014.2323633.



Sateesh Ambesange, with 20+ years of industry expertise, is presently linked with PRAGYAN SMARTAI TECHNOLOGY LLP. A committed researcher in Data Science, Machine Learning, Deep Learning, Natural Language Processing (NLP), and Generative AI, he excels in these domains. As a leader, he takes pride in contributing significantly to advancements in artificial intelligence and data-driven solutions.



Parikshit Mahalle, a senior IEEE member and Professor at Vishwakarma Institute of Information Technology, Pune, holds 23+ years in teaching and research. With a Ph.D. from Aalborg University, Denmark, and a postdoc at CMI, Copenhagen, he boasts 15 patents, 200+ publications, and 59 books. As an editorial leader at IGI Global's International Journal of Rough Sets and Data Analysis, he significantly influences Machine Learning, Data Science, and IoT, earning accolades like the "Best Faculty Award".

Biographies



Fatima M Inamdar, Assistant Professor at Vishwakarma Institute of Information Technology, holds 18 years of teaching and 7 months of industry expertise. Proficient in Software Engineering, Data Science, AI-ML+DL, Web Technology, she's a Fellow at Eudoxia Research University. Recognized for research excellence, awards from SAW and the Global Research Foundation underscore her impact. The Women Leaders Forum award highlights her commitment to medical sector challenges.



Nilesh P. Sable is a senior member IEEE and working as Associate Professor, Head Department of Computer Science & Engineering (Artificial Intelligence) at Vishwakarma Institute of Information Technology, Pune, India. He has 16+ years of teaching and research experience. He is guiding 4 Ph.D. students in the area of Machine Learning, Federated Learning and IoT under his supervision from SPPU. He is working as Research Advisory Committee (RAC) Member for various Research Centres. His research interests are Data Mining, Image Processing,

Machine Learning, Cognitive Computing, Internet of Things, Networking and Security. He has published 75+ papers in National, International conferences and Journals. He had Filed and Published 16 Patents and 18 Copyrights. He has authored books published by National/International publishers.



Ritesh Bachhav, currently pursuing a B.Tech in Information Technology at Vishwakarma Institute of Information Technology, seamlessly blends academic excellence with active participation in hackathons and coding competitions. Driven by a profound passion for technology, he specializes in Android development and Blockchain.



Chaitanya Ganjiwale, currently pursuing a B.Tech in Information Technology at Vishwakarma Institute of Information Technology, actively engages in competitive coding, hackathons, and coding competitions. Passionate about technology, especially Android development, networking, and machine learning, I possess a quick-learning mindset and an unwavering commitment to excellence, thriving in diverse challenges and opportunities.



Shantanu Badmanji, a graduate pursuing a B.Tech in Information Technology from Vishwakarma Institute of Information Technology, seamlessly integrates a robust academic background with active involvement in hackathons and coding competitions. Fueled by a passion for technology, an innovative mindset, and a commitment to excellence, I am poised to be a promising contributor to the IT industry.



Sarthak Agase, pursuing B.Tech in IT from VIIT Pune, looking forward to improving my managerial abilities and leadership skills, interested in Machine Learning, Web Development and Blockchain Technologies.

Beam-Tracking Challenges in THz Communications

*Giorgos Stratidakis, Sotiris Droulias and Angeliki Alexiou**

Abstract: In recent years, the demand for high data rates has increased drastically. In response to this increasing demand, higher frequencies, such as millimeter wave (mmWave) and terahertz (THz) have been considered that offer much larger bandwidth and potentially much higher data rates. Nevertheless, with the increase in frequency, the pathloss increases significantly, with the non-line-of-sight (nLoS) incurring attenuation levels that can dramatically reduce the quality of a wireless link. The use of directional antennas to counteract the increased pathloss is widely accepted. However, highly directional beams are prone to misalignment and a method to track the tracking object (TO) of the link with consistency, accuracy and low overhead is needed. To this end, the design of beam-tracking algorithms has been proposed. In this work, the main parameters that affect the reliability of beam-tracking are presented along with the challenges that beam-tracking algorithms need to address. Furthermore, the performance merits with respect to three reliability parameters are presented in the case of a simple beam-tracking algorithm.

Keywords: Beam-forming, beam-tracking, reliability, terahertz.

1. Introduction

For a link to be established with directional antennas, they must be properly aligned or the received power will be reduced which can lead to outage [1–3]. If both the transmitter (TX) and receiver (RX) are static, the alignment of their antennas is simple as it can be done once and make corrective actions to fix any future misalignment (caused by earthquakes, wind, etc). On the other hand, if either the TX or RX is mobile, and at least one of them is equipped with a directional antenna, the one with directional antenna requires a method to track the direction of the other part of the link in order to properly align the beam (or beams). This can be realized with either localization or beam-training/tracking. Localization, beam-training and tracking must be highly accurate for the misalignment to be low and the received power high. The accuracy threshold is dependent on the half-power beamwidth (HPBW) of the beams and is not universal. For example, a localization or beam-training/tracking error that leads to a directional

error of 5° can be acceptable with an HPBW of 10° but not with an HPBW of 2° . In other words, the narrower the HPBW the lower the acceptable directional error. Depending on the scenario, localization may require additional equipment (e.g. sensors) and is not always accurate enough by itself. Beam-training offers high reliability at the cost of increased overhead. Beam-tracking aims to offer the same reliability with beam-training but with reduced overhead [4, 5]. To do this a prediction is used, which requires a number of pre-acquired samples to begin. These samples are acquired through beam-training or localization. The accuracy of the prediction depends on the trajectory type of the user equipment (UE) motion in relation to the access point (AP) location. The direction of the UE motion affects the perception of the directional/angular variations relative to the tracker. For example, there are no angular variations if the UE moves towards the tracker in a straight line. The speed and AP-UE distance are closely related as the closer the UE is to the AP the faster it appears to move for this AP and vice versa. The trajectory of the UE directly affects the accuracy of the prediction as any abrupt changes in direction can make the prediction fail. Furthermore, the sampling rate of the trajectory by the tracker affects the mapping resolution of the trajectory, which in turn affects the prediction accuracy. For example, with low sampling rate the mapped trajectory resembles random points in the area. With high sampling rate, the mapped trajectory begins to resemble the actual trajectory. With very high sampling rate and assuming there are no estimation errors, the mapped and actual trajectories are identical. However, with increased sampling rate comes an increased tracking overhead that decreases the data transmission time. Finally, blockage can disrupt communication [6, 7], and therefore the beam-tracking process and make the prediction fail as it creates blanks in the mapped trajectory.

In this work, the parameters that affect the beam-tracking reliability are identified, including the obvious ones such as the UE motion, the area geometry, tracking orientation frequency and some less obvious such as the antenna orientation of the UE. Furthermore, a simple beam-tracking algorithm is introduced, along with some simple solutions that counteract the randomness of the angular variations of the user's motion. Finally, the performance of the presented solutions is presented and compared through simulations.

2. System Model

In this work a typical THz massive multiple-input multiple-output (MIMO) system is considered, with one tracker and one TO as shown in Figure 1. The tracker is equipped with a directional

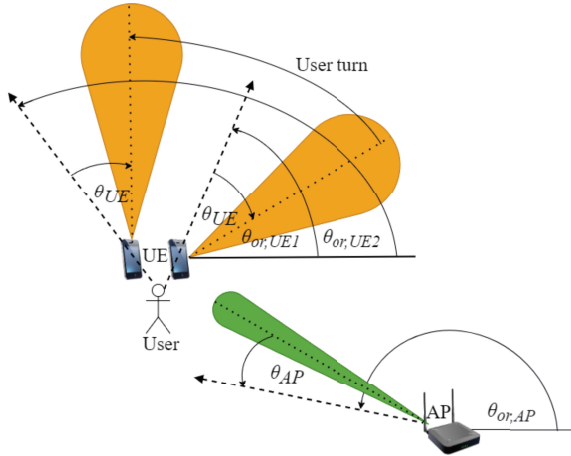
Department of Digital Systems, University of Piraeus, Piraeus 18534, Athens, Greece

E-mail: giostrat@unipi.gr; sdroulias@unipi.gr; alexiou@unipi.gr

* Corresponding Author

Manuscript received 06 August 2023, accepted 02 January 2024, and ready for publication 21 March 2024.

© 2024 River Publishers


Figure 1.

System model of an indoor scenario with 1 AP, and 1 UE. The dashed lines depict the orientation of the AP and UE antennas. The straight solid lines depict the horizontal axis from which the orientation angles $\theta_{or,AP}$ and $\theta_{or,UE}$ are estimated. The angles θ_{AP} and θ_{UE} are the angles that the beam-tracking process attempts to estimate. As the user turns the antenna orientation of the UE changes from $\theta_{or,UE1}$ to $\theta_{or,UE2}$. Therefore the direction of the beam changes but the angle θ_{UE} stays the same.

antenna of N antenna elements, and the TO with an omnidirectional antenna. Assuming that there is no blockage, the baseband equivalent received signal vector for the TO can be obtained as

$$\mathbf{y} = \mathbf{h}^H \mathbf{W}_k^H \mathbf{P} \mathbf{s} + \mathbf{n}, \quad (1)$$

where $\mathbf{h} \in N \times 1$, $\mathbf{W} \in N \times N$ and $\mathbf{P} \in N \times 1$ stand for the MIMO channel vector, the codebook matrix and the digital precoding vector. Moreover, \mathbf{s} denotes the transmitted signal vector with normalized power, and $\mathbf{n} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I})$ indicates the additive Gaussian noise (AWGN) vector. In this work, the Saleh-Valenzuela channel model is adopted [8]. In THz frequencies scattering induces more than 20 dB attenuation in the nLoS components [9], and therefore only the LoS component is taken into account, which is multiplicative to the array steering vector

$$\mathbf{a}(\psi) = \frac{1}{\sqrt{N}} [e^{-j2\pi\psi m}], \quad (2)$$

where $m \in [-\frac{N-1}{2}, \frac{N-1}{2}]$. The spatial direction can be calculated as $\psi \triangleq \frac{d}{\lambda} \sin(\theta)$, where θ is the actual direction of the TO, d is the spacing between the antenna elements, and λ is the wavelength. It is assumed that $d = \lambda/2$.

3. Frame Structure

A widely used generic frame structure for THz communications consists of multiple timeslots, the first of which are occupied by the beam-tracking and channel estimation procedures, followed by

the data transmission [10]. The beam-tracking timeslot includes the entire beam-tracking process, such as coarse beam-tracking and angle estimation phases [11], or beam-tracking and localization [12]. The frame begins with the beam-tracking procedure to find the direction of the TO, which is followed by the channel estimation in the direction estimated by the tracking process and then the data transmission. The duration of beam-tracking must be shorter than the duration of the data transmission for the beam-tracking efficiency to be high.

4. Reliability Parameters

The parameters that affect the reliability of beam-tracking can be divided into two main categories, the ones that affect the trajectory mapping resolution, and the ones that affect the angular resolution as shown in Figure 2. Specifically, the trajectory mapping resolution refers to the successful tracking of the trajectory (either location or angular) of the UE motion, which is affected by the prediction of the location/direction of the TO by the tracker, and the angular scanning area. The accuracy of the prediction is affected by the angular variation of the UE motion, the prediction algorithm and the tracking estimation frequency. The angular variation expresses how fast the direction of the TO changes in relation to the position of the tracker and the tracking estimation frequency, expresses how often the tracking process is performed. The angular variation is the result of the UE motion (speed, direction, and trajectory type), in relation to the location of the AP, i.e. the point of view (POV) of the AP (see Figure 3) if the AP is tracking the UE, and the direction of the UE motion. In the case of the UE tracking the AP, the point of view of the UE is affected greatly by the antenna orientation of the UE which changes along with the direction of the UE motion. Furthermore, the UE motion can be limited by the environment and in particular the area geometry. For example, the UE motion is more limited in a narrow corridor than in an open area such as a mall. The limited motion in a corridor, along with the appropriate placement of the AP translates to small angular variation and as a result, small prediction errors, which are easy to compensate for. The trajectory type refers to UE moving in a linear way, or a more complex way with fast direction changes (or trajectory variation) as shown in Figure 4(a). As expected, the trajectory type of the UE motion affects the angular variations in a direct way as shown in Figure 4(b) and (c), where the angular variation from the point of view of a static and a mobile tracker is presented. If the UE uses a location prediction for the location of the AP, the changes in the antenna orientation of the UE must be known to the UE. The tracking estimation frequency affects the continuity of the trajectory that is perceived by the tracker and the detected angular variations. In other words, it is the sampling rate of the trajectory/direction. The angular scanning area refers to area that is searched by the tracker and is the result of the antenna beamwidth of the tracker and the number of pilots that are sent to find the TO. Angular resolution refers to the number of directions that are successfully scanned and it is affected by the angular scanning area, and blockage. Blockage occurs when an object or human is between the tracker and the TO, and can obstruct the scanning of specific directions. The motion of the blocking object/human is independent of the AP and the UE, but like them is restricted by the area geometry.

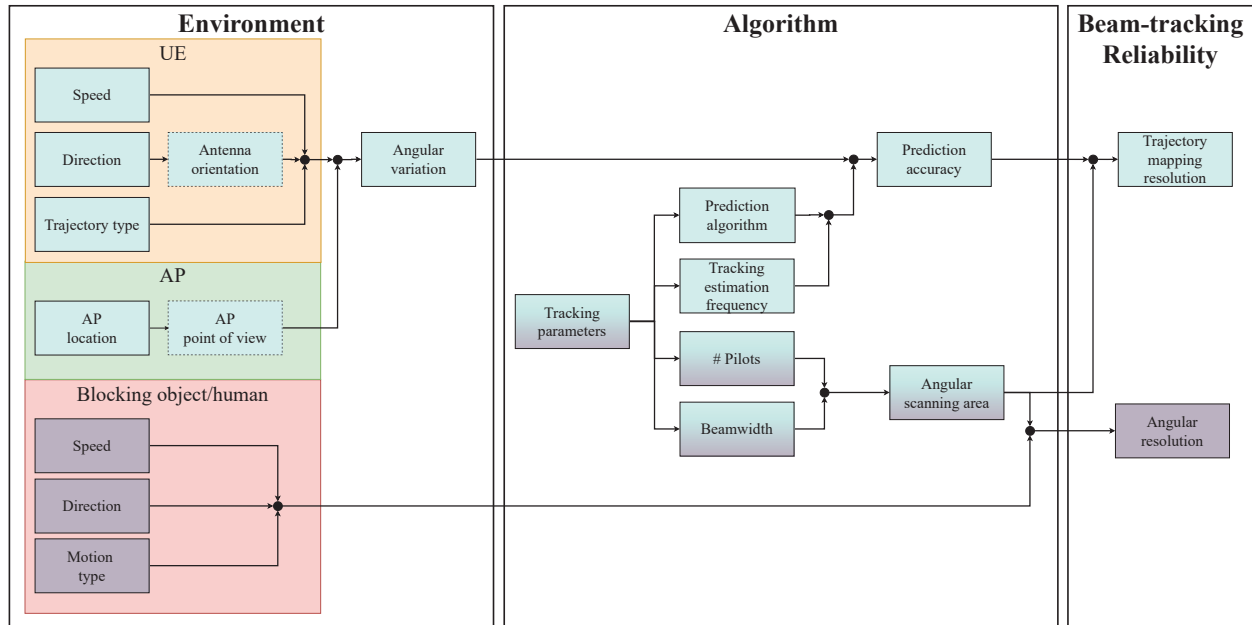


Figure 2.

Tracking reliability parameters. The parameters are divided to two main categories, the ones that affect the trajectory tracking resolution and the ones that affect the angular resolution. The dashed boxes, “Antenna orientation”, and “AP point of view” are only relevant when the AP is tracking the UE and when the UE is tracking the AP, respectively. The symbol # in Pilots denotes the number of pilots.

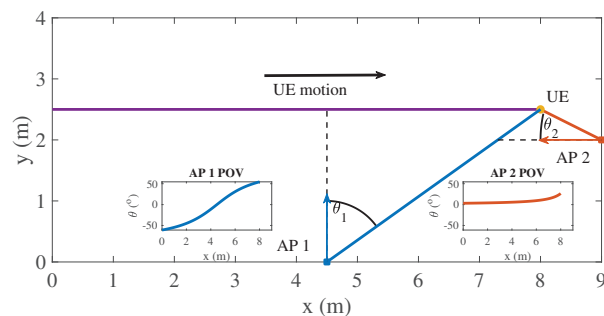


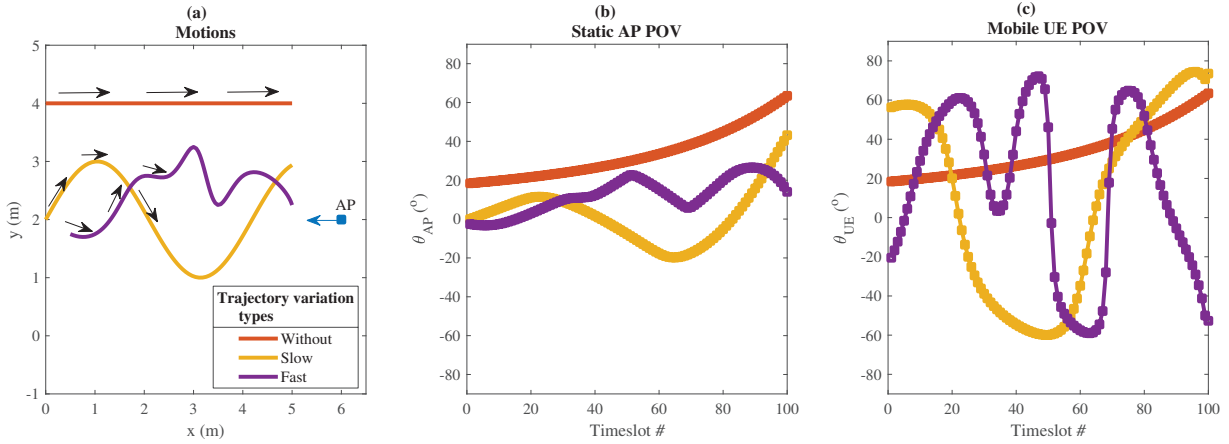
Figure 3.

AP/Tracker location point of view with regards to the UE motion. The colored arrows denote the antenna orientation of each AP and the black colored arrow denotes the general direction of the UE motion. In the insets, the angular variation of the UE motions from the different point of views of AP 1 and 2 is shown.

5. Challenges

There are many challenges in tracking a mobile UE, the most obvious of which is the motion of the UE, which can be broken down to speed, direction and type of trajectory (linear, irregular, etc). In Figure 4(a), three motions are presented with different trajectory variation types (without, slow and fast), and their angular variation from the point of view of the AP and the UE is shown in Figure 4(b) and (c). It can be observed that different trajectories have different angular variation in relation to the a single AP. For ease of reference, the trajectory without variation

will be named “Trajectory 1”, the trajectory with slow variation “Trajectory 2” and the trajectory with fast variation “Trajectory 3”. Trajectory 1 is linear and the angular variation it causes in relation to the AP location is slow. Trajectory 2 is sinusoidal and the angular variation it causes is faster. Trajectory 3 is a random trajectory and causes slow angular variation at the start and faster towards the end. Furthermore, it has more changes in direction than the Trajectory 2. Both Trajectory 2 and 3 are harder to follow than Trajectory 1. The main parameter that affects the performance of beam-tracking is the beamwidth. Wide beams cover a relatively large area and therefore make tracking easier, but with low accuracy and antenna gain. On the other hand, narrow beams cover a smaller area making tracking harder but with increased accuracy and antenna gain. In the case where the UE tracks the AP, the direction of the UE motion is especially important as in most cases it affects the antenna orientation of the UE, and any changes to it significantly affect the angular variation, as shown in Figure 4(c), which can impair the beam-tracking performance. If a direction prediction is used, the antenna orientation changes are included in the prediction. However, if the UE uses a location-based prediction, then the antenna orientation of the UE must be known for the UE to form a beam in the direction of the AP. For example, the UE antenna forms a beam at 10° right of the broadside to the direction of the AP. The user then turns 10° left and the direction of the UE beam stops pointing to the direction of the AP and the link is broken (see for example Figure 1). In order to restore the link, the UE antenna must form a beam 20° to the right. This can happen even with static user. There are methods that allow the UE to know the direction changes and in turn the antenna orientation changes, such as using a gyroscope, [13], or the camera of a smartphone [14]. Using the camera however, raises privacy


Figure 4.

Trajectory variation types of a UE motion, with a static AP. (a) Variation types, (b) angular variation from the point of view of a static tracker (AP), and (c) angular variation from the point of view of a mobile tracker (UE). The antenna orientation of the AP is depicted with the blue arrow, and the antenna orientation of the UE changes with the motion and is depicted with the black arrows for each motion. It is assumed that there is no blockage. The markers in panels (b) and (c) show how sudden the changes in direction are.

concerns. Another factor that greatly affects beam-tracking is blockage and particularly human blockage, since objects are mostly static. Blockage can significantly impede beam-tracking, especially when combined with highly directional beams as the probability of a link being blocked completely (total blockage) instead of partially increases with narrower beams. Moreover, user-blockage makes the area behind the user non-scannable. One method to cope with blocked links is to scan around the presumed blocker and wait for the blockage to end [10], assuming the duration is short.

The use of multiple APs, relays and RISs is the most obvious solution to avoid blocked links [15]. Self-healing is another promising solution to blockage as it allows for the beam to regenerate behind the obstacle [16]. This way the required number of APs, relays and RISs can be reduced significantly and in the best case scenario only 1 AP (per room if indoors) will be required.

6. Beam-Tracking Algorithm

In this section a simple beam-tracking algorithm is presented for the purpose of explaining the required steps in a beam-tracking process. It is assumed that the beam-tracking is performed by the AP. In general, a beam-tracking algorithm consists of an initialization phase, in which the necessary input is obtained (e.g. some direction/location estimations) and the beam-tracking phase that consists of the prediction and the tracking in a small amount of directions around the predicted direction. Other phases, such as a correction phase can be added if necessary. In general, the initialization phase includes an exhaustive search. In this paper, the initialization phase lasts for 2 timeslots, where the tracker estimates the direction of the UE with exhaustive search using the hierarchical codebook. Specifically, binary search is used and reduces total number of pilots required for the exhaustive search to $2K$. Note that the 2 codewords of the first codebook level can be skipped in favor of using 4 codewords in the second level, for the increased antenna gain without increasing the pilot overhead.

After the initialization phase, the tracker starts predicting the next direction of the UE relative the position of the tracker. The prediction assumes that the angular variation of the UE motion relative to the position of the tracker in two timeslots is consistent (i.e. the trajectory is circular). The simplicity of the prediction allows for the accurate assessment of other methods in increasing the tracking reliability.

7. Beam-Tracking Reliability

In this section some simulation results with 3 reliability increase methods are presented. Unless otherwise stated, the central frequency is 150 GHz, and the number of antenna elements of the AP is $N = 256$. It is assumed that the number of active antenna elements follows 2^k , where $k = 1, 2, \dots, K$ is the codebook level used and $K = \log_2(N)$ is the last codebook level, in this case $K = 8$. Furthermore, the codebook in [17] is used. The antenna element spacing is $\lambda/2$, the initialization phase lasts 2 timeslots as it was mentioned in Section 6, and the number of pilots used in the beam-tracking process is 8. The number of pilots required by the initialization phase is 16 and serves as a benchmark. The number of pilots required by the beam-tracking process to fully track a trajectory can be estimated as $N_p = N_p(T_s - T_i)$, where N_p is the number of pilots per timeslot, T_s is the number of timeslots in which the beam-tracking process is performed and T_i is the number of timeslots in which the initialization is performed, which in this work is 2. The antenna of the UE is assumed to be omnidirectional. The beam-tracking process does not take corrective measures are taken when it fails to find the UE, in order to make the performance assessment easier. In this work, all 3 trajectories are divided in 10 parts, the start of each one representing a sample taken with the beam-tracking process in 1 timeslot, for a total of 10 samples and timeslots of beam-tracking. The estimation frequency, Q_I , is a multiplicative factor in the number of samples. For example, when $Q_I = 1$, 10 samples are taken from the trajectory in 10 timeslots, when $Q_I = 2$, 20 samples in 20 timeslots, etc.

7.1. Reliability Increase Methods

In Figure 5, an example of beam-tracking from the point of view of the AP for the three trajectories shown in Figure 4, namely Trajectory 1, Trajectory 2 and Trajectory 3 is presented. The tracking estimation frequency, $Q_I = 2$, the maximum number of active antenna elements is 256 and the number of pilots (i.e. the number of directions that can be scanned) used in the beam-tracking process is $N_p = 8$. Trajectory 1 can be fully tracked due to the slow angular variation, however the beam-tracking process in Trajectories 2 and 3 fails at the start due to the fast angular variation. In order to track the Trajectories 2 and 3 entirely, the accuracy of the prediction algorithm must increase. To do this without changing the prediction algorithm, the angular variation between two consecutive timeslots must become slower and the scanning area of the tracker must be increased. The angular variation can become slower by increasing the tracking estimation frequency, Q_I . By increasing Q_I , the distance travelled by the UE between two consecutive timeslots decreases and so does the angular variation that is perceived by the AP. To increase the scanning area of the tracker, the number of directions to scan (and therefore the number of pilots sent by the AP, N_p) must be increased, or the beamwidth of the AP antenna must be increased by reducing the number of active antenna elements, N_a . Although the above-mentioned methods have the potential to fully track the UE trajectories, the requirements to do so are different for each method. Furthermore, higher requirements mean higher overhead. The trade-off between the beam-tracking reliability and the overhead that is required to achieve it is the focus of most works on beam-tracking. In Figure 6, the last timeslot in which the beam-tracking process successfully found the UE is presented as a function of (a) Q_I , (b) the number of active antenna elements, and

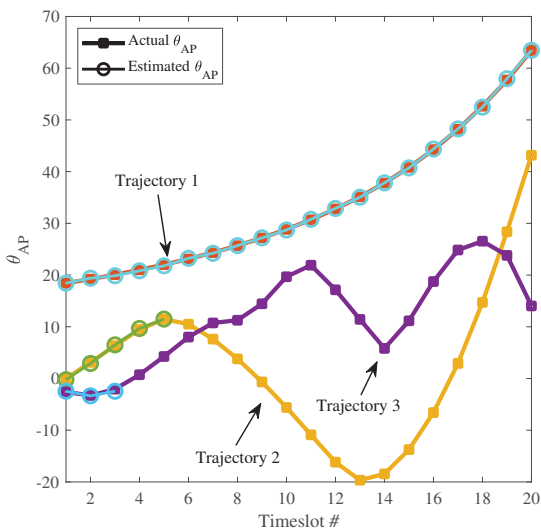


Figure 5. Beam-tracking example from the point of view of the AP for the trajectories shown in Figure 4. The beam-tracking process is successful in Trajectory 1 due to the slow angular variation, but fails in Trajectories 2 and 3 due to the fast angular variation.

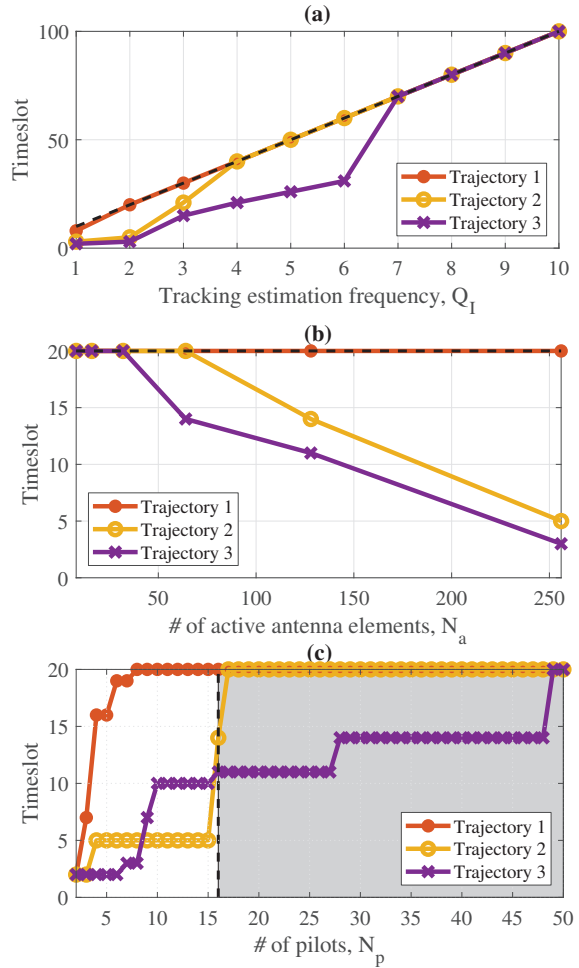


Figure 6. Performance of the 3 reliability increase methods. (a) Tracking estimation frequency, (b) number of active antenna elements, and (c) number of pilots vs the last timeslot in which the beam-tracking process succeeded. In (a) and (b), the dashed line depicts the maximum number of timeslots for the beam-tracking process. In (c), the dashed line at 16 pilots marks the number of pilots used in the initialization and acts as a benchmark.

(c), the number of pilots used in tracking, for all 3 trajectories shown in Figure 4. Unless otherwise stated, the tracking estimation frequency $Q_I = 2$, the number of pilots $N_p = 8$ and the number of active antenna elements is $N_a = 256$. It is observed that increasing the tracking estimation frequency in (a), increases the number of timeslots in which the UE is successfully tracked. As expected, the tracking of the entire Trajectory 1 is successful from $Q_I = 2$ and for $Q_I = 1$, only a small part of the trajectory is not tracked. However, as the angular variation speed of the trajectory becomes faster, tracking the entire trajectory becomes more difficult. For Trajectory 2 and 3, the tracking of the entire trajectory is achieved with $Q_I = 5$ and 7, respectively. Further increase of Q_I increases the pilot overhead without increasing the performance. For a specific trajectory, the increase is not linear due to the changes in the angular variation from the point of

view of the AP. For example, the angular variation of Trajectory 2 becomes faster as the UE moves towards the AP. In (b), increasing the number of active elements while using the same number of pilots decreases the scanning area due to the narrower beams and can decrease the probability of tracking the UE. Again, Trajectory 1 is always entirely tracked, while for Trajectory 2 and Trajectory 3, increasing the number of active antenna elements reduces the number of timeslots that can be tracked. Trajectory 1 can be tracked with all 256 antenna elements, while Trajectories 2 and 3 can be tracked with 64 and 32 antenna elements respectively. In (c), increasing the number of pilots increases the number of timeslots of successfully tracking the UE due to the larger scanning area. The increase is not linear as the prediction is not accurate enough and a larger scanning area is required. For example, Trajectory 1 requires 8 pilots, Trajectory 2 requires 17 to be entirely tracked and Trajectory 3 requires 49 pilots. As the angular variation of Trajectory 2 at the start is slower than the angular variation of Trajectory 3, it can be tracked with fewer pilots. However, as both trajectories progress, the speed of their angular variation changes. In the next part of both trajectories, the angular variation of Trajectory 3 is slower than that of Trajectory 2 and therefore Trajectory 3 can be tracked up to this point with fewer pilots than Trajectory 2. Overall, the angular variation of Trajectory 2 is slower than that of Trajectory 3 and therefore can be tracked with fewer pilots. It should be noted that increasing the number of pilots to 16 and higher is inefficient as the initialization requires 16 pilots and is more probable to find the UE due to using the binary search method, which is a version of the exhaustive search. It is observed that all solutions can help the AP track the 3 trajectories entirely, but with different cost, e.g. pilot overhead. For example, the total number of pilots, excluding the initialization phase, without using a reliability increase method is $N_p = 8 * (20 - 2) = 144$, with increasing Q_I to 7 in (a) it is $N_p = 8 * (70 - 2) = 544$, with reducing N_a to 32 in (b) it is $N_p = 8 * (20 - 2) = 144$, and with increasing N_p to 49 in (c) it is $N_p = 49 * (20 - 2) = 882$. In this case, the increased performance relative to the pilot overhead of reducing N_a seems to be the best among the presented solutions but the reduced gain that follows can reduce the viability of the solution in cases where high antenna gain or high accuracy is required. Increasing the number of pilots seems to be the worst solution as the pilot overhead required for the AP to track all 3 trajectories is the highest. Increasing Q_I is the middle ground between increasing N_p and decreasing N_a in the beam-tracking process. It requires significantly fewer pilots than increasing N_p to fully track all 3 UE trajectories and without decreasing the antenna gain. Although, all solutions can fully track the 3 UE trajectories, the optimal solution depends on the scenario. If high gain or high accuracy is required, then increasing the tracking estimation frequency is the best solution among the three. On the other hand, if the antenna gain or accuracy requirements are not high, then increasing the beamwidth by decreasing the number of active antenna elements is the best solution between the three. Increasing the number of pilots is preferred only if it is lower than the number of pilots required by the initialization. Note that a more accurate prediction will significantly increase the performance of beam-tracking with low overhead, especially if used in conjunction with one of the above-mentioned methods.

7.2. Hierarchical Codebook in Beam-Tracking

One simple way of decreasing the number of pilots without decreasing the scanning area is by making use of the codebook. Specifically, with the use of multiple codebook levels instead of only the last one, the pilot overhead can be reduced substantially by taking advantage of the binary tree structure of the codebook and the binary search method. For example, instead of using 8 pilots in the last codebook level, the same scanning area can be achieved with using 4 pilots in the $K - 1$ codebook level and 2 pilots in the last level for a total of 6 pilots, which is a 25% decrease.

8. Discussion & Conclusions

This paper provides an overview on beam-tracking methods for finding the direction of the user. The parameters that affect the reliability of beam-tracking algorithms were identified, and divided into two main categories, the trajectory mapping resolution and the angular resolution. Trajectory mapping resolution was defined as the successful tracking of the UE over time and the angular resolution the number of directions that can be successfully scanned each time. In contrast to localization, beam-tracking aims to offer greater accuracy and with lower overhead than beam-training. However, the accuracy of beam-tracking is related to the accuracy of the prediction and the scanning area. The accuracy of a prediction algorithm can be increased by increasing the tracking estimation frequency. The scanning area can be increased by decreasing the number of active antenna elements to form wider beams or by increasing the number of pilots sent by the tracker. These 3 methods were presented and compared with each other in terms of performance, overhead and requirements, to find how much they affect tracking and the cost that comes with them. It was shown that greatly increasing the scanning area by increasing the number of pilots to high levels is not a viable solution due to the high pilot overhead, although finding the ideal number of pilots for the UE trajectory is vital. Furthermore, increasing the beamwidth by decreasing the number of active antenna elements increases the tracking reliability, but also decreases the antenna gain which can be undesirable at mmWave and THz frequencies. Increasing the tracking estimation frequency was observed to be a good solution as it increases the accuracy of the prediction and therefore the tracking reliability with relatively low increase of the pilot overhead. The tracking reliability can be significantly improved with a more accurate prediction algorithm. Furthermore, the addition of more APs, relays and RISs can increase the tracking reliability in cases of blockage, if at least 1 AP-UE link is LoS. The cost however (e.g. total pilot overhead) can increase significantly.

Acknowledgement

This work was supported by the European Commission's Horizon Europe Programme under the Smart Networks and Services Joint Undertaking (TERA6G project Grant Agreement 101096949).

References

- [1] S. Priebe, M. Jacob, and T. Kürner, "The impact of antenna directivities on THz indoor channel characteristics," in *2012 6th European Conference on Antennas and Propagation (EUCAP)*, pp. 478–482, 2012.
- [2] E. N. Papatiriu, A.-A. A. Boulogeorgos, and A. Alexiou, "Performance Analysis of THz Wireless Systems in the Presence of Antenna Misalignment and Phase Noise," *IEEE Commun. Lett.*, pp. 1–1, 2020.
- [3] A.-A. A. Boulogeorgos, E. N. Papatiriu, and A. Alexiou, "Analytical Performance Assessment of THz Wireless Systems," *IEEE Access*, vol. 7, no. 1, pp. 1–18, Jan. 2019.
- [4] Y. Wang, Z. Wei, and Z. Feng, "Beam Training and Tracking in MmWave Communication: A Survey," 2022.
- [5] B. Ning, Z. Chen, Z. Tian, C. Han, and S. Li, "A Unified 3D Beam Training and Tracking Procedure for Terahertz Communication," *IEEE Transactions on Wireless Communications*, vol. 21, no. 4, pp. 2445–2461, 2022.
- [6] G. R. MacCartney, S. Deng, S. Sun, and T. S. Rappaport, "Millimeter-Wave Human Blockage at 73 GHz with a Simple Double Knife-Edge Diffraction Model and Extension for Directional Antennas," in *IEEE 84th Vehicular Technology Conference (VTC-Fall)*, IEEE, Sep 2016.
- [7] M. Jacob, S. Priebe, R. Dickhoff, T. Kleine-Ostmann, T. Schrader, and T. Kurner, "Diffraction in mm and Sub-mm Wave Indoor Propagation Channels," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 3, pp. 833–844, Mar 2012.
- [8] A. Saleh and R. Valenzuela, "A Statistical Model for Indoor Multipath Propagation," *IEEE Journal on Selected Areas in Communications*, vol. 5, no. 2, pp. 128–137, 1987.
- [9] E. N. Papatiriu, J. Kokkonen, A.-A. A. Boulogeorgos, J. Lehtomaki, A. Alexiou, and M. Juntti, "A new look to 275 to 400 GHz band: Channel model and performance evaluation," in *IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, IEEE, Sep. 2018.
- [10] J. Tan and L. Dai, "Wideband Beam Tracking in THz Massive MIMO Systems," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 6, pp. 1693–1710, 2021.
- [11] S. Kim, G. Kwon, and H. Park, "High-resolution multi-beam tracking with low overhead for mmWave beamforming system," *ICT Express*, vol. 7, no. 1, pp. 28–35, 2021.
- [12] G. Stratidakis, A.-A. A. Boulogeorgos, and A. Alexiou, "A cooperative localization-aided tracking algorithm for THz wireless systems," in *IEEE Wireless Communications and Networking Conference (WCNC)*, Marrakech, Morocco, Apr. 2019.
- [13] S. Yean, B. S. Lee, C. K. Yeo, C. H. Vun, and H. L. Oh, "Smartphone Orientation Estimation Algorithm Combining Kalman Filter With Gradient Descent," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 5, pp. 1421–1433, 2018.
- [14] W. Elloumi, K. Guissous, A. Chetouani, R. Canals, R. Leconte, B. Emile, and S. Treuillet, "Indoor navigation assistance with a Smartphone camera based on vanishing points," in *International Conference on Indoor Positioning and Indoor Navigation*, pp. 1–9, 2013.
- [15] G. Stratidakis, S. Droulias, and A. Alexiou, "A beam-tracking framework for THz networks," *Frontiers in Communications and Networks*, vol. 3, 2022.
- [16] P. Zhou, X. Fang, Y. Fang, R. He, Y. Long, and G. Huang, "Beam Management and Self-Healing for mmWave UAV Mesh Networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 2, pp. 1718–1732, 2019.
- [17] Z. Xiao, T. He, P. Xia, and X. Xia, "Hierarchical Codebook Design for Beamforming Training in Millimeter-Wave Communication," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3380–3392, May 2016.

Biographies



Giorgos Stratidakis was born in Athens, Greece, in 1990. He received the bachelor's degree in telecommunications engineering from the Department of Telecommunications Science and Technology, University of Peloponnese, in 2016, and the master's degree in digital communications and networks from the Department of Digital Systems, University of Piraeus, in 2018. He is currently pursuing the Ph.D. degree in wireless communications. In 2017, he joined the Department of Digital Systems, University of Piraeus, where he conducts research in the area of wireless communications.



Sotiris Droulias received the diploma in electrical and computer engineering and the Ph.D. degree in nonlinear photonics from the National Technical University of Athens, Athens, Greece, in 2001 and 2007, respectively. He is currently a Research Associate at the Department of Digital Systems, ICT School, University of Piraeus, Piraeus, Greece. During the period 2012–2020 he was a member of the Photonic- Phonic- and Meta-materials group at FORTH-IESL, Crete, Greece, and during the period 2009–2012 we worked as an Adjunct Lecturer at the University of Patras. He has worked on several EC funded projects and his research interests include metamaterials, photonic crystals, metasurfaces, nanolasers, plasmonics, active media. He is an author of more than 30 papers and 2 book chapters, he has received more than 15 talk invitations in prestigious conferences and he serves as a reviewer in several scientific journals. In 2019 he received the best poster award for his work on metasurface lasers in META 2019, Lisbon, Portugal, and in 2020 he was recognized as an outstanding reviewer from the Institute of Physics (IOP).



Angeliki Alexiou is a professor at the department of Digital Systems, ICT School, University of Piraeus. She received the Diploma in Electrical and Computer Engineering from the National Technical University of Athens in 1994 and the PhD in Electrical Engineering from Imperial College of Science, Technology and Medicine, University of London in 2000. Since May 2009 she has been a faculty member at the Department of Digital Systems, where she conducts research and teaches undergraduate and postgraduate courses in Broadband Communications and Advanced Wireless Technologies. Prior to this appointment she was with Bell Laboratories, Wireless Research, Lucent Technologies, (later Alcatel-Lucent, now NOKIA), in Swindon, UK, first as a member of technical staff (January 1999–February 2006) and later as a Technical Manager (March 2006–April 2009). Professor Alexiou is a co-recipient of Bell Labs President’s Gold Award in 2002 for contributions to Bell Labs Layered Space-Time (BLAST) project and the Central Bell Labs Teamwork Award in 2004 for role model teamwork and technical achievements in the IST FITNESS project. Professor Alexiou is the Chair of the Working Group on Radio Communication Technologies and of the Working Group on High Frequencies Radio Technologies of the Wireless World Research Forum. She is a member of the IEEE and the Technical Chamber of Greece. Her current research interests include radio interface for systems beyond 5G, MIMO, THz wireless technologies and Reconfigurable Intelligent Surfaces, efficient resource management for Ultra Dense wireless networks, machine-to-machine communications and Artificial Intelligence and Machine Learning for future wireless systems. She is the project coordinator of the H2020 TERRANOVA project (ict-terranova.eu) and the technical manager of H2020 ARIADNE project (ict-ariadne.eu).

Coherence Bandwidth of Rice Channels for Millimeter Wave and Sub-Terahertz Applications

Werner Mohr

Abstract: 6G research in the millimeter wave and sub-Terahertz domain is targeting very wideband systems with significantly higher data throughput than for 5G systems. Multipath propagation under shadowing conditions is affecting radio propagation, where multipath propagation results in frequency-selective fading, which is characterized by the coherence bandwidth and the time variation by the coherence time. In these frequency ranges shadowing can be overcome by additional means in the network deployment such as reflectors, RIS arrays or repeaters, which provide at the receiver a channel impulse response with a strong component (Rice type channel). Coherence bandwidth and coherence time are well-known for Rayleigh channels. However, both parameters for Rice channels versus the Rice factor K' are not available. This paper is investigating the coherence bandwidth and time for Rice channels based on an approximative approach for the fading statistics. With the proposed correlation criterion, the coherence bandwidth and time tend to infinity from a Rice factor around $K' \geq 4 \approx 6$ dB. These relations are provided by approximative functions for $K' \geq 0$ starting at the Rayleigh channel.

Keywords: Coherence bandwidth, coherence time, Rayleigh channel, rice channel.

1. Introduction

6G (sixth generation of mobile communication) is targeting high aggregated carrier throughput rates in the order of several hundred Gbit/s or up to a Terabit/s (e.g., [1]), which requires very wideband radio channels. Such wide carrier bandwidth will only be available in the millimeter wave or sub-Terahertz domain. Pathloss in these frequency ranges is very high, because additional effects of atmospheric, rain and foliage attenuation need to be considered. Radio propagation is interrupted by obstacles or shadowing ([2] to [5]). In addition, multipath propagation needs to be considered, where the amplitude statistics of the received signal can be described in cases without a strong component in the channel impulse response by a Rayleigh distribution or in cases with a strong component by a Rice distribution ([5] to [7]).

Under shadowing conditions, usually a Rayleigh channel can be expected. If by additional technical means like reflectors, RIS arrays (Reconfigurable Intelligent Surfaces) or repeaters a strong component in the channel impulse response is created to shape the

radio channel [8] in order to increase the received power compared to the Rayleigh channel, the channel can be characterized by a Rice channel.

Multipath propagation results in frequency-selective fading, where for certain frequencies in the channel transfer function deep fades may occur. The frequency-selectivity is characterized by the channel coherence bandwidth. For very wideband systems and delays in the channel impulse response, which are significantly bigger than the signal symbol duration, a huge channel equalization effort at the receiver is required. Therefore, a big coherence bandwidth is desired.

This paper is investigating the coherence bandwidth of Rice channels and especially its relation to the Rice factor K' to increase the coherence bandwidth compared to a Rayleigh channel under similar propagation conditions but without additional technical means (reflectors, RIS or repeaters).

Section 2 is describing the interference scenario and the signal model, Section 3 provides the approach to calculate the channel coherence bandwidth with the envelope correlation function and Section 4 is calculating/estimating the coherence bandwidth and coherence time of Rayleigh and Rice channels. Results are summarized in Section 5.

2. Interference Scenario and Signal Model

2.1. Channel Impulse Response for Multipath Propagation

The multipath propagation scenario between the transmitter and receiver of user k is shown in Figure 1. The different I multipath components may be caused by reflecting areas R_1 to R_I like walls, metallic reflectors, RIS arrays or repeaters. In general, there is also a direct component between transmitter and receiver or a strong component under shadowing conditions, which is enforced by metallic reflectors, RIS arrays or repeaters [8]. The channel impulse response follows in (1).

$$h_{mp}(t) = h_s \cdot \delta(t) + \sum_{i=1}^I h_{r,i} \cdot \delta(t - \tau_i) \quad (1)$$

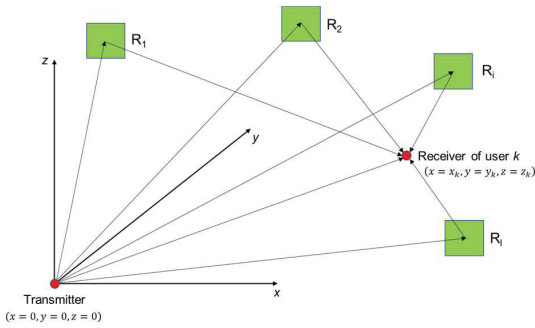
In the following it is assumed that the direct or strong component h_s occurs at the beginning of the channel impulse response at relative delay $\tau = 0$. If a RIS is applied as a reflecting area, each multipath component $h_{r,i}$ comprises $M \cdot N$ sub-components from the reflection of each RIS element, which may add a slightly different delay $\Delta\tau_{RIS,i,m,n,k}$. In the following it is assumed that all involved RIS arrays may have the same size $M \cdot N$. It is assumed that the different RIS elements are electromagnetically mutually

Technical University of Berlin, Germany

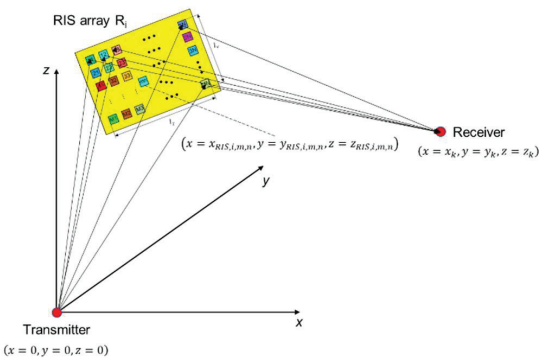
E-mail: werner.mohr@campus.tu-berlin.de

Manuscript received 28 July 2023, accepted 23 November 2023, and ready for publication 21 March 2024.

© 2024 River Publishers


Figure 1.

General multipath propagation scenario between transmitter and receiver of user k .


Figure 2.

Propagation scenario for RIS array R_i between transmitter and receiver of user k .

independent from each other for a spacing $\geq \lambda/2$ (λ wavelength) like for antenna arrays [9], pp. 254.

Figure 2 is describing the propagation scenario for a RIS array R_i . The reflection coefficient or Radar cross section $\varrho(\text{material}, \phi_i, \phi_r)_{i,m,n,k}$ of each RIS element is changing – depending on the material – the amplitude and angle ϕ_r of the reflected wave in relation to the amplitude and angle ϕ_i of the incoming wave. A RIS array shows some beamforming capabilities to direct the outgoing wave to a specific direction.

The sum of all contributions $b_{tr,i,k}(t)$ via a RIS array R_i is given in (2) with the contributions $b_{tR,i,m,n}$ and delay $\tau_{tR,i,m,n}$ from the transmitter to RIS element i, m, n and $b_{Rr,i,m,n,k}$ and delay $\tau_{Rr,i,m,n,k}$ from the RIS element i, m, n to the receiver of user k and a RIS element internal delay $\Delta\tau_{i,m,n,k}$.

$$b_{tr,i,k}(t) = \sum_{m=1}^M \sum_{n=1}^N \{ b_{tR,i,m,n} \cdot \varrho(\text{material}, \phi_i, \phi_r)_{i,m,n,k} \cdot b_{Rr,i,m,n,k} \cdot \delta(t - \tau_{tR,i,m,n} - \Delta\tau_{i,m,n,k} - \tau_{Rr,i,m,n,k}) \} \approx b_{r,i} \cdot \delta(t - \tau_i) \quad (2)$$

Due to the rather small physical RIS size compared to the distance between transmitter and receiver the delay differences between the different RIS elements do not result in a significant

frequency dependency. The RIS elements can be used for beam steering towards a specific direction.

However, the entire impulse response $b_{tr,k}(t)$ between transmitter and receiver of user k follows as the sum of all I multipath components plus the direct component in (3).

$$b_{tr,k}(t) = b_{s,k} \cdot \delta(t) + \sum_{i=1}^I \sum_{m=1}^M \sum_{n=1}^N \{ b_{tR,i,m,n} \cdot \varrho(\text{material}, \phi_i, \phi_r)_{i,m,n,k} \cdot b_{Rr,i,m,n,k} \cdot \delta(t - \tau_{tR,i,m,n} - \Delta\tau_{i,m,n,k} - \tau_{Rr,i,m,n,k}) \} \approx b_s \cdot \delta(t) + \sum_{i=1}^I b_{r,i} \cdot \delta(t - \tau_i) \quad (3)$$

In general, the bigger delay differences between the different I multipath components remain, which are responsible for the frequency dependency of the RIS transmission [10] which is sensitive to the user location.

If beam steering in the RIS system by appropriate outgoing angles of reflected waves is applied, a strong multipath component towards user k may be created by avoiding additional strong components, which is reducing the frequency dependency. However, these reflected waves in other directions may cause interference or strong frequency dependency for other users.

2.2. Signal Model to Derive the Coherence Bandwidth of the Rice Channel

The applied signal model for the Rice channel is based on [11], pp. 45. In the following the channel impulse response with a dominant component is described by (1). If the number of arriving partial waves $I \rightarrow \infty$, (1) is replaced by (4).

$$b_{Rice,mp}(t) = b_s \cdot \delta(t) + b(t) \quad (4)$$

$b_s \cdot \delta(t)$ corresponds to the dominant component and $b(t)$ describes the scattering part from many different sources.

With the notation in [11], pp. 13 the transmitted signal is described in (5) with the E-field component in z-direction.

$$s_t(t) = E_{z,0} \cdot \cos(\omega_0 \cdot t) \quad (5)$$

The received signal $s_r(t)$ is affected by Doppler shift $\omega_{d,j}$ versus different angles of arrival ϑ_j and wavelength λ or frequency f , speed of light c_0 and mobile speed v [11], p. 20.

$$\beta = \frac{2\pi}{\lambda} = \frac{2\pi \cdot f}{c_0} = \frac{\omega}{c_0} \quad \omega_{d,j} = \frac{2\pi}{\lambda} \cdot v \cdot \cos \vartheta_j \quad (6)$$

At each Doppler shift $\omega_{d,j}$ corresponding to $\pm\vartheta_j$ signals with different delay $\tau_{j,i}$ may occur or at each delay $\tau_{j,i}$ partial waves with different Doppler shift $\omega_{d,j}$ are received.

With these considerations the received signal is described as the sum of all different wave components versus Doppler shift $\omega_{d,0}$ of the dominant component and versus delay $\tau_{j,i}$ and Doppler shift

$\omega_{d,j}$ for the scattering part with the amplitude coefficients $C_{0,0}$ and $C_{j,i}$ with I and $J \rightarrow \infty$.

$$E_z(\omega, t) = E_{z,0} \cdot C_{0,0} \cdot \cos(\omega \cdot t + \omega_{d,0} \cdot t) + E_{z,0} \cdot \sum_{j=1}^J \sum_{i=1}^I C_{j,i} \cdot \cos(\omega \cdot t + \omega_{d,j} \cdot t - \omega \cdot \tau_{j,i}) \quad (7)$$

The amplitude coefficients are defined via the square root of the received power normalized to the transmitted power, which corresponds to the pathloss. With the receiver antenna azimuth diagram $G(\vartheta)$ and the received power $P(\vartheta_0, \tau_s)$ of the dominant component and the fraction of received power $p(\vartheta_j, \tau_{j,i})$ within $d\vartheta d\tau$ for the scattered part the coefficients $C_{0,0}$ and $C_{j,i}$ can also be interpreted as the amplitudes of the channel impulse response of the partial waves (8). The sum of all partial waves with the same delay but different angle of arrival or Doppler shift is given in (9), which is related to (1) and (4).

$$\frac{C_{0,0}}{\sqrt{G(\vartheta_0)}} = \sqrt{P(\vartheta_0, \tau_s)} = |b_s|$$

$$\frac{C_{j,i}}{\sqrt{G(\vartheta_j)}} = \sqrt{p(\vartheta_j, \tau_{j,i}) d\vartheta d\tau} = |b_{j,i}| \quad (8)$$

$$b_{r,i}(\tau_i) = \sum_{j=1}^J b_{j,i}(\tau_i) \quad (9)$$

For simplification it is assumed that the phase angles in (7) are related to the phase angle of the dominant component. Then $E_z(\omega, t)$ in (7) is replaced by $E'_z(\omega, t)$ in (10) and can be represented with [12], p. 156 by in-phase $x_1(t)$ and quadrature components $x_2(t)$ in (11).

$$E'_z(\omega, t) = E_{z,0} \cdot C_{0,0} \cdot \cos(\omega \cdot t) + E_{z,0} \cdot \sum_{j=1}^J \sum_{i=1}^I C_{j,i} \cdot \cos(\omega \cdot t + (\omega_{d,j} - \omega_{d,0}) \cdot t - \omega \cdot \tau_{j,i}) \quad (10)$$

$$E'_z(\omega, t) = E_{z,0} \cdot C_{0,0} \cdot \cos(\omega \cdot t) + E_{z,0} \cdot \sum_{j=1}^J \sum_{i=1}^I C_{j,i} \cdot \{ \cos(\omega \cdot t) \cdot \cos((\omega_{d,j} - \omega_{d,0}) \cdot t - \omega \cdot \tau_{j,i}) - \sin(\omega \cdot t) \cdot \sin((\omega_{d,j} - \omega_{d,0}) \cdot t - \omega \cdot \tau_{j,i}) \}$$

$$= x_1(t) \cdot \cos(\omega \cdot t) - x_2(t) \cdot \sin(\omega \cdot t) \quad (11)$$

With the assumptions in [11], pp. 47 on the stationarity of the parameters and $E'_z(\omega, t)$ corresponds to a random process, where the phase angles $\psi_{j,i}$ of sample functions of the individual waves are uniformly distributed random independent variables with $0 \leq \psi_{j,i} \leq 2\pi$, $E'_z(\omega, t)$ is regarded as a wide-sense stationary process with respect to ensemble averages. However, it is not stationary with respect to time averages, because over time and distance time delays and angles of arrivals are slowly changing. Therefore, the process is not strictly ergodic. With an increasing number of waves – which is assumed here – the difference between ensemble

and time averages decreases. Therefore, in the following ensemble averages are used for the calculation.

The definition of a specific model for the channel impulse response and propagation conditions follows the same assumptions as in [11], pp. 46. For simplification it is assumed that at the receiver location partial waves arrive from all directions with a uniform distribution.

$$p(\vartheta) = \frac{1}{2\pi} \quad (12)$$

From practical experience the shape of the density of the square of the channel impulse response versus delay τ (power delay profile) based on many scatterers without a dominant component corresponds to an exponential function and is interpreted as a probability distribution function $p(\tau)$ of the received power versus delay normalized to the transmitted power.

$$p(\tau) = \frac{1}{s} \cdot e^{-\tau/s} \quad \text{with} \quad \int_0^\infty \frac{1}{s} \cdot e^{-\tau/s} d\tau = 1 \quad (13)$$

The parameter s in (13) directly corresponds to the delay spread [4], pp. 160 of the scattering part of the channel impulse response without the direct component. Both probability distribution functions are assumed to be statistically independent. Therefore, it follows [11], p. 50:

$$p_s(\vartheta, \tau) = \frac{1}{2\pi} \cdot \frac{1}{s} \cdot e^{-\tau/s} \quad (14)$$

The direct component in the Rice channel has the probability distribution in (15).

$$p_s(\vartheta, \tau) = \delta(\vartheta - \vartheta_0) \cdot \delta(\tau) \quad (15)$$

With (8) this is related to the channel impulse response.

$$P(\vartheta_0, \tau = 0) = \frac{C_{0,0}^2}{G(\vartheta_0)} \cdot p_s(\vartheta, \tau) = |b_s|^2 \cdot \delta(\vartheta - \vartheta_0) \cdot \delta(\tau)$$

$$p(\vartheta, \tau) = \frac{C(\vartheta, \tau)^2}{G(\vartheta)} = |b(\vartheta, \tau)|^2 = \frac{1}{2\pi} \cdot \frac{1}{s} \cdot e^{-\tau/s} \quad (16)$$

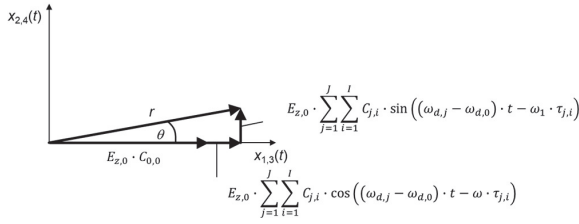
In the following the amplitude of the transmitted E-field signal is set to (17) and for the mobile station an omni-directional antenna is assumed in (18).

$$E_{z,0} = 1 \quad (17)$$

$$G(\vartheta) = 1 \quad (18)$$

3. Approach for Coherence Bandwidth Calculation with Envelope Correlation Function

The frequency and time dependency of the radio channel in (11) can be characterized by the coherence bandwidth B_c and coherence time Δt_c , which depend on the correlation between the signal envelopes at different frequencies ω_1 and ω_2 (frequency shift $\Delta\omega = \omega_2 - \omega_1$) and time instants t_1 and t_2 (time shift $\Delta t = t_2 - t_1$). The Doppler shift at ω_1 and ω_2 is basically the same for only slightly different carrier frequencies (small relative coherence bandwidth). However, the phase at different time instants is different. The detailed mathematical calculation is provided in Annex 1.


Figure 3.

Representation of (A3) to (A6) by amplitude r and phase θ in polar coordinates from the in-phase and quadrature components $x_1(t)$ and $x_2(t)$ or $x_3(t)$ and $x_4(t)$, respectively.

Annex 1.A shows the detailed description of the in-phase and quadrature components $x_1(t)$, $x_2(t)$, $x_3(t)$ and $x_4(t)$ at ω_1 and ω_2 and also their representation by amplitude (envelope) r and phase θ in polar coordinates (Figure 3).

The envelope cross-correlation function R_e between the envelopes in (A9) follows from the ensemble average of r_1 and r_2 by means of the joint probability distribution $p(r_1, r_2)$ in (Annex 1.B). $p(r_1, r_2)$ is derived in Annex 1.C in (A19) from the joint probability distribution $p(x_1, x_2, x_3, x_4)$ in (A16) of the quadrature components by variable substitution to polar coordinates r and θ (Figure 3).

4. Coherence Bandwidth of Radio Channels

4.1. Coherence Bandwidth of the Rayleigh Channel

The coherence bandwidth for the Rayleigh channel can analytically be calculated by means of the complete Elliptic integral $E(k)$ of the second kind. The basic complex calculation is shown in [11], pp. 45, however, without several intermediate steps.

The envelope cross-correlation function $R_e(\Delta\omega, \Delta t)$ follows with [11], p. 51 and λ according to (A15) in (19).

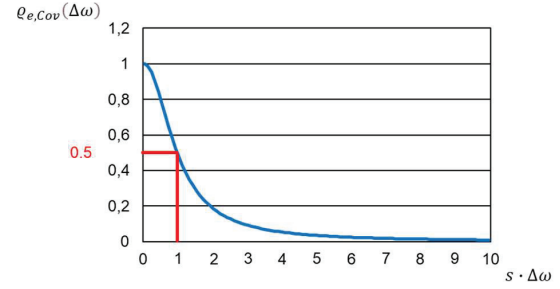
$$R_e(\Delta\omega, \Delta t) = \mu \epsilon \cdot (1 + \lambda) \cdot E\left(\frac{2 \cdot \sqrt{\lambda}}{1 + \lambda}\right) \quad (19)$$

In [11], p. 51 the coherence bandwidth B_c and time Δt_c is defined by the correlation coefficient $\rho_{e,Cov}(\Delta\omega, \Delta t)$ or covariance (CoV) between the two Rayleigh distributed envelopes r_1 and r_2 , which are separated versus frequency $\Delta\omega$ or time Δt , as defined in [15], p. 108 and 150. The correlation between both envelopes can also be expressed by the normalized cross-correlation function (CCF) $\rho_{e,CCF}(\Delta\omega, \Delta t)$, which will be relevant for the Rice channel.

$$\rho_{e,Cov}(\Delta\omega, \Delta t) = \frac{\langle r_1 \cdot r_2 \rangle - \langle r_1 \rangle \cdot \langle r_2 \rangle}{\sqrt{(\langle r_1^2 \rangle - \langle r_1 \rangle^2) \cdot (\langle r_2^2 \rangle - \langle r_2 \rangle^2)}} \quad (20a)$$

$$\rho_{e,CCF}(\Delta\omega, \Delta t) = \frac{\langle r_1 \cdot r_2 \rangle}{\sqrt{\langle r_1^2 \rangle \cdot \langle r_2^2 \rangle}} \quad (20b)$$

The covariance in (20a) is characterizing the fluctuation around the mean of the process, where the normalized cross-correlation function in (20b) ([15], pp. 215) is characterizing the fluctuation of the entire random process including the mean value.


Figure 4.

Covariance for the Rayleigh channel versus frequency shift $s \cdot \Delta\omega$ from (22).

With the different elements in (20a) and (20b) and the variance $\sigma^2 = \mu$ of the Gaussian processes x_1, x_2, x_3 and x_4 in (21a) to (21d) $\rho_{e,Cov}(\Delta\omega, \Delta t)$ and $\rho_{e,CCF}(\Delta\omega, \Delta t)$ result in (22) and (23).

$$\langle r_1 \cdot r_2 \rangle = R_e(\Delta\omega, \Delta t) \quad (19) \quad (21a)$$

$$\langle r_1 \rangle = \langle r_2 \rangle = \sqrt{\frac{\pi}{2}} \cdot \sigma \quad [6], \text{ p. 65} \quad (21b)$$

$$\langle r_1^2 \rangle = \langle r_2^2 \rangle = 2 \cdot \sigma^2 \quad [6], \text{ p. 65} \quad (21c)$$

$$\mu = \langle x_1^2 \rangle = \sigma^2 \quad [6], \text{ p. 64 and (A13)} \quad (21d)$$

$$\rho_{e,Cov}(\Delta\omega, \Delta t) = \frac{(1 + \lambda) \cdot E\left(\frac{2 \cdot \sqrt{\lambda}}{1 + \lambda}\right) - \frac{\pi}{2}}{2 - \frac{\pi}{2}} \quad (22)$$

$$\rho_{e,CCF}(\Delta\omega, \Delta t) = \frac{(1 + \lambda) \cdot E\left(\frac{2 \cdot \sqrt{\lambda}}{1 + \lambda}\right)}{2} \quad (23)$$

In [11], p. 51 the coherence bandwidth B_c is defined for $\rho_{e,Cov}(\Delta\omega, \Delta t) = 0.5$. From the numerical evaluation of (22) this condition is achieved at $\lambda_c = 0.725$ with (A15) and the approximation in (24) for practical applications [11], p. 51. For $\lambda_c = 0.725$ the normalized cross correlation function $\rho_{e,CCF}(\Delta\omega, \Delta t) = 0.893$.

$$\lambda_c^2 = \frac{J_0^2\left(\omega \cdot \frac{\nu}{c_0} \cdot \Delta t\right)}{1 + s^2 \cdot \Delta\omega^2} = 0.725^2 = 0.53 \approx 0.5 \quad (24)$$

(25) and (26) provide the coherence bandwidth B_c for mobile speed $\nu = 0$ and/or time shift $\Delta t = 0$ and the coherence time Δt_c as a function of maximum Doppler shift f_m for $\Delta\omega = 0$ [4], p. 165.

$$B_c \approx \frac{\Delta\omega}{2 \cdot \pi} = \frac{1}{2 \cdot \pi \cdot s} \quad (25)$$

$$\sqrt{\frac{1}{2}} \approx J_0\left(\omega \cdot \frac{\nu}{c_0} \cdot \Delta t\right) \quad \Delta t_c \approx 1.13 \cdot \frac{1}{2 \cdot \pi \cdot f_m} = \frac{9}{16 \cdot \pi} \cdot \frac{1}{f_m} \quad (26)$$

The covariance in (22) and the normalized cross-correlation function in (23) for the correct solution by means of the complete Elliptic integral is shown in Figures 4 and 5.

The relation between covariance $\rho_{e,Cov}(\Delta\omega, \Delta t) = 0.5$ corresponding to the normalized cross-correlation $\rho_{e,CCF}(\Delta\omega, \Delta t) = 0.893$ will be applied to the Rice channel.

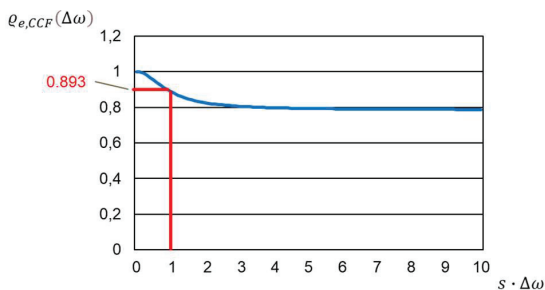


Figure 5. Normalized cross-correlation for the Rayleigh channel versus frequency shift $s \cdot \Delta\omega$ from (23).

4.2. Coherence Bandwidth of the Rice Channel

The method in Section 4.1 with an analytical solution cannot directly be applied to the Rice channel, because the joint probability density function in (A18) is very complex for $|b_s| > 0$. Therefore, an approximative approach is used for the Rice factor $K' \gg 1$ for the Rice distribution in Annex 2.A and the derivation of the envelope cross-correlation function in Annex 2.B. K' is defined in (27) [5], pp. 134 and $\sigma^2 = \mu$ in (A13). With the condition (17) the amplitude of the strong component r_s is replaced by $r_s = |b_s|$.

$$K' = \frac{r_s^2}{2 \cdot \sigma^2} = \frac{|b_s|^2}{2 \cdot \sigma^2} \quad (27)$$

The comparison in Figure 6 between the normalized Rice distribution in (A20) and the normalized Gaussian distribution in (A25) as approximation versus the Rice factor K' shows a good similarity for $K' > 2$ but deviations for $K' < 2$.

The covariance $\ell_{e,Cov}(\Delta\omega, \Delta t)$ and the normalized cross-correlation function $\ell_{e,CCF}(\Delta\omega, \Delta t)$ follow directly from (20) versus K' , which are provided for the coherence bandwidth and coherence time.

$$\begin{aligned} \ell_{e,Cov}(\Delta\omega, \Delta t) &= \frac{\langle r_1 \cdot r_2 \rangle - \langle r_1 \rangle \cdot \langle r_2 \rangle}{\sqrt{(\langle r_1^2 \rangle - \langle r_1 \rangle^2) \cdot (\langle r_2^2 \rangle - \langle r_2 \rangle^2)}} \\ &= \frac{\langle r \rangle^2 + \mu'_1 - \langle r \rangle \cdot \langle r \rangle}{\sqrt{(\langle r^2 \rangle - \langle r \rangle^2) \cdot (\langle r^2 \rangle - \langle r \rangle^2)}} \\ &= \frac{\mu'_1}{\langle r^2 \rangle - \langle r \rangle^2} = \frac{\mu'_1}{\sigma^2} \end{aligned} \quad (28)$$

$$\begin{aligned} \ell_{e,CCF}(\Delta\omega, \Delta t) &= \frac{\langle r_1 \cdot r_2 \rangle}{\sqrt{\langle r_1^2 \rangle \cdot \langle r_2^2 \rangle}} \\ &= \frac{\langle r \rangle^2 + \mu'_1}{\langle r^2 \rangle} = \frac{\langle r \rangle^2 + \mu'_1}{\sigma^2 + \langle r \rangle^2} = \frac{\langle r \rangle^2 + \mu'_1}{\mu'} \end{aligned} \quad (29)$$

The covariance $\ell_{e,Cov}(\Delta\omega, \Delta t)$ is independent of the Rice factor K' (corresponding to [24]), where the normalized cross-correlation function depends on K' . However, with increasing K' the different representations of the random processes r_1 and r_2 are increasingly similar and for $K' \rightarrow \infty$ both are the same, i.e., the

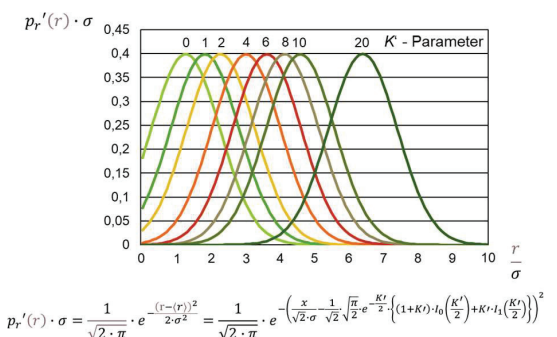
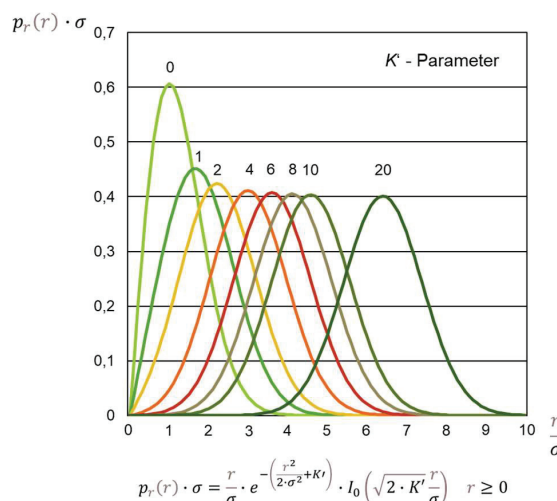


Figure 6. Upper Figure: Normalized Rice distribution $p_r(r) \cdot \sigma$ according to (A20) versus K' . Lower Figure: Normalized Gaussian approximation $p'_r(r) \cdot \sigma$ according to (A25) versus K' with the mean of the Rice distribution in (A21).

correlation between r_1 and r_2 tends to 1. Therefore, for the Rice channel with a strong component the normalized cross-correlation function in (29) is a reasonable measure to derive the coherence bandwidth.

Coherence bandwidth $\Delta\omega = 2\pi B_c$: In this case the mobile speed is assumed $v = 0$. Then, the angle of arrival ϑ_0 of the strong component does not matter. With (28) $\ell_{e,Cov}(\Delta\omega)$ with the delay spread s of the channel impulse response without the strong component is provided in (30). Figure 7 is comparing (30) as the Rice approximation with the correct formula for the Rayleigh channel in (22) and the Rayleigh approximation in (24). There is nearly no difference between the different representations.

$$\ell_{e,Cov}(\Delta\omega) = \frac{1}{1 + s^2 \cdot \Delta\omega^2} \quad (30)$$

The normalized cross-correlation function $\ell_{e,CCF}(\Delta\omega, K')$ with (29) and $\langle r(K') \rangle$ in (A21) as a function of K' is given in (31). Figure 8 is showing the evaluation of (31) and comparing it with (20b) and for the Rayleigh channel in (23). As can be expected from the comparison of the Rice distribution and the Gaussian approximation in Figure 6, the Gaussian approximation for $\ell_{e,CCF}(\Delta\omega, K')$ leads to a too low correlation for $K' < 2$.

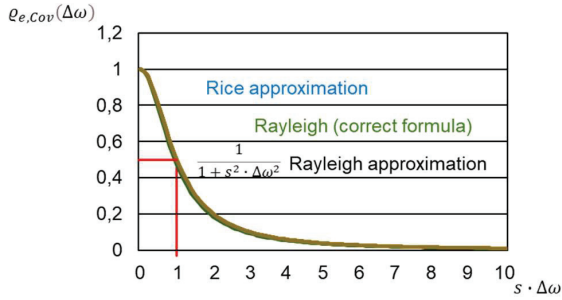


Figure 7. Coherence bandwidth $2\pi B_c$ based on the covariance of the cross-correlation function, comparison of (30) with (22) and (24).

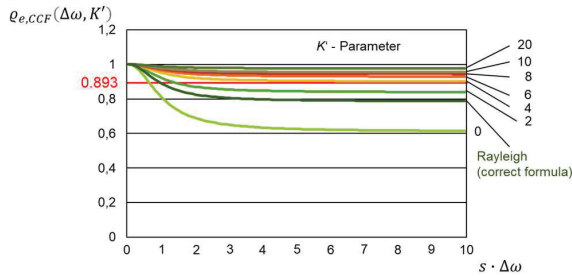


Figure 8. Coherence bandwidth $2\pi B_c$ based on the normalized cross-correlation function with the mean according to (A21), K' from 0 to 20, comparison of (31) with (20b) and (23).

For $K' > 2$ the correlation $\rho_{e,CCF}(\Delta\omega, K')$ provides a good approximation.

$$\rho_{e,CCF}(\Delta\omega, K') = \frac{\frac{\langle r \rangle^2}{\sigma^2} + \frac{1}{1+s^2 \cdot \Delta\omega^2}}{1 + \frac{\langle r \rangle^2}{\sigma^2}} \quad (31)$$

With the condition $\rho_{e,CCF}(\Delta\omega, K') = 0.893$ applied to (31) as for the Rayleigh channel the coherence bandwidth can be calculated from numerical evaluation, which provides a too small coherence bandwidth for $K' < 2$. With $\Delta\omega \cdot s \approx 1$ in (24) for the Rayleigh channel an approximative formula is derived, which is valid in the entire range $K' \geq 0$ (32). From $K' \geq 3.63$ the coherence bandwidth $2\pi B_c \cdot s \rightarrow \infty$ with the definition above. Figure 9 shows the good fit of (32) for the entire range including the limiting case of the Rayleigh channel at $K' = 0$.

$$s \cdot 2\pi B_c \approx \left(\frac{3.63}{3.63 - K'} \right)^{0.5250} \quad 0 \leq K' \leq 3.63 \quad (32)$$

Coherence time Δt_c : In this case the frequency shift is assumed $\Delta\omega = 0$ as well as the angle of arrival $\vartheta_0 = 0$ of the strong component. With (28) $\rho_{e,Cov}(\Delta t)$ with the delay spread s of the channel impulse response without the strong component is provided in (33). This differs for bigger time shifts from (24) from the Rayleigh channel.

$$\rho_{e,Cov}(\Delta t) = J_0 \left(\omega \cdot \frac{v}{c_0} \cdot \Delta t \right) \cdot \cos \left(\omega \cdot \frac{v}{c_0} \cdot \Delta t \right) \quad (33)$$

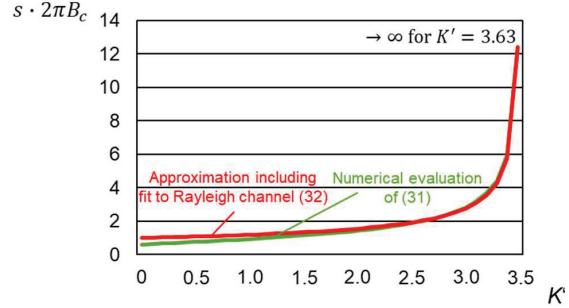


Figure 9. Normalized coherence bandwidth versus K' based on (31) and the approximation in (32).

The normalized cross-correlation function $\rho_{e,CCF}(\Delta t, K')$ with (29) and $\langle r(K') \rangle$ in (A21) as a function of K' is given in (34). As can be expected from the comparison of the Rice distribution and the Gaussian approximation in Figure 6, the Gaussian approximation for $\rho_{e,CCF}(\Delta t, K')$ leads to a too low correlation for $K' < 2$. For $K' > 2$ the correlation $\rho_{e,CCF}(\Delta\omega, K')$ provides a good approximation.

$$\rho_{e,CCF}(\Delta t, K') = \frac{\frac{\langle r \rangle^2}{\sigma^2} + J_0 \left(\omega \cdot \frac{v}{c_0} \cdot \Delta t \right) \cdot \cos \left(\omega \cdot \frac{v}{c_0} \cdot \Delta t \right)}{1 + \frac{\langle r \rangle^2}{\sigma^2}} \quad (34)$$

With the condition $\rho_{e,CCF}(\Delta t, K') = 0.893$ applied to (34) as for the Rayleigh channel the coherence time can be calculated from numerical evaluation, which provides a too small coherence time for $K' < 2$. With $\omega \cdot \frac{v}{c_0} \cdot \Delta t_c \approx 1.13$ in (26) for the Rayleigh channel an approximative formula is derived, which is valid in the entire range $K' \geq 0$ (35). From $K' \geq 4.075$ the coherence time $\omega \cdot \frac{v}{c_0} \cdot \Delta t_c \rightarrow \infty$ with the definition above.

$$\omega \cdot \frac{v}{c_0} \cdot \Delta t_c \approx 1.13 \cdot \left(\frac{4.075}{4.075 - K'} \right)^{0.117} \quad 0 \leq K' \leq 4.075 \quad (35)$$

4.3. Relation Between the Delay Spread of the Rayleigh and the Rice Channel

With the definition of the delay spread in [4], pp. 160 a relation between the delay spread between the Rayleigh and the Rice channel can be derived in (36) under the assumption that the strong component in the Rice channel occurs at the beginning of the channel impulse response at $\tau = 0$. With increasing K' s_{Rice} is decreasing for constant $s_{Rayleigh}$.

$$s_{Rice}|_{\tau=0} = s_{Rayleigh} \cdot \sqrt{\frac{1}{K'+1} \cdot \left(2 - \frac{1}{K'+1} \right)} \quad (36)$$

4.4. Practical Applications

Especially in the millimeter wave and sub-Terahertz domain coverage under shadowing conditions can only be ensured by

additional technical means in network deployment such as reflectors, RIS arrays or repeaters [8]. Such systems provide a strong component in the channel impulse response, which is seen by the receiver. In these frequency ranges the roughness of reflecting or scattering surfaces such as wall material plays an increasing role ([3], p. H6, [5], pp. 26). If the roughness is in the order of the carrier frequency wavelength, the non-specular reflection or scattering in many different directions is increasing, which results in a reduced reflection coefficient for rough surfaces.

If reflectors, RIS arrays or repeaters are deployed in such areas, a strong component in the channel impulse response with a Rice factor $K' > 4$ can be ensured or at least targeted, which results in a significant increase of the channel coherence bandwidth or even in nearly infinite coherence bandwidth as well as increased coherence time. The radio channel is approaching effectively a single path channel. This is relevant for very wideband systems in 6G and allows reduced effective inter-symbol interference, significantly reduced signal processing requirements for channel estimation, equalization effort and a reduced update rate.

5. Conclusions

In the millimeter wave and sub-Terahertz frequency range very wideband systems are intended as part of 6G systems. Therefore, the coherence bandwidth of radio channels under practical conditions or by additional technical means (reflectors, RIS arrays or repeaters) for network deployment is of interest to reduce requirements on the radio interface design and its potential simplification. The cross-correlation between different fading signals, which are separated in frequency and time, is investigated for Rice channels and compared with the analytical solution for Rayleigh channels.

The analytical approach for Rayleigh channels cannot directly be applied to Rice channels. Therefore, the Rice distribution is approximated by a Gaussian distribution for bigger Rice factors K' to calculate the cross-correlation function. This provides reasonably good results for the coherence bandwidth and coherence time for $K' > 2$. For $K' < 2$ the Gaussian distribution deviates significantly from the Rice distribution ($K' \rightarrow 0$ corresponds to the Rayleigh distribution) and therefore deviations of the coherence bandwidth and time are understandable.

In the literature (e.g., [11]) the coherence bandwidth is defined by the covariance. However, it is shown that for Rice channels this definition is misleading, because the impact of a strong component is ignored. In a Rice channel the fading fluctuations are around the direct component and the relative fluctuation is decreasing. For a Rice factor $K' \rightarrow \infty$ the relative fluctuations normalized to the strong component will disappear. It is therefore proposed to use the normalized cross-correlation function as a measure for the coherence bandwidth and coherence time. The coherence bandwidth and coherence time is here defined for the normalized cross-correlation function according to the condition in (37), which corresponds to a covariance of around 0.5 for Rayleigh channels as proposed in [11].

$$\rho_{e,CCF}(\Delta\omega, \Delta t) \geq 0.893 \quad (37)$$

The results of the covariance for the Rice channel for frequency shift in the approximation in Section 4.2 are similar like in [11] and independent of K' . With the condition in (37)

the normalized cross-correlation function provides the coherence bandwidth and time, which depends on K' .

Based on the numerical evaluation the relation between the Rice factor K' and the coherence bandwidth and time is derived and presented. Approximative formulas are developed, which are adapted for small K' to the well-known result in [11] for $K' = 0$ for the Rayleigh channel. These formulas show a good estimate for the coherence bandwidth and time. For $K' = 3.63$ corresponding to 5.6 dB for the coherence bandwidth and $K' = 4.075$ corresponding to 6.1 dB for the coherence time both are becoming infinite under the definition in (37). Therefore, a Rice channel can be regarded as nearly frequency independent for $K > 6$ dB.

With the consideration on the roughness of reflecting/scattering surfaces in the environment (Section 4.4) and additional technical means like reflectors, RIS arrays or repeaters ([8]) a Rice type channel can be enforced even under shadowing conditions. Such conditions allow reduced inter-symbol interference and relax the requirements on signal processing for channel estimation, equalization and update rates significantly for very wideband systems.

Annex 1: Derivation of Envelope Correlation Function

A. Signal Description with Frequency and Time Shift

With (11) the signals at both frequencies ω_1 and ω_2 are described with the narrowband in-phase and quadrature components in (A1) and (A2).

$$E_z'(\omega_1, t) = x_1(t) \cdot \cos(\omega_1 \cdot t) - x_2(t) \cdot \sin(\omega_1 \cdot t) \quad (A1)$$

$$E_z'(\omega_2, t + \Delta t) = x_3(t) \cdot \cos(\omega_2 \cdot t) - x_4(t) \cdot \sin(\omega_2 \cdot t) \quad (A2)$$

with

$$x_1(t) = E_{z,0} \cdot C_{0,0} + E_{z,0} \cdot \sum_{j=1}^J \sum_{i=1}^I C_{j,i} \cdot \cos((\omega_{d,j} - \omega_{d,0}) \cdot t - \omega_1 \cdot \tau_{j,i}) \quad (A3)$$

$$x_2(t) = E_{z,0} \cdot \sum_{j=1}^J \sum_{i=1}^I C_{j,i} \cdot \sin((\omega_{d,j} - \omega_{d,0}) \cdot t - \omega_1 \cdot \tau_{j,i}) \quad (A4)$$

$$x_3(t) = E_{z,0} \cdot C_{0,0} + E_{z,0} \cdot \sum_{j=1}^J \sum_{i=1}^I C_{j,i} \cdot \cos((\omega_{d,j} - \omega_{d,0}) \cdot (t + \Delta t) - \omega_2 \cdot \tau_{j,i}) \quad (A5)$$

$$x_4(t) = E_{z,0} \cdot \sum_{j=1}^J \sum_{i=1}^I C_{j,i} \cdot \sin((\omega_{d,j} - \omega_{d,0}) \cdot (t + \Delta t) - \omega_2 \cdot \tau_{j,i}) \quad (A6)$$

With I and $J \rightarrow \infty$ in the sums in (11) the processes $x_1(t)$, $x_2(t)$, $x_3(t)$ and $x_4(t)$ are regarded as Gaussian random processes according to the Central Limit Theorem [13], pp. 51, where

- $x_1(t)$ and $x_3(t)$ have the mean $E_{z,0} \cdot C_{0,0}$ and
- $x_2(t)$ and $x_4(t)$ have zero mean.

In [11], p. 48 the variables in (A7) are defined at the same fixed time t for further considerations.

$$\begin{aligned} x_1 &\hat{=} x_1(t) & x_2 &\hat{=} x_2(t) \\ x_3 &\hat{=} x_3(t + \Delta t) & x_4 &\hat{=} x_4(t + \Delta t) \end{aligned} \quad (\text{A7})$$

$x_1(t)$, $x_2(t)$, $x_3(t)$ and $x_4(t)$ are represented by amplitude r and phase θ in polar coordinates (Figure 3).

$$\begin{aligned} x_1 &= r_1 \cdot \cos \theta_1 & x_2 &= r_1 \cdot \sin \theta_1 \\ x_3 &= r_2 \cdot \cos \theta_2 & x_4 &= r_2 \cdot \sin \theta_2 \end{aligned} \quad (\text{A8})$$

B. Envelope Cross-correlation Function

According to [3], pp. D12, [11], p. 50 and [14], pp. 44 the envelope cross-correlation function R_e between the envelopes r_1 and r_2 versus frequency shift $\Delta\omega$ and time shift Δt can be derived from the ensemble average in (A9) by the integral of the joint probability distribution function $p(r_1, r_2, \theta_1, \theta_2)$ versus θ_1 and θ_2 (c.f. A19) under the simplified assumption of ergodic random processes.

$$R_e(\Delta\omega, \Delta t) = \langle r_1 \cdot r_2 \rangle = \int_0^\infty \int_0^\infty r_1 \cdot r_2 \cdot p(r_1, r_2) \, dr_1 dr_2 \quad (\text{A9})$$

C. Joint Probability Distribution Functions

The Rice radio channel at different frequencies and time instances above is described in (A1) and (A2) by quadrature components with the Gaussian processes $x_1(t)$, $x_2(t)$, $x_3(t)$ and $x_4(t)$. This leads at first to the joint probability distribution function of these processes $p(x_1, x_2, x_3, x_4)$. However, this needs then to be transformed with (A8) to the joint probability density function $p(r_1, r_2, \theta_1, \theta_2)$ versus amplitude r and phase θ .

The n -dimensional normal distribution is described in [15], pp. 126, [16], pp. 366, [17], pp. 191 and [18], pp. 16 and is given for the case of $n = 4$ in (A10) with the means m .

$$p(x_1, x_2, x_3, x_4) = \frac{1}{(2\pi)^2 \cdot \sqrt{|M|}} \cdot e^{-\frac{1}{2 \cdot |M|} \sum_{j,k} |M_{jk}| \cdot (x_j - m_j) \cdot (x_k - m_k)} \quad (\text{A10})$$

The matrix M

$$M = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} & \lambda_{14} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} & \lambda_{24} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} & \lambda_{34} \\ \lambda_{41} & \lambda_{42} & \lambda_{43} & \lambda_{44} \end{pmatrix} = \begin{pmatrix} \lambda_{11} & 0 & \lambda_{13} & \lambda_{14} \\ 0 & \lambda_{11} & -\lambda_{14} & \lambda_{13} \\ \lambda_{13} & -\lambda_{14} & \lambda_{11} & 0 \\ \lambda_{14} & \lambda_{13} & 0 & \lambda_{11} \end{pmatrix} \quad (\text{A11})$$

corresponds with [17], p. 115 and (A12) to the symmetric matrix of moments with the elements $\lambda_{jk} = \lambda_{kj}$. The λ_{jk} are the covariances of all pairs of the random processes and $\lambda_{ii} = \sigma_i^2$ are the variances

of these processes, which are the same. The λ_{jk} follow from (A3) to (A6) with (17) and (18).

$$\begin{aligned} \lambda_{11} &= \lambda_{22} = \lambda_{33} = \lambda_{44} = \langle x_1^2 \rangle - |b_s|^2 = \langle x_3^2 \rangle - |b_s|^2 \\ &= \langle x_2^2 \rangle = \langle x_4^2 \rangle \\ \lambda_{12} &= \lambda_{21} = \langle x_1 \cdot x_2 \rangle = \langle x_2 \cdot x_1 \rangle = 0 \\ \lambda_{13} &= \lambda_{31} = \lambda_{24} = \lambda_{42} = \langle x_1 \cdot x_3 \rangle - |b_s|^2 = \langle x_3 \cdot x_1 \rangle - |b_s|^2 \\ &= \langle x_2 \cdot x_4 \rangle = \langle x_4 \cdot x_2 \rangle \\ \lambda_{14} &= \lambda_{41} = -\lambda_{23} = -\lambda_{32} = \langle x_1 \cdot x_4 \rangle = \langle x_4 \cdot x_1 \rangle \\ &= -\langle x_2 \cdot x_3 \rangle = -\langle x_3 \cdot x_2 \rangle \\ \lambda_{34} &= \lambda_{43} = \langle x_3 \cdot x_4 \rangle = \langle x_4 \cdot x_3 \rangle = 0 \end{aligned} \quad (\text{A12})$$

With the notation in [11], p. 49, [19], pp. 133, [20], p. 451, No. 13 and 18 and [21], p. 360 and the Bessel function J_0 these moments follow in (A13).

$$\begin{aligned} \mu &= \lambda_{11} = \langle x_1^2 \rangle - |b_s|^2 = \sigma^2 = \frac{1}{2} \\ \mu_1 &= \lambda_{13} = \langle x_1 \cdot x_3 \rangle - |b_s|^2 = \frac{1}{2} \cdot J_0 \left(\omega \cdot \frac{\nu}{c_0} \cdot \Delta t \right) \\ &\quad \cdot \frac{\cos \left(\left(\omega \cdot \frac{\nu}{c_0} \cdot \cos \vartheta_0 \right) \cdot \Delta t \right) - s \cdot \Delta \omega \cdot \sin \left(\left(\omega \cdot \frac{\nu}{c_0} \cdot \cos \vartheta_0 \right) \cdot \Delta t \right)}{1 + s^2 \cdot \Delta \omega^2} \\ \mu_2 &= \lambda_{14} = \langle x_1 \cdot x_4 \rangle = -\frac{1}{2} \cdot J_0 \left(\omega \cdot \frac{\nu}{c_0} \cdot \Delta t \right) \\ &\quad \cdot \frac{\sin \left(\left(\omega \cdot \frac{\nu}{c_0} \cdot \cos \vartheta_0 \right) \cdot \Delta t \right) + s \cdot \Delta \omega \cdot \cos \left(\left(\omega \cdot \frac{\nu}{c_0} \cdot \cos \vartheta_0 \right) \cdot \Delta t \right)}{1 + s^2 \cdot \Delta \omega^2} \end{aligned} \quad (\text{A13})$$

The determinant $|M|$ of matrix M in (A11) results in (A14).

$$|M| = [\lambda_{11}^2 - (\lambda_{13}^2 + \lambda_{14}^2)]^2 = \mu^2 \cdot \left(1 - \frac{\mu_1^2 + \mu_2^2}{\mu^2} \right) \quad (\text{A14})$$

With [11], p. 49 and [17], pp. 191 the following term λ^2 is interpreted as correlation coefficient of the random variables, which is independent of the angle of arrival ϑ_0 of the direct component.

$$\lambda^2 = \frac{\mu_1^2 + \mu_2^2}{\mu^2} = \frac{J_0^2 \left(\omega \cdot \frac{\nu}{c_0} \cdot \Delta t \right)}{1 + s^2 \cdot \Delta \omega^2} \quad (\text{A15})$$

The terms $|M_{jk}|$ in (A10) correspond to the algebraic complements or adjuncts of the covariance matrix of the matrix elements λ_{jk} in M and are calculated with [22], p. 39.

The joint probability distribution function $p(x_1, x_2, x_3, x_4)$ in (A10) is derived in (A16).

$$p(x_1, x_2, x_3, x_4) = \frac{1}{(2\pi)^2 \cdot \mu^2 \cdot (1 - \lambda^2)}$$

$$\begin{aligned} & \cdot \exp \left[-\frac{1}{2 \cdot \mu^2 \cdot (1 - \lambda^2)} \cdot \{ \mu \cdot [(x_1 - |b_s|) \right. \\ & \cdot (x_1 - |b_s|) + (x_2) \cdot (x_2) + (x_3 - |b_s|) \\ & \cdot (x_3 - |b_s|) + (x_4) \cdot (x_4)] - 2 \cdot \mu_1 \\ & \cdot [(x_1 - |b_s|) \cdot (x_3 - |b_s|) + (x_2) \cdot (x_4)] - 2 \\ & \left. \cdot \mu_2 \cdot [(x_1 - |b_s|) \cdot (x_4) - (x_2) \cdot (x_3 - |b_s|)] \right\} \end{aligned} \quad (A16)$$

The required probability density $p(r_1, r_2, \theta_1, \theta_2)$ for polar coordinates is calculated in (A18) by substitution of variables with (A8) and the Jacobi-determinant J [23], pp. 37 in (A17).

$$\begin{aligned} J &= \begin{vmatrix} \frac{\partial r_1 \cdot \cos \theta_1}{\partial r_1} & \frac{\partial r_1 \cdot \sin \theta_1}{\partial r_1} & \frac{\partial r_2 \cdot \cos \theta_2}{\partial r_1} & \frac{\partial r_2 \cdot \sin \theta_2}{\partial r_1} \\ \frac{\partial r_1 \cdot \cos \theta_1}{\partial \theta_1} & \frac{\partial r_1 \cdot \sin \theta_1}{\partial \theta_1} & \frac{\partial r_2 \cdot \cos \theta_2}{\partial \theta_1} & \frac{\partial r_2 \cdot \sin \theta_2}{\partial \theta_1} \\ \frac{\partial r_1 \cdot \cos \theta_1}{\partial r_2} & \frac{\partial r_1 \cdot \sin \theta_1}{\partial r_2} & \frac{\partial r_2 \cdot \cos \theta_2}{\partial r_2} & \frac{\partial r_2 \cdot \sin \theta_2}{\partial r_2} \\ \frac{\partial r_1 \cdot \cos \theta_1}{\partial \theta_2} & \frac{\partial r_1 \cdot \sin \theta_1}{\partial \theta_2} & \frac{\partial r_2 \cdot \cos \theta_2}{\partial \theta_2} & \frac{\partial r_2 \cdot \sin \theta_2}{\partial \theta_2} \end{vmatrix} \\ &= r_1 \cdot r_2 \end{aligned} \quad (A17)$$

$$\begin{aligned} p(r_1, r_2, \theta_1, \theta_2) &= \frac{r_1 \cdot r_2}{(2\pi)^2 \cdot \mu^2 \cdot (1 - \lambda^2)} \\ & \cdot \exp \left[-\frac{r_1^2 + r_2^2 + 2 \cdot |b_s|^2 \cdot \left(1 - \frac{\mu_1}{\mu}\right)}{2 \cdot \mu \cdot (1 - \lambda^2)} \right. \\ & + \frac{2 \cdot |b_s| \cdot (r_1 \cdot \cos \theta_1 + r_2 \cdot \cos \theta_2)}{2 \cdot \mu \cdot (1 - \lambda^2)} \\ & + \frac{2 \cdot \lambda \cdot r_1 \cdot r_2 \cdot \cos(\theta_2 - \theta_1 - \varphi)}{2 \cdot \mu \cdot (1 - \lambda^2)} \\ & - \frac{2 \cdot \lambda \cdot |b_s| \cdot r_1 \cdot \cos(\theta_1 + \varphi)}{2 \cdot \mu \cdot (1 - \lambda^2)} \\ & \left. - \frac{2 \cdot \lambda \cdot |b_s| \cdot r_2 \cdot \cos(\theta_2 - \varphi)}{2 \cdot \mu \cdot (1 - \lambda^2)} \right] \end{aligned} \quad (A18)$$

(A18) results with $|b_s| = 0$ directly in the case of the Rayleigh channel [11], p. 49. The required joint probability density function $p(r_1, r_2)$ in (A9) follows from integration of (A18) versus θ_1 and θ_2 .

$$p(r_1, r_2) = \int_0^{2\pi} \int_0^{2\pi} p(r_1, r_2, \theta_1, \theta_2) d\theta_1 d\theta_2 \quad (A19)$$

Annex 2: Derivation of Envelope Correlation Function for Rice Channels

A. Approximation of the Rice Distribution for Bigger K'

For $|b_s| > 0$ with a strong component the amplitude distribution of the received signal under multipath propagation conditions is

described by the Rice distribution [3], p. H3, [5], pp. 134, [6], pp. 26 and [18], pp. 34.

$$p_r(r) = \frac{r}{\sigma^2} \cdot e^{-\frac{r^2 + |b_s|^2}{2\sigma^2}} \cdot I_0\left(\frac{r \cdot |b_s|}{\sigma^2}\right) \quad r \geq 0 \quad (A20)$$

The following means apply with $|b_s| = \sigma \cdot \sqrt{2 \cdot K'}$ ([6], pp. 64 and [18], p. 17).

$$\begin{aligned} \langle r_1 \rangle &= \langle r_2 \rangle = \langle r \rangle = \sigma \cdot \sqrt{\frac{\pi}{2}} \cdot e^{-\frac{K'}{2}} \\ & \cdot \left\{ (1 + K') \cdot I_0\left(\frac{K'}{2}\right) + K' \cdot I_1\left(\frac{K'}{2}\right) \right\} \end{aligned} \quad (A21)$$

$$\langle r_1^2 \rangle = \langle r_2^2 \rangle = \langle r^2 \rangle = 2 \cdot \sigma^2 \cdot (1 + K') \quad (A22)$$

For $K' \gg 1$ the Rice distribution can be approximated by a Gaussian distribution. The modified Bessel function $I_0(z)$ in (A20) with $z = \frac{r \cdot |b_s|}{\sigma^2} \gg 1$ corresponding to $K' \gg 1$ is replaced by series expansion [21], p. 377.

$$I_0(z) = \frac{e^z}{\sqrt{2 \cdot \pi \cdot z}} \cdot \left(1 + \frac{1}{8 \cdot z} + \frac{9}{21 \cdot (8 \cdot z)^2} + \dots \right) \quad (A23)$$

If only the first term of the series expansion in (A23) is considered, $I_0(z)$ is approximated by (A24).

$$I_0\left(\frac{r \cdot |b_s|}{\sigma^2}\right) \approx \frac{e^{\frac{r \cdot |b_s|}{\sigma^2}}}{\sqrt{2 \cdot \pi \cdot \frac{r \cdot |b_s|}{\sigma^2}}} \quad (A24)$$

The distribution function of the envelope in (A20) is then approaching a Gaussian distribution [7], p. 27 and [17], p. 177, because $r \approx r_s$ or $r \approx |b_s|$. The mean of the Gaussian approximation is the same as in (A21) for the Rice distribution, where r is now a Gaussian variable.

$$\begin{aligned} p'_r(r) \cdot \sigma &= \frac{1}{\sqrt{2 \cdot \pi}} \cdot e^{-\frac{(r-r_s)^2}{2 \cdot \sigma^2}} \\ &= \frac{1}{\sqrt{2 \cdot \pi}} \\ & \cdot e^{-\left(\frac{r}{\sqrt{2} \cdot \sigma} - \frac{1}{\sqrt{2}} \cdot \sqrt{\frac{K'}{2}} \cdot e^{-\frac{K'}{2}} \cdot \left\{ (1+K') \cdot I_0\left(\frac{K'}{2}\right) + K' \cdot I_1\left(\frac{K'}{2}\right) \right\} \right)^2} \end{aligned} \quad (A25)$$

This approximation shows the probability distribution of the envelope r . In this approximation for $K' \gg 1$ no quadrature components are considered as in Figure 3.

B. Approximation of the Envelope Cross-correlation for Rice Channels

For the calculation of the cross-correlation function $R_e(\Delta\omega, \Delta t)$ between the envelopes in (A9) the joint probability function $p'(r_1, r_2)$ is needed with the approximation in (A25).

$$R_e(\Delta\omega, \Delta t) = \langle r_1 \cdot r_2 \rangle \approx \int_0^\infty \int_0^\infty r_1 \cdot r_2 \cdot p'(r_1, r_2) dr_1 dr_2 \quad (A26)$$

The approximation $p'(r_1, r_2)$ follows with [17], p. 191.

$$p'(r_1, r_2) = \frac{1}{2 \cdot \pi \cdot \sigma^2 \cdot \sqrt{1 - \rho^2}} \cdot \exp \left\{ -\frac{1}{2 \cdot (1 - \rho^2)} \cdot \left[\frac{r_1^2 - 2 \cdot \langle r \rangle \cdot r_1 + \langle r \rangle^2}{\sigma^2} - 2 \cdot \rho \cdot \frac{r_1 \cdot r_2 - r_1 \cdot \langle r \rangle - r_2 \cdot \langle r \rangle + \langle r \rangle^2}{\sigma^2} + \frac{r_2^2 - 2 \cdot \langle r \rangle \cdot r_2 + \langle r \rangle^2}{\sigma^2} \right] \right\} \quad (A27)$$

ρ corresponds to the correlation coefficient between the two Gaussian processes r_1 and r_2 with the elements in (A28).

$$\rho(r_1, r_2) = \frac{\langle r_1 \cdot r_2 \rangle - \langle r_1 \rangle \cdot \langle r_2 \rangle}{\sqrt{(\langle r_1^2 \rangle - \langle r_1 \rangle^2) \cdot (\langle r_2^2 \rangle - \langle r_2 \rangle^2)}} = \frac{\mu'_1}{\sigma^2} = J_0 \left(\omega \cdot \frac{v}{c_0} \cdot \Delta t \right) \cdot \frac{\cos \left(\left(\omega \cdot \frac{v}{c_0} \cdot \cos \vartheta_0 \right) \cdot \Delta t \right) - s \cdot \Delta \omega \cdot \sin \left(\left(\omega \cdot \frac{v}{c_0} \cdot \cos \vartheta_0 \right) \cdot \Delta t \right)}{1 + s^2 \cdot \Delta \omega^2} \quad (A28)$$

with

$$\langle r_1 \cdot r_2 \rangle = \langle r \rangle^2 + \mu'_1 \quad (A13), (A21)$$

$$\langle r_1 \rangle = \langle r_2 \rangle = \langle r \rangle \quad (A21)$$

$$\langle r_1^2 \rangle = \langle r_2^2 \rangle = \sigma^2 + \langle r \rangle^2 = \mu' \quad (A13), (A21)$$

In the case of a joint Gaussian distribution the correlation coefficient (A28) already corresponds to the covariance. The proof is shown by the solution of the integral (A26) by [19], p. 65, No.314.6c and No. 314.6d with $p'(r_1, r_2)$ in (A27), $\langle r \rangle$ in (A21) and ρ in (A28). This results in (A29).

$$R_e(\Delta\omega, \Delta t) \approx \langle r_1 \cdot r_2 \rangle = \langle r \rangle^2 + \rho \cdot \sigma^2 \quad (A29)$$

References

[1] ITU-R, "WP 5D Workshop on IMT for 2030 and beyond," June 14, 2022, Geneva, Switzerland, <https://www.itu.int/en/ITU-R/study-groups/rsg5/rwp5d/Pages/default.aspx>.

[2] W. Mohr, "Range and capacity considerations for Terahertz-Systems for 6G mobile and wireless communication," *Journal of Mobile Multimedia*, Vol. 19, Issue 1, 2022, pp. 215, published September 20, 2022, DOI: <https://doi.org/10.13052/jmm1550-4646.19111>, citation link: <https://journals.riverpublishers.com/index.php/JMM/article/view/18511>. (Range and Capacity Considerations for Terahertz-Systems for 6G Mobile and Wireless Communication | Journal of Mobile Multimedia (riverpublishers.com)).

[3] Meinke/Gundlach, *Taschenbach der Hochfrequenztechnik*, Berlin: Springer-Verlag 1986.

[4] T.S. Rappaport, *Wireless Communications – Principles & Practice*, New Jersey: Prentice Hall 1996.

[5] J.B. Parsons J.B., *The Mobile Radio Propagation Channel*, Pentech Press, London, 1992.

[6] W.C.Y. Lee, *Mobile Communications Engineering*, McGraw-Hill Book Company, New York, 1982.

[7] W.C.Y. Lee, *Mobile Communications Design Fundamentals*, Second edition, John Wiley & Sons, New York, 1993.

[8] W. Mohr, "Coverage improvements for sub-Terahertz systems under shadowing conditions," *Journal of Telecommunications and Information Technology (JTIT)*, September 2023, (doi: 10.26636/jtit.2023.3.1301), <https://doi.org/10.26636/jtit.2023.3.1301>.

[9] O. Zinke and H. Brunswig, *Lehrbuch der Hochfrequenztechnik, Vol. I*, Springer Verlag, Berlin, Heidelberg, New York, 1973.

[10] W. Mohr, "Achievable bandwidth of Reconfigurable Intelligent Surfaces (RIS) concepts towards 6G communications," *12th CONASENSE Symposium*, June 27 and 28, 2022, Munich Germany, River Publishers, https://www.riverpublishers.com/research_article_details.php?book_id=1051&cid=2.

[11] W.C. Jakes, *Microwave Mobile Communications*, John Wiley & Sons, New York, 1974.

[12] I. Bronstein and K.A. Semendjajew, *Taschenbuch der Mathematik*, Verlag Harri Deutsch, Zürich and Frankfurt/Main, 1974.

[13] J.B. Thomas, *An Introduction to Statistical Communication Theory* John Wiley & Sons, New York, 1969.

[14] J.G. Proakis, *Digital Communications*, McGraw-Hill Book Company, New York, second edition, 1989.

[15] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill Book Company, International Student Edition, 1984.

[16] R.G. Gallager, *Stochastic Processes – Theory for Applications*, Cambridge University Press, Cambridge, UK, 2013.

[17] M. Fisz, *Wahrscheinlichkeitsrechnung und mathematische Statistik*, VEB Deutscher Verlag der Wissenschaften, Berlin, 1980.

[18] M. Pätzold, *Mobile Fading Channels*, John Wiley & Sons, Ltd, Chichester, 2002.

[19] W. Gröbner and N. Hofreiter, *Integraltafel – Zweiter Teil Bestimmte Integrale*, 5th edition, Springer-Verlag, Vienna, New York, 1973.

[20] I. Gradstein and I. Ryzhik, *Tables of series, products and integrals. Vol. I*, Verlag Harri Deutsch, Thun Frankfurt/M, 1981.

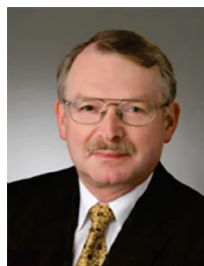
[21] M. Abramowitz and I.S. Stegun, *Handbook of mathematical functions*, Dover Publications, Inc., New York, 1972.

[22] R. Zurmühl and S. Falk, *Matrizen 1 – Grundlagen*, Springer Verlag, Berlin, 6th edition, 1992.

[23] W.B. Davenport and W.L. Root, *An Introduction of the Theory of Random Signals and Noise*, McGraw-Hill Book Company, New York, 1958.

[24] J.R. Mendes, M.D. Yacoub and D. Benevidas da Costa, "Closed-form generalized power correlation coefficient of Ricean channels," *European Transactions on Telecommunications, Euro. Trans. Telecomm. (in press)* Published online in Wiley InterScience (www.interscience.wiley.com) DOI: 10.1002/ett.1150, 2006.

Biography



Werner Mohr (M'91–SM'96) was graduated from the University of Hannover, Germany, with the Master Degree in electrical engineering in 1981 and with the Ph.D. degree in 1987.

He joined Siemens AG, Mobile Network Division in Munich, Germany in 1991, where He was involved in several EU funded projects and ETSI standardization groups on UMTS and systems beyond 3G. He coordinated several EU and Eureka Celtic funded projects on 3G (FRAMES project), LTE and IMT-Advanced radio interface (WINNER I, II and WINNER+ projects), which developed the basic concepts for future radio standards. Since April 2007 he was with Nokia Solutions and Networks (now Nokia) in Munich Germany, where he was Head of Research Alliances. In addition, he was chairperson of the NetWorld2020 European Technology Platform until December 2016. Werner Mohr was Chair of the Board of the 5G Infrastructure Association in 5G PPP of the EU Commission from its launch until December 2016. He was chair of the “Wireless World Research Forum – WWRF” from its launch in August 2001 up to December 2003. He is co-author of a book on “Third Generation Mobile Communication Systems” (Artech House, Boston – London, 2000) a book on “Radio Technologies and Concepts for IMT-Advanced” (John Wiley & Sons, Chichester, 2009) and a book “Mobile and Wireless Communications for IMT-Advanced and Beyond” (John Wiley & Sons, Chichester, UK, 2011).

Dr. Werner Mohr is member of VDE (Germany) and IEEE. He was member of the board of ITG in VDE from 2006 to 2014. In 1990 he received the ITG Award and in December 2016 the IEEE Communications Society Award for Public Service in the Field of Telecommunications, in November 2018 the VDE ITG Fellowship 2018 and in May 2019 the WWRF Fellowship. In March 2021 he retired from Nokia and is now active as consultant (mainly for 6G Infrastructure Association in the Smart Networks and Services Joint Undertaking of the European Commission) and lecturer at Technical University of Berlin, Germany. His research interests are 6G and radio systems.

Adversarial Machine-Learning-Enabled Anonymization of OpenWiFi Data

Sambita Kuili¹, Kareem Dabbour¹, Irtiza Hasan¹, Andrea Herscovich¹, Burak Kantarci^{1,}, Marcel Chenier² and Melike Erol-Kantarci¹*

Abstract: Data privacy and protection through anonymization is a critical issue for network operators or data owners before it is forwarded for other possible use of data. With the adoption of Artificial Intelligence (AI), data anonymization augments the likelihood of covering up necessary sensitive information; preventing data leakage and information loss. OpenWiFi networks are vulnerable to any adversary who is trying to gain access or knowledge on traffic regardless of the knowledge possessed by data owners. The odds for discovery of actual traffic information is addressed by applied conditional tabular generative adversarial network (CTGAN). CTGAN yields synthetic data; which disguises as actual data but fostering hidden acute information of actual data. In this paper, the similarity assessment of synthetic with actual data is showcased in terms of clustering algorithms followed by a comparison of performance for unsupervised cluster validation metrics. A well-known algorithm, K-means outperforms other algorithms in terms of similarity assessment of synthetic data over real data while achieving nearest scores 0.634, 23714.57, and 0.598 as Silhouette, Calinski and Harabasz and Davies Bouldin metric respectively. On exploiting a comparative analysis in validation scores among several algorithms, K-means forms the epitome of unsupervised clustering algorithms ensuring explicit usage of synthetic data at the same time a replacement for real data. Hence, the experimental results aim to show the viability of using CTGAN-generated synthetic data in lieu of publishing anonymized data to be utilized in various applications.

Keywords: Anonymization, clustering techniques, cluster validation, generative CTGAN.

1. Introduction

WiFi-enabled services have been booming with the advent of widespread adoption of wirelessly connected devices. Simultaneously, a potential vulnerability persists in ensuring the privacy of

usage information. Previous studies employed data privacy techniques such as k-anonymity [1], l-diversity [2], t-closeness [3] and differential privacy [4]. However, those techniques can be complemented by considering data correlation, which possesses utmost significance while exploiting big data [5]. The heterogeneous characteristics underlying in OpenWiFi data possess identical traits of Big Data. Hence, adoption of novel machine learning techniques can eventually enable data privacy through anonymization.

Data anonymization is an essential task prior to making the data public for numerous domain applications. Apart from assuring privacy, it is also challenging to develop anonymized traces, which will resemble the original data. Therefore, retaining the resemblance of original in anonymized data by incorporating abundant noise alters the variance among features of data. This in turn raises issue on usability of data; which is where a trade-off exists while ensuring both data privacy and utility simultaneously on published data [6]. Previously, anonymization has been carried out mostly on healthcare-related domain pertaining to privacy for medical records of patients [7]. To the best of our knowledge, this paper paves the way for anonymization of WiFi usage streams for the first time.

A collaboration of unsupervised and generative adversarial network (GAN) is considered a holistic approach to initiate privacy preservation. Several clustering algorithms, namely K-means, Density Based Spatial Clustering of Application with Noise (DBSCAN) [8], Gaussian Hidden Markov Model [9] and Agglomerative are employed to analyze the actual pattern distribution of the usage traffic. While examining real-time data, there are no prior information on labels corresponding to each stream of data per timestamps. Therefore, employing clustering techniques provides useful insights to comprehend the behavioural patterns among samples in a multivariate data set. On the other hand, adopting a deep neural generative model GAN [10, 11] harbors generation of synthetic samples.

Commonly GAN is quite well-known in computer vision [12, 13]. There are other robust architectures for GANs which can yield remarkable performance on tabular data sets or non-image data sets. In light of these, this work aims to generate synthetic samples by leveraging a conditional tabular GAN (CTGAN) [14]; which is a modified version of architecture to traditional tabular GAN (TGAN) [15]. Hence, leveraging both clustering algorithms and computational power of CTGAN, we narrow down our work to anonymization by testing and validating in terms of quality of clusters. Therefore with support of cluster metrics it becomes easier to validate the performance of CTGAN; which aids us in assessing the quality of generated synthetic

¹School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, Canada

²NetExperience Inc., Ottawa, Canada

E-mail: skuil016@uottawa.ca; kdabb095@uottawa.ca; ihasa074@uottawa.ca; ahers063@uottawa.ca; burak.kantarci@uottawa.ca; marcel@netexperience.com; melike.erolkantarci@uottawa.ca

*Corresponding Author

Manuscript received 20 July 2023, accepted 02 January 2024, and ready for publication 21 March 2024.

© 2024 River Publishers

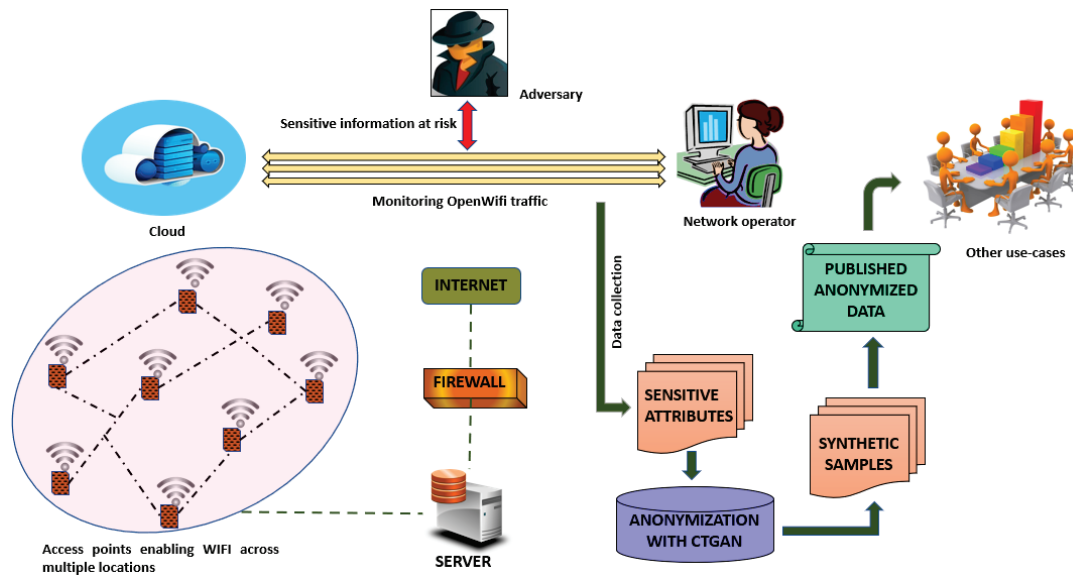


Figure 1.
Anonymization of OpenWiFi network traffic data.

samples. At the same time, it is crucial to test the amount of distortion being carried out in original samples. Simultaneously, we want to ensure that the synthetic samples resembles statistical properties of the original data. Figure 1 showcases topology for demonstration of data anonymization.

Key contributions of this paper are listed below with the ultimate goal to successfully build data anonymization while maintaining data utility and data privacy:

- The initiation of anonymization is being undergone by leveraging heuristic clustering algorithms: K-means, DBSCAN, Gaussian Hidden Markov Model (GHMM) and Agglomerative on a normalized data to showcase the potential cluster labels from the standpoint of distance, density and probability function.
- Leveraging unsupervised algorithms to present the formation of cluster labels and considering clustering membership information as a discrete variable and one of the pivotal parameters to train CTGAN.
- In addition, production of synthetic samples by training CTGAN on the real normalized samples and adopting effective measures to validate similarity performance of synthetic over real samples in terms of cluster metrics, for instance, Silhouette, Calinski and Harabasz (CH) and Davies-Bouldin (DB) scores.

The aforementioned state-of-the-art aligns to aim for high-quality data anonymization technique while advocating for the reliability and trustworthiness of using synthetic data from the standpoint of best-unsupervised algorithm and a number of cluster validation scores regardless of the distorted distribution of data. Over the course of this study, meticulous background information is provided in Section II highlighting the importance of data privacy algorithms and details of use-cases for generative adversarial networks (GANs) in terms of the generation of synthetic data. Moreover, Section III discusses the coherency of the data anonymization undertaken gradually on a multivariate time series

OpenWiFi data. Furthermore, an intensive evaluation showcases the significant performance comparison of adopted unsupervised clustering algorithms with regard to several validation metrics in Section IV, evaluating the quality of anonymized data in comparison to that of the real data. Finally, Section V provides the concluding remarks summarizing the research undergone in this study.

2. Related Work

Data anonymization is defined as the strategy of encrypting sensitive information specifically personally identifiable data in a dataset. In addition, the goal is to assure privacy protection of any information belonging to an attribute in the data by preventing information leakage when being exploited by others performing numerous use-cases. Among various methods of anonymization, for instance, generalization, masking, suppression, perturbation, and usage of synthetic data, the latter is given the major importance by advocating a machine learning model such as a Conditional Tabular Generative Adversarial Network (CTGAN). Consequently, prominent divergence infused on the real data reveals statistical alteration of the data while ensuring commonalities between synthetic and real data.

2.1. Data Privacy Algorithms

Previously data privacy preserving methods such as k -anonymity, l -diversity, t -closeness and δ -disclosure are leveraged to initiate data privacy by compromising certain amendments to the attributes present in data set. Before exploring the process of those methods, it is crucial to understand following key terminologies identifiers, quasi-identifiers (QIDs) and sensitive attributes [16]. Generally, key identifiers signify those attributes with a unique code or

number; which are lacking by default in our data. However, quasi-identifier signifies attributes with discrete value denoting multiple parameters; such as MAC address, location IDs, equipment IDs, and timestamps in our work. Lastly, sensitive attributes are those consisting of other attributes which can neither be termed as identifiers or quasi-identifiers (QIDs). Sweeney [17] proposed an algorithm k-anonymity by elaborating on data; associated with person specific information to ensure privacy of identity and simultaneously considering the issue of re-identification attack into account. In fact, this algorithm works suitably well for selective attributes. Therefore, there is always a way for launching an adversarial attack using other attributes which are not being anonymized using k-anonymity. k-anonymity performs poorly in term of guaranteeing privacy due to presence of unequivocal homogeneity and background knowledge attacks. This is exactly when l-diversity plays a humongous role in ensuring privacy. On the contrary, the adversary seems to develop a thorough knowledge of background distribution of data even after relying on l-diversity algorithm. In order to avoid this consequence, Li et al. [2] introduced a novel algorithm t-closeness, which aligns the distribution of sensitive attributes with other attributes and therefore diminishing plausible background knowledge. Differential privacy is another well-known technique which shows improvement while incorporating privacy measures related to health care domains. Jain et al. [4] discusses on the amount of distortion incorporated by database into the data; which is determined in maintaining privacy and consequently can be found useful by data analysts.

2.2. Conditional Tabular Generative Adversarial Network

The above mentioned techniques still lack non trivial knowledge to convince data privacy for big data. As discussed earlier, OpenWiFi follows 5th Vs of big data and continual deployment of diverse data streams in high volume, therefore, those techniques are not sufficient to ensure data anonymization. Generative adversarial networks [11] are receiving notability lately for creation of fake synthetic samples and are employed for anonymization application by researchers. The architecture of this generative model comprises of two components: generator and discriminator. Those components can be designed by adopting different types of neural networks based on the applicability. Due to the complex architecture of GANs, classifying a sample into fake or original simply eases the entire process. This is the adversarial minmax game where the terminology of GANs turns up. The idea behind minmax game demonstrates that generator will minimize its performance of generating new synthetic training samples, on the contrary, discriminator will aim for maximizing its performance in classifying the generated and original samples into accurate scalar label.

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{data}(x)} [\log(D(x))] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))] \quad (1)$$

$p_z(z)$ is input noise distribution variable representing generator's distribution p_g over data samples x . A data mapping $G(z; \theta_g)$ denotes differential function G with parameters θ_g . Similarly, there is another differential function D with θ_d as parameters, which represents a mapping space for discriminator $D(x; \theta_d)$. In addition,

$D(x)$ signifies likelihood of x coming from real samples and not from distribution p_g . While training GANs, the motive is to train discriminator in assigning accurate labels for samples coming from each distribution mapping space. At the same time, train the generator to minimize the variable $\log(1 - D(G(z)))$. More clarification about the variables referred to (1) can be understood in [11].

There are other variant of generative models which are exploited for use case of generation of synthetic samples. Park et al. [16] introduce the adoption of tableGAN; which includes the architecture of deep convolutional generative adversarial networks (DCGAN) [12]. The main focus is being shown on generating relational synthetic tables based on existence of data attributes associated with convolutional neural networks [16].

Another interesting study undergone by Hajihassani et al. [18] which highlights their major contribution on anonymizing time series sensor data obtained from IoT devices with the help of Variational Autoencoder (VAE). Additionally, their work shares knowledge on understanding the anonymization into two aspects: algorithmic and systemic solutions. Considering algorithmic solutions involving machine learning, anonymization is being previously explored on the factors of security and protection with the assistance of deep neural networks and PrivacyNet [19]. Further insights on systemic solutions can be referred on [18] since those are not related with this work. As Xu and Veeramachaneni [15] propose the novel architecture of conditional tabular GAN emphasizing on the marginal analysis of data with an adoption of recurrent neural network, their work also shares other ways to obtaining synthetic data.

Adoption of other GANs, for instance, RCGAN and RCGAN [20] are popular for generation of time series data. McCoy in [21] share the importance of CTGAN while training recommender systems which are useful to preserve user's privacy. Furthermore, this work also focuses on the reduction of actual raw training samples and in fact, expect an augmentation of actual samples in disguise of synthetic samples. Therefore, reliance on large retrieval of raw data is significantly reduced while still ensuring user privacy.

3. Methodology

Prior to the experiment on anonymization, processing the raw data forms primary step of pipeline. This dataset comprises of 20,000 service records collected from a small network operated over 4 weeks. Moreover, this network is composed of 2 access points averaging 2 connected clients a day. Furthermore, these access points are up and running efficiently, with clients in ON/OFF mode. By setting the desired duration and timestamps, 20K records are collected via REST APIs Cloud Service Portal for OpenWiFi until sufficient information on the active access points and clients connected to them is retrieved. In terms of processing and setting raw data ready for CTGAN, removal of any specific quasi-identifiers; including but not limited to any equipment IDs, location IDs, or customer IDs are implemented. In addition, scaling raw data samples address concerns with mode-specific normalization of CTGAN. Thus, modelling a CTGAN builds on existing numerical or sensitive attributes that contain at most 15,000 traffic records forming a multivariate time series data set.

Moreover, collected data does not contain any prior information on labels. Therefore, adoption of unsupervised learning

blends perfectly to comprehend underlying behavioural distribution of sensitive attributes. Those distributions are analyzed over clustering techniques such as K-means, Density Based Spatial Clustering of Application with Noise (DBSCAN) [22], Gaussian Hidden Markov Model (GHMM) [9]; establishing three different outlooks with respect to distance, density and probability factors. Moreover, an aggregation of clusters is visualized by selecting Agglomerative clustering algorithm. Quality of cluster validation for each unsupervised algorithm is assessed via Silhouette, CH and DB scores. At each stage of execution, potential clusters are demonstrated via Principal Component Analysis (PCA)-based dimensional reduction for more coherent understanding of original observable samples as well as synthetic samples.

Our processed multivariate data set consists of 15 feature variables with (11,900+) instances. While selection of optimal number of clusters for k-means clustering is challenging, setting a range from 2 to 10 is chosen. A trial and error execution is implemented to determine a best optimal number of cluster out of the arbitrary selection of the aforementioned range. To obtain the optimal value k , a cluster validation metric Silhouette score is determined for each value k . Consequently, the value k for k-means turns out to be 2 with the highest Silhouette score ensuring grouping of observations based on distance heuristic with the focus on statistical similarity in data distribution of a multivariate dataset. Furthermore, the Silhouette score is an unsupervised clustering metric which quantifies the quality of the clustering of samples into the chosen optimal number of clusters i.e. k . The higher the Silhouette score (with a maximum score of 1) indicates a higher quality of grouping within each cluster and sufficient separability between them. On the other hand, estimating optimal parametric values for DBSCAN (ϵ and minPts) undergoes an exhaustive search approach highlighting highest Silhouette score [22]. The parameters ϵ and minPts are data dependant where ϵ stands for the radius to be followed with neighboring points from a particular data point in a cluster. In addition, minPts signifies the minimum neighboring data points required to be at a distance ϵ forming a dense distribution of data in a cluster. DBSCAN further identifies some samples as outliers which is bound to occur on high dimensional heterogeneous real data.

Therefore, minimum neighbouring points (minPts) 5 and a distance threshold ϵ 0.038 are selected. The number of states obtained for GHMM is domain-dependent. Kullback Leibler (KL) divergence has been used to obtain empirical probability distribution functions on HMM models with states 2, 3, and 4 on 1000 real RSSI WiFi signal samples [23]. Since there is a small difference between the KL divergence when using 2, 3, or 4 hidden states [24], a GHMM with 3 states is used to model this time series data. Additionally, in order to obtain a more cohesive view of the data, using agglomerative clustering with Ward linkage is also carried out in this work. Generally, agglomerative clustering follows an iterative hierarchical process starting from a single data point as a cluster and eventually merges into another cluster by identifying the closest pair of clusters.

4. Experimental Results

Tuning of parameters to achieve the highest Silhouette score enables analysis of the points exhibiting similar statistical property grouped into clusters. Initially, the original behaviour of real

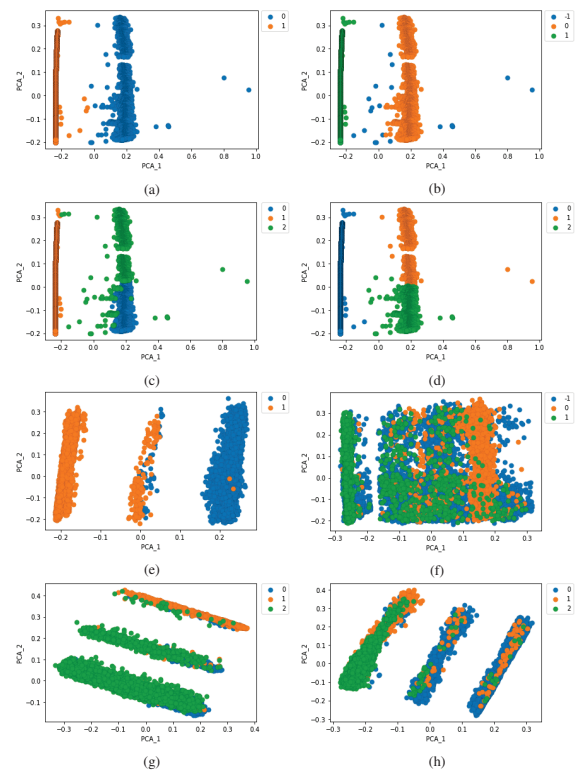


Figure 2.

Clustering visualization of real: (a–d) and synthetic data: (e–h).

data set right after preprocessing is learnt through clustering techniques as demonstrated in Figure 2(a–d); representing K-means, DBSCAN, GHMM and agglomerative algorithms respectively. On examining the Silhouette scores obtained for each technique, it is observed K-means acquires the highest Silhouette score with optimal cluster value 2 (k). Similarly, optimal cluster values for DBSCAN, GHMM and agglomerative obtains $k = 3$ respectively. At the same time, training a complex deep neural network architecture like CTGAN takes longer than expected while searching for best parameters suited for trained model. It is also challenging to validate the performance of synthetic samples generated by CTGAN. A well trained CTGAN is analyzed by understanding the performance loss convergence corresponding to two neural networks: generator and discriminator respectively as shown in Figure 3.

Most common problems which thrive while training GANs are mode collapse and loss convergence. This incorporates adversarial game between two neural networks where either there is a decrement in performance of generator loss and simultaneously an increment on discriminator loss. In addition, initial performance on acquiring the cluster validation scores for real data is analyzed as in Table 1. Furthermore, CTGAN is being trained to generate synthetic samples on being conditioned to discrete column of a data set; which contains cluster labels from each clustering algorithms at a time. The main motive is not limited to simply generate synthetic samples by exploiting generative model; but to check whether synthetic data can be relied upon as compared to real data. In order to assess the quality of synthetic data, clustering

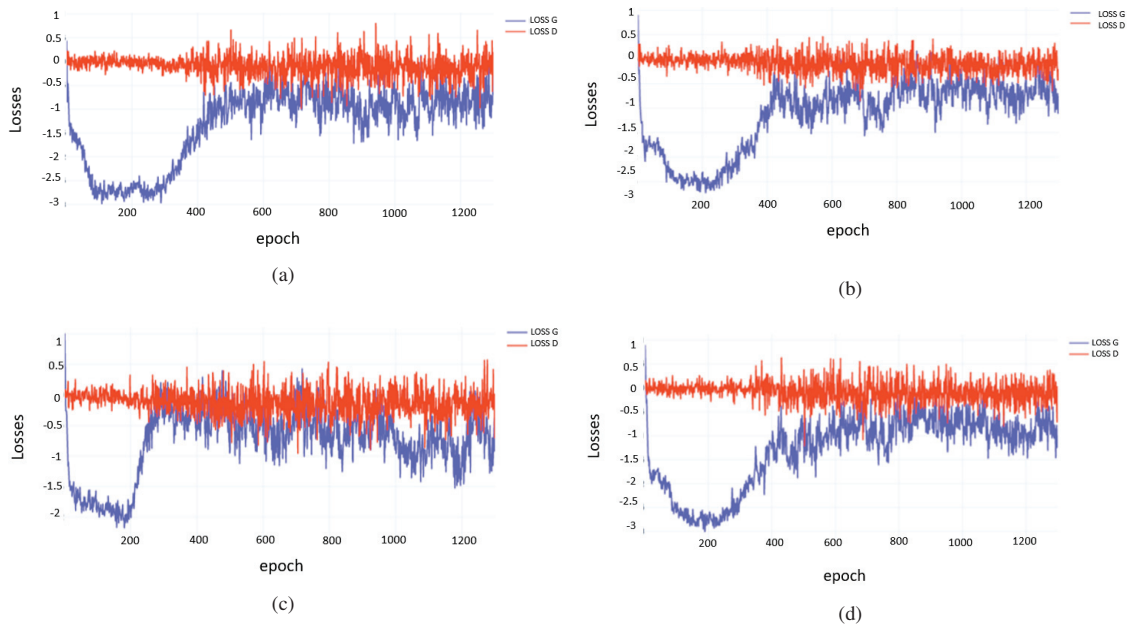


Figure 3. Convergence of loss curves for CTGAN.

Table 1. Cluster validation scores of unsupervised algorithms on real data

Unsupervised Algorithms	Real Data		
	Silhouette	Calinski Harabasz	Davies Bouldin
<i>K-means</i>	0.642	23342.92	0.599
<i>DBSCAN</i>	0.552	11659.34	5.559
<i>GHMM</i>	0.629	24788.87	0.529
<i>Agglomerative</i>	0.635	26109.21	0.501

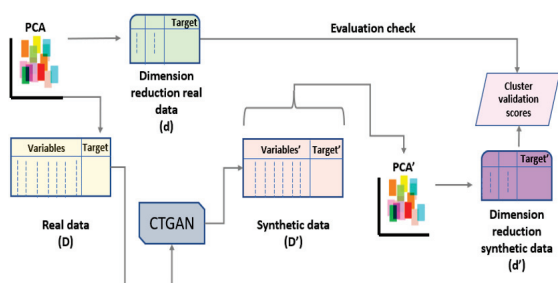


Figure 4. Initiation of data anonymization leveraging clustering algorithms.

algorithms are employed which reflect the performance score for each synthetic and real data. As a result, more similarity or close the value for each metric, higher is the indication of replacing real with synthetic data. The initial real data (D) in Figure 4 is set with target attribute; which consists of labels obtained from K-means.

Moreover, we retain dimensionality reduction of real data (D) through PCA obtaining low dimensional data (d); which also

mitigates the problem collinearity among features. The core task for training generator of CTGAN requires conditional vectors. However, target attribute with labels 0 and 1 of K-means is considered as discrete variable while continuous variables as rest of the attributes or features existing on real data. Therefore, labels in discrete variable are utilized as conditional vectors to generate synthetic samples. A synthetic data (D') is obtained after training CTGAN while tuning for convergence with specific hyper-parameters demonstrated in the form of a loss curve in Figure 3(a).

This synthetic data (D') is further visualized with the help of PCA; forming another low dimensional data set (d'). Accordingly, evaluation of cluster metrics on (d') are computed. This exact method is carried out followed by altering the target variable subsequently with labels of DBSCAN, GHMM and agglomerative clustering algorithms, and correspondingly scores are calculated. Consequently, loss curves demonstrated in Figure 3(b), (c) and (d) are also checked for convergence after setting the labels for DBSCAN, GHMM and agglomerative algorithms as conditional vectors for generator. The cluster validation scores computed on synthetic data can be understood from Table 2.

While comparing the validation scores of K-means on both Tables 1 and 2, there is decent clarity which shows resemblance in scores for all three metrics. However, there is significant amount of deviation in scores for DBSCAN with respect to Silhouette, CH and DB. Similar to the Silhouette score, CH and DB scores aim to quantify quality of cluster formation. For instance, a higher CH score signifies better quality of clustering from the perspective of factors inclusive of within-cluster variance and between-cluster variance. Moreover, DB is quite the opposite of Silhouette and CH scores as it measures the average computation of inter-cluster distance and within-cluster dispersion of a cluster. More importantly, a lower DB score denotes greater cluster formation

Table 2.

Cluster validation scores of unsupervised algorithms on synthetic data			
Unsupervised Algorithms	Silhouette	Synthetic Data	
		Calinski Harabasz	Davies Bouldin
<i>K-means</i>	0.634	23714.57	0.598
<i>DBSCAN</i>	0.171	3742.45	1.836
<i>GHMM</i>	0.409	8865.44	0.770
<i>Agglomerative</i>	0.507	12188.80	0.769

quality. While comparing scores for GHMM, CH score reflects high deviation as compared to Silhouette and DB. On the contrary, agglomerative achieves reasonable scores for all metrics of both real and synthetic data. In this study, we intend to look for similarity between the scores for both real and synthetic data. For instance, K-Means preserves the statistical properties in synthetic data even after incorporating noise or conditional vectors by CTGAN into real data.

K-means is compatible while analyzing the performance of cluster validation scores from the perspective of data anonymization [25]. Additionally, a baseline algorithm K-Means on real data is selected to compare the cluster validation scores with each algorithm from synthetic data. It is observed that K-means exhibits resemblance in scores while comparing cluster validation metrics of real and synthetic data as shown in Tables 1 and 2, respectively. This coherently indicates that there is no significant statistical implication in the application context of data anonymization. However, DBSCAN, GHMM, and agglomerative algorithms result in the variation of both low Silhouette and Calinski Harabasz scores, and higher Davies Bouldin scores. In the context of application data anonymization and deviation in scores, it is understandable that DBSCAN is sensitive to outliers which results in performance downgrade under synthetic data when compared with the baseline algorithm. Similarly, the poor performance of GHMM and agglomerative algorithms highlights presence of noise in the synthetic data and involvement of model complexity with parameter initialization based on the data distribution. To comprehend deviation, we establish a visual analysis forming clusters of synthetic data for each algorithm shown in Fig. 2 (e-h); denoting K-means, DBSCAN, GHMM and agglomerative algorithms. On examine those figures, we encounter a skewed distribution of cluster labels as compared to Figure 2(a-d); which is rational to occur on a synthetic data. The reason of such distortion depends on the conditional vectors provided to train the generator of CTGAN. The resemblances in validation scores as seen on both real and synthetic data also provide assurance of preserving statistical properties similar to samples of real data; which is an essential factor as discussed in the earlier sections. Furthermore, this ensures that the synthetic data can be relied upon and effectively utilized for numerous applications.

5. Open Issues, Challenges, and Opportunities

The usage of multivariate time series OpenWiFi data where this study aims to adopt data anonymization and forecast the generation of the skewed distribution of synthetic samples already lack prior information of ground truth. Simultaneously, the analysis

focuses on mimicking cluster labels / ground truth based on the data distribution present in real and synthetic data setting an example to comprehend the problem only from the standpoint of the statistical context. However, other ways to augment the trustworthiness of synthetic data include privacy risk assessment, evaluating downstream tasks, and incorporating the review by domain experts who are knowledgeable about real data and will provide valuable feedback on synthetic data.

Every new data anonymization technique needs to cope with several challenges. Re-identifying instances of anonymized data, and safeguarding data utility is just one to mention. The latter can be understood if there is a loss of pivotal information in the real data, which can pave the way for an effective decision capability while exploiting the anonymized data. In addition, the dynamicity of multivariate time series data needs to be taken into account since it consists of temporal dependency and may impact anonymized data. Moreover, the scalability of the proposed framework is also of paramount importance which needs to be addressed while dealing with large volume of data. However, consistent performance of data anonymization is expected while safeguarding the data utility regardless of exploiting big data. Furthermore, scalability may also be studied from the use of CTGAN, which expects a large volume of data to train itself, therefore would ensure better prediction ability by studying the distinctions between synthetic and real data.

The idea of adversarial machine learning-enabled anonymization imparts the efficacy of data sharing across different domains such as ethical use of data by adopting a measure of privacy protection of sensitive information in the data.

6. Conclusion

Data anonymization is a crucial task for operators to augment privacy of information; which furthermore meets required user demands. This work summarizes the generation of synthetic samples by deploying conditional tabular GAN to replace real data instances with synthetic samples. Incorporation of clustering mechanism highlighted heterogeneous distribution of OpenWiFi traffic on the grounds of underlying factors such as distance, density and probability of samples existing in both real and synthetic data. In addition, computation of cluster validation metrics by well-known Silhouette, Calinski and Harabasz and Davies-Bouldin scores have been undergone to comprehend reliability of synthetic data. A wise similarity comparison of those scores has been checked while assessing synthetic with real data. Moreover, two dimensional visualization at every stage of implementation enables us to understand the original and the skewed distribution of synthetic data. This work has particularly focused on resemblance in the unsupervised validation metrics, which invokes the preservation of statistically correlated properties among samples. By leveraging this knowledge, extension of this work includes incorporation of other machine learning and privacy-based metrics to improve data anonymization.

Acknowledgment

This work was supported in part by Mathematics of Information Technology and Complex Systems (MITACS) Accelerate Program under Project Number IT24702.

References

- [1] A. Machanavajjhala, D. Kifer, J. Gehrke, and M. Venkatasubramanian, "l-diversity: Privacy beyond k-anonymity," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 1, no. 1, pp. 3–es, 2007.
- [2] N. Li, T. Li, and S. Venkatasubramanian, "t-closeness: Privacy beyond k-anonymity and l-diversity," in *2007 IEEE 23rd international conference on data engineering*. IEEE, 2007, pp. 106–115.
- [3] X. Yang, T. Wang, X. Ren, and W. Yu, "Survey on improving data utility in differentially private sequential data publishing," *IEEE Transactions on Big Data*, vol. 7, no. 4, pp. 729–749, 2017.
- [4] P. Jain, M. Gyanchandani, and N. Khare, "Big data privacy: a technological perspective and review," *Journal of Big Data*, vol. 3, no. 1, pp. 1–25, 2016.
- [5] S. Biswas, N. Khare, P. Agrawal, and P. Jain, "Machine learning concepts for correlated big data privacy," *Journal of Big Data*, vol. 8, no. 1, pp. 1–32, 2021.
- [6] W. Liao, J. He, S. Zhu, C. Chen, and X. Guan, "On the tradeoff between data-privacy and utility for data publishing," in *2018 IEEE 24th International Conference on Parallel and Distributed Systems (ICPADS)*. IEEE, 2018, pp. 779–786.
- [7] J. Coutinho-Almeida, P. P. Rodrigues, and R. J. Cruz-Correia, "Gans for tabular healthcare data generation: A review on utility and privacy," in *International Conference on Discovery Science*. Springer, 2021, pp. 282–291.
- [8] E. Schubert, J. Sander, M. Ester, H. P. Kriegel, and X. Xu, "Dbscan revisited, revisited: why and how you should (still) use dbscan," *ACM Transactions on Database Systems (TODS)*, vol. 42, no. 3, pp. 1–21, 2017.
- [9] L. R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.
- [10] J. Yoon, L. N. Drumright, and M. Van Der Schaar, "Anonymization through data synthesis using generative adversarial networks (adsgan)," *IEEE journal of biomedical and health informatics*, vol. 24, no. 8, pp. 2378–2388, 2020.
- [11] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," *Advances in neural information processing systems*, vol. 27, 2014.
- [12] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," *CoRR*, vol. abs/1511.06434, 2015. [Online]. Available: <https://api.semanticscholar.org/CorpusID:11758569>.
- [13] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2223–2232.
- [14] L. Xu, M. Skoularidou, A. Cuesta-Infante, and K. Veeramachaneni, "Modeling tabular data using conditional gan," *Advances in Neural Information Processing Systems*, vol. 32, 2019.
- [15] L. Xu and K. Veeramachaneni, "Synthesizing tabular data using generative adversarial networks," *arXiv preprint arXiv:1811.11264*, 2018.
- [16] N. Park, M. Mohammadi, K. Gorde, S. Sajodia, H. Park, and Y. Kim, "Data synthesis based on generative adversarial networks," *Proc. VLDB Endow.*, vol. 11, pp. 1071–1083, 2018. [Online]. Available: <https://api.semanticscholar.org/CorpusID:47017667>.
- [17] L. Sweeney, "k-anonymity: A model for protecting privacy," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 10, no. 05, pp. 557–570, 2002.
- [18] O. Hajihassani, O. Ardakanian, and H. Khazaei, "Anonymizing sensor data on the edge: a representation learning and transformation approach," *ACM Transactions on Internet of Things*, vol. 3, no. 1, pp. 1–26, 2021.
- [19] V. Mirjalili, S. Raschka, and A. Ross, "Privacynet: semi-adversarial networks for multi-attribute face privacy," *IEEE Transactions on Image Processing*, vol. 29, pp. 9400–9412, 2020.
- [20] C. Esteban, S. L. Hyland, and G. Rätsch, "Real-valued (medical) time series generation with recurrent conditional gans," *arXiv preprint arXiv:1706.02633*, 2017.
- [21] S. V. McCoy, "Exploration of user privacy preservation via ctgan data synthesis for deep recommenders."
- [22] S. Kuili, B. Kantarci, M. Erol-Kantarci, M. Chenier, and B. Herscovici, "A holistic machine learning approach to identify performance anomalies in enterprise wifi deployments," in *Big Data IV: Learning, Analytics, and Applications*, vol. 12097. SPIE, 2022, pp. 105–117.
- [23] Ó. Belmonte-Fernández, E. Sansano-Sansano, A. Caballer-Miedes, R. Montoliu, R. García-Vidal, and A. Gascó-Compte, "A generative method for indoor localization using wi-fi fingerprinting," *Sensors*, vol. 21, no. 7, p. 2392, 2021.
- [24] M. Vidyasagar, "Bounds on the kullback-leibler divergence rate between hidden markov models," in *2007 46th IEEE Conference on Decision and Control*. IEEE, 2007, pp. 6160–6165.
- [25] M. E. Ferrão, P. Prata, and P. Fazendeiro, "Utility-driven assessment of anonymized data via clustering," *Scientific Data*, vol. 9, no. 1, p. 456, 2022.

Biographies



Samhita Kuili is a PhD candidate and a researcher in Smart Connected Vehicles Innovation Centre (SCVIC) at the University of Ottawa. Her research direction includes AI-enabled anomaly detection in 5G and local wireless network connection WiFi 6. She collaborated on her research by working alongside an enterprise in the field of OpenWiFi networks. She is working on a 5G wireless network to detect security attacks and anomalies. She is one of the founding members of ACM-Women Student Chapter at University of Ottawa. She is an active student member of the IEEE Communication Society and IEEE Communication Systems Integration and Modeling Technical Committee.



Kareem Dabbour graduated with an honours Bachelor of Science in Computer Science specializing in Data Science from the University of Ottawa. He is a computer science and math

enthusiast that is committed to leveraging the power of both disciplines for innovative problem-solving.



Irtiza Hasan holds a Bachelor of Science degree in Computer Science from the University of Ottawa, specializing in Data Science. He is interested in exploring solutions that prioritize ethical considerations while advancing the development of artificial intelligence at a graduate level.

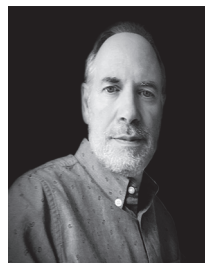


Andrea Herscovich works as a software developer at CoLab software. She has a Bachelor of Science in Computer Science with Honours from the University of Ottawa.



Burak Kantarci (S'05, M'09, SM'12) received the Ph.D degree in computer engineering in 2009. He is a Full Professor and the Founding Director of Smart Connected Vehicles Innovation Centre (SCVIC), and the Founding Director of the Next Generation Communications and Computing Networks (NEXTCON) Research Lab, University of Ottawa. He has coauthored over 250 publications in established journals and conferences, and 15 book chapters. He is well known for his contributions to the quantification of data trustworthiness in mobile crowd-sensing (MCS) systems, and game theoretic incentives to promote user participation in MCS campaigns with

high value data; as well as AI-backed access control, authentication and machine learning-backed intrusion detection solutions in sensing environments. In 2022, he has been awarded a Minister's Award of Excellence in Innovation and Entrepreneurship from Ontario Ministry of Colleges and Universities. He served as the Chair of IEEE Communications Systems Integration and Modeling Technical Committee, and has served as the Technical Program Co-Chair/Symposium Co-chair of more than twenty international conferences/symposia/workshops, including IEEE Global Communications Conference (GLOBECOM) – Communications Systems QoS, Reliability and Modeling (CQRM) symposium. In 2021, he has been elected as the new Secretary of IEEE Social Networks Technical Committee. He is an Editor of the IEEE Communications Surveys & Tutorials, IEEE Internet of Things Journal, Vehicular Communications (Elsevier), and an Associate Editor for IEEE Networking Letters, and Journal of Cybersecurity and Privacy. He is Editor for several IEEE and Elsevier journals. He was an ACM Distinguished Speaker in 2019–2021, currently an ACM Senior Member. He is IEEE Systems Council and IEEE Communications Society Distinguished Lecturer.



Marcel Chenier is CTO and co-founder of NetExperience where he brings three decades of technology leadership experience in the field of public and enterprise wireline, wireless and cloud networking. In his role, he is responsible for technology direction and ensuring product feedback from Net Experience's Service Provider customers.

Prior to NetExperience, Marcel took the role of CTO and VP of Engineering for KodaCloud building a Cloud Network control solution for MSP WLAN networks. Previous to KodaCloud, Marcel lead product architecture for the WLAN portfolio and Small Cell Plug&Play Solutions at Ericsson. Prior to Ericsson, Marcel was Vice-President of engineering at BelAir Networks, a WLAN carrier-class small-cell equipment provided which was acquired by Ericsson. Prior to BelAir, Marcel was head of engineering at Zhone Technologies Canada from early company stage to IPO responsible for voice and video packet products. As VP of R&D for Premisys, a smart DLC equipment provider, Marcel was part of the leadership team that lead the acquisition by Zhone. Marcel started his career with Nortel Networks working on silicon, hardware and software research and development and as an R&D Executive for the Signaling System Product Line. Marcel is a co-inventor on several patents in the area of wireless networking architecture.

Marcel holds a Bachelor of Engineering from the University of Ottawa. He was a member of the Intel Technical Advisory board from 2011 to 2014.



Melike Erol-Kantarci is Chief Cloud RAN AI/ML Data Scientist at Ericsson and Canada Research Chair in AI-enabled Next-Generation Wireless Networks and Full Professor at the School of Electrical Engineering and Computer Science at the University of Ottawa. She is the founding director of the Networked Systems and Communications Research (NETCORE) laboratory. She has received numerous awards and recognitions. Dr. Erol-Kantarci is the co-editor of three books on smart grids, smart cities and intelligent transportation. She has over 200 peer-reviewed publications. She has delivered 70+ keynotes, plenary talks and tutorials around the globe. Dr. Erol-Kantarci is on the editorial board of the IEEE Transactions on Communications, IEEE Transactions on Cognitive Communications and Networking and IEEE Networking Letters. She has acted as the general chair and technical program chair for many international conferences and workshops. Her main research interests are AI-enabled wireless networks, 5G and 6G wireless communications, smart grid and Internet of Things. Dr. Erol-Kantarci is an IEEE ComSoc Distinguished Lecturer, IEEE Senior member and ACM Senior Member.

On the Impact of CDL and TDL Augmentation for RF Fingerprinting Under Impaired Channels

Omer Melih Gul¹, Michel Kulhandjian¹, Burak Kantarci^{1,}, Claude D'Amours¹, Azzedine Touazi² and Cliff Ellement²*

Abstract: Cyber-physical systems have recently been used in several areas (such as connected and autonomous vehicles) due to their high maneuverability. On the other hand, they are susceptible to cyber-attacks. Radio frequency (RF) fingerprinting emerges as a promising approach. This work aims to analyze the impact of decoupling tapped delay line and clustered delay line (TDL+CDL) augmentation-driven deep learning (DL) on transmitter-specific fingerprints to discriminate malicious users from legitimate ones. This work also considers 5G-only-CDL, WiFi-only-TDL augmentation approaches. RF fingerprinting models are sensitive to changing channels and environmental conditions. For this reason, they should be considered during the deployment of a DL model. Data acquisition can be another option. Nonetheless, gathering samples under various conditions for a train set formation may be quite hard. Consequently, data acquisition may not be feasible. This work uses a dataset that includes 5G, 4G, and WiFi samples, and it empowers a CDL+TDL based augmentation technique in order to boost the learning performance of the DL model. Numerical results show that CDL+TDL, 5G-only-CDL, and WiFi-only-TDL augmentation approaches achieve 87.59%, 81.63%, 79.21% accuracy on unobserved data while TDL/CDL augmentation technique and no augmentation approach result in 77.81% and 74.84% accuracy on unobserved data, respectively.

Keywords: Data augmentation, deep learning, RF fingerprinting, secure design, uncrewed aerial vehicles.

1. Introduction

Cyber-physical systems which are networked systems with physical and computational capabilities can be applied to many areas. They include wireless communications, search and rescue, crop spraying and monitoring, and agricultural, fire, or wildlife surveillance. Their integration with wireless communication capabilities brings many advantages. Nevertheless, they may face security issues. To name a few for a sample cyber-physical system, an uncrewed aerial

vehicle (UAV) may be susceptible to tampering, masquerading, spoofing, and replay attacks [1, 2].

Considering their various features and characteristics of the systems, various security solutions are suggested at various levels. An intrusion detection system is used to identify suspicious activities in application and network layers. Furthermore, applying message encryption at different levels is of paramount importance to protect data at rest and in transmission. On the basis of this motivation, authentication and encryption-based security solutions for cyber-physical systems have been widely investigated [1]. Recently, Alladi et al. [2] introduce an authentication system based on a physical unclonable function. Bassey et al. [3] discuss the use of authentication codes for an Internet of Things (IoT) device, and by leveraging the Kolmogorov-Smirnov test, it determines the legitimacy of the user of an IoT device.

Despite the existence of encryption and authentication techniques, cyber-physical systems are vulnerable to spoofing attacks that can be coped if system-level security-complemented encryption and authentication-based solutions. For instance, Shi et al. [4] present an anticipatory study via a generative adversarial network (GAN)-based spoofing attacks that aim to bypass the authentication barrier by mimicking legitimate transmitters. To do so, the GAN-based adversarial attack captures channel effects and waveforms. Consequently, it increases the success probability of such an attack. GANs can also be used to safeguard systems against spoofing attacks. Sagduyu et al. [5] exploit the classification outputs of a neural network to deceive adversaries. At lower layers such as in the radio frequency (RF) domain, security by design can be offered to cyber-physical systems.

RF security entails hardware-specific characteristics and imperfections of transmitters. For example, in a connected vehicle network, each vehicle transmitter reveals hardware imperfections during communications via RF transmissions, which is referred to as RF fingerprint of their transmitters. For ultra-reliable low latency communications (URLLC) purposes, RF fingerprinting can suggest security solutions to accelerate security services in comparison to network-layer (L3) solutions. RF fingerprinting can be defined as mining hardware imperfection patterns in manufacturing defaults [6]. DL methods are widely studied and shown to be successful at extracting specific features. Therefore, DL models can be utilized to detect particular RF signatures [7].

Channel impairments affect the performance of DL techniques to identify the RF signatures of a transmitter [8, 9]. It is not feasible to collect data in all circumstances. For this reason, applying augmentation methods for RF fingerprinting is a viable solution to deal with channel impairments during the DL training

¹Department of EECS, University of Ottawa, Ottawa, ON, Canada

²Artificial Intelligence Solutions, ThinkRF, Ottawa, ON, Canada
E-mail: ogul@uottawa.ca; mkulhand@uottawa.ca; burak.kantarci@uottawa.ca; cdamours@uottawa.ca; azzedine.touazi@thinkrf.com; cliff.ellement@thinkrf.com

* Corresponding Author

Manuscript received 20 July 2023, accepted 23 November 2023, and ready for publication 21 March 2024.

© 2024 River Publishers

process. As a special type of augmentation technique, adversarial machine learning can be used prior to training DL classifiers for RF fingerprinting [10].

This study that was initially presented in [23] differs from the existing works with the following points:

- To the best of our knowledge, this is the first work focusing on applying different types of data augmentation on raw time-domain signals in heterogeneous environments with real-world data.
- We decouple the tapped delay line (TDL) and clustered delay line (CDL) augmentation approaches such that 5G data is augmented by CDL whereas WiFi data is augmented by transforming it by TDL (referred to as CDL+TDL hereafter).

This study extends our work in [23] with the following points:

- We propose a 5G-only-CDL augmentation approach where only 5G data is augmented with CDL transform. Thus, we can observe its difference from CDL/TDL augmentation.
- We propose a WiFi-only-TDL augmentation approach where only WiFi data is augmented with TDL transform. Thus, we can observe its difference from CDL/TDL augmentation.

This paper is structured as follows. Related work is reviewed in Section 2. Section 3 gives the system model and problem definition. Section 4 introduces the 5G-only-CDL and WiFi-only-TDL augmentation techniques. Section 5 introduces the decoupled CDL+TDL augmentation technique. Section 6 presents simulation results. Section 8 concludes this work and presents future work.

2. Related Work

This section presents the most relevant literature on RF fingerprinting. The work presented by Cekic et al. primarily focuses on multiple WiFi protocols (IEEE 802.11a and IEEE802.11g) and automatic dependent surveillance-broadcast (ADS-B) protocol [11] to identify two main transmitter characteristics, namely in-phase and quadrature-phase (I/Q) imbalance, and power amplifier nonlinearity. A comparison of additive white Gaussian noise (AWGN) and without augmentation is presented. Shea et al. in [12] report that complex-value networks may not be sufficient for radio signal classification in many real-world applications.

Sankhe et al. [13] leverage raw I/Q samples to characterize static channels, which eliminates the coefficient prediction requirement. On the other hand, complex demodulated symbols are utilized to remove channel effects. By exploiting direct current (DC) offset and I/Q imbalance, a convolutional neural network (CNN) is trained to recognize radio signal signatures under static and varying channel characteristics.

Ozturk et al. [14] present a CNN-based approach that takes spectrograms and time-series images of RF signals as inputs. Transient signals with low signal-to-noise-ratio (SNR) are prone to noise effects; hence, it is not viable to use a time series-based model. As a result of the frequency range consideration by spectrograms, more accurate outputs can be obtained in comparison to the time series-based model.

RF signatures of UAVs are aimed to be identified by Mohanti et al. [15] under a CNN with VGG backbone. To tackle I/Q imbalance, amplitude changes are fed into real and imaginary components. It is worth noting that DL-based fingerprinting mechanisms can recognize UAV radio signals only if the channel conditions do not change across training and test phases. To cope with this phenomenon, the authors propose a processing block to arrange I/Q instances.

Soltanie et al. [16] propose two alternate augmentation methods that build on two finite impulse response filters. The first method is trained under a dataset generated in MATLAB and validated under a DARPA dataset. These methods have been shown to run with raw I/Q instances without quadrature phase shift keying (QPSK) modulation. The high throughput task group (TGn) of wireless local area network (WLAN) channel data samples are augmented with AWGN to mimic more realistic channel behavior. This method can perform well with channels that have not been seen before. Testing of this method is performed by considering the boundaries of single-channel models.

For RF fingerprinting purposes, Reus-Muns et al. [17] built a dataset that comprises WiFi, 5G, and LTE samples and feed raw I/Q samples into a neural network. The focal point of that work is to investigate the accuracy degradation between training and test data acquired on different days with different channel conditions. When the train and test data belong to the same day, the accuracy is sufficient whereas when the test is performed in a different day, the accuracy is remarkably degraded. To remediate this issue, leveraging triplet loss functions is proposed by Reus-Muns et al. [17] so that the accuracy is acceptable even if the training and test data are collected on different days, i.e., with different propagation conditions. Zhao et al. [18] present an analysis of wireless signals for UAV model classification by combining auxiliary classifier generative adversarial network (ACGAN) and Wasserstein generative adversarial network (WGAN) where root mean squared propagation is applied as an accelerator and signals obtained by applying a band-pass filter.

Table 1 presents a comparative list of the relevant works that are closest to the work presented in this paper. It is observed that the majority of the relevant works rely on either synthetic datasets generated in MATLAB or datasets that contain modulated data. Training and testing on different days (i.e., varying

Table 1.

Brief comparison of techniques proposed in close literature				
	Demodulated or Real Data	Different Day Tests	Multiple Waveforms	CDL&TDL Augmentation Analysis
Cekic et al. [11]	no	yes	yes	no
Sankhe et al. [13]	no	yes	no	no
Mohanti et al.[15]	no	yes	no	no
Soltani et al. [16]	no	yes	no	no
Muns et al. [17]	yes	yes	yes	no
Zhao et al. [18]	yes	no	no	no
Comert et al. [19]	yes	yes	yes	no
Gul et al. [23]	yes	yes	yes	yes
This work	yes	yes	yes	yes

channel conditions) is pursued by few studies. Since CDL/TDL augmentation was shown to perform well in the previous work in [19], this work aims to pursue a detailed investigation of decoupling CDL and TDL on 5G and WiFi waveforms, respectively under data-scarce scenarios.

3. System Model and Problem Definition

This section introduces the system model before proceeding with the methodological details. Figure 1 illustrates the system model.

Autonomous vehicles, similar to other cyberphysical systems, are vulnerable to spoofing attacks. Each wireless device operates with slight hardware imperfections that are caused during manufacturing. These hardware imperfections can be recognized on radio signals as their wireless signatures, which are referred to as the RF fingerprint of the device. Alongside the enhanced mobile broadband (eMBB) and massive machine type communications (mMTC), the ultra-reliable low latency communications (URLLC) use case of 5G, requires security by design, for which RF fingerprinting is an ideal candidate.

The problem formulation for RF fingerprinting is as follows: A receiver aims to identify transmitters based on their RF signatures in time-domain raw signal samples.

The objective of this study is to obtain high accuracy under Day-2 data when the radio signal classifier is trained under Day-1 data. Thus, we aim to provide insights into decoupling the channel condition effects and the transmitter impairments. The radio signal instances in the data consist of 5G, WiFi, and LTE waveforms in the POWDER Dataset that were originally presented in [17]. The dataset includes 60 .bin files for Day-1 and 60 .bin files for Day-2 that are equally split between the 5G, 4G/LTE, and WiFi waveforms for both days.

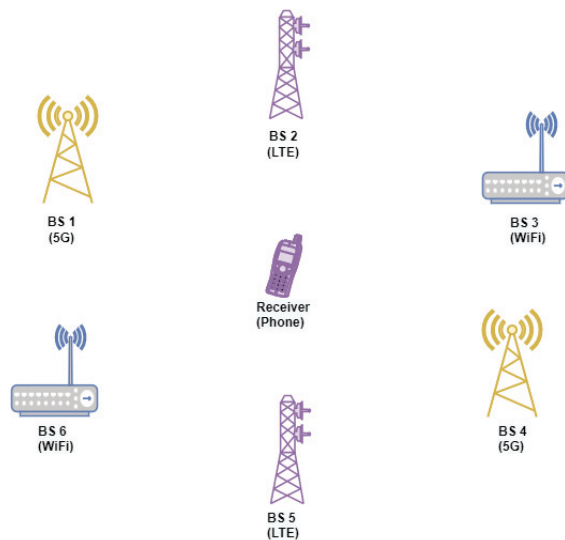


Figure 1.

In this network, base stations B5 and B6 communicates with mobile phone by resembling B2 and B3 (spoofing attack).

4. 5G-only-CDL and WiFi-only-TDL Augmentation Approaches

As previously stated, RF fingerprinting is susceptible to a variety of channel and environmental conditions; as a result, RF fingerprinting requires effective solutions that are robust (in terms of accuracy) in a variety of circumstances. Data augmentation-based solutions are required to develop neural network architectures that take the channel conditions into consideration. In fact, the goal of data augmentation is to develop a method that can replicate radio signal traces in a variety of channel and environmental settings.

In this section, we propose two augmentation approaches where each of CDL and TDL transforms are applied to just one of the communication technologies data (5G, WiFi) instead of applying the same transform to all data. In the first subsection, we first propose a 5G-only-CDL augmentation approach different from CDL/TDL approach. Then, in the second subsection, we propose a WiFi-only-TDL augmentation approach different from CDL/TDL approach.

4.1. 5G-only-CDL Augmentation Approach

In this subsection, we attempt to observe the effect of applying CDL transform on just a subset (5G data) of the dataset instead of applying CDL transform to all data. This provides more insight for us to present the CDL+TDL augmentation approach, which is proposed in [23], in the following section.

We used the MATLAB 5G Toolbox (found in [20]) to apply CDL-transform on only 5G signals for data augmentation to train the model before the unseen data with channel impairments. A minimalist illustration of 5G-only-CDL augmentation is given in Figure 2.

4.2. WiFi-only-TDL Augmentation Approach

In this subsection, we would like to observe the effect of applying TDL transform on just a subset (WiFi data) of the dataset instead of applying TDL transform to all data. This provides more insight for us to present the CDL+TDL augmentation approach, which is proposed in [23], in the following section.

We suggest using the MATLAB 5G Toolbox (found in [20]) to apply TDL-transform on only WiFi signals for data augmentation to train the model before the unseen data with channel impairments.

A minimalist illustration of WiFi-only-TDL augmentation is given in Figure 3.

5. Decoupled CDL and TDL (CDL+TDL) Augmentation Approach

As stated earlier, RF fingerprinting is vulnerable to varying channel and environmental conditions; hence, RF fingerprinting calls for effective solutions that are robust (in terms of accuracy) under these varying conditions. To account for the channel conditions in the design of a neural network architecture, data augmentation-based solutions are needed. Indeed, the rationale for data augmentation is to come up with a technique that can

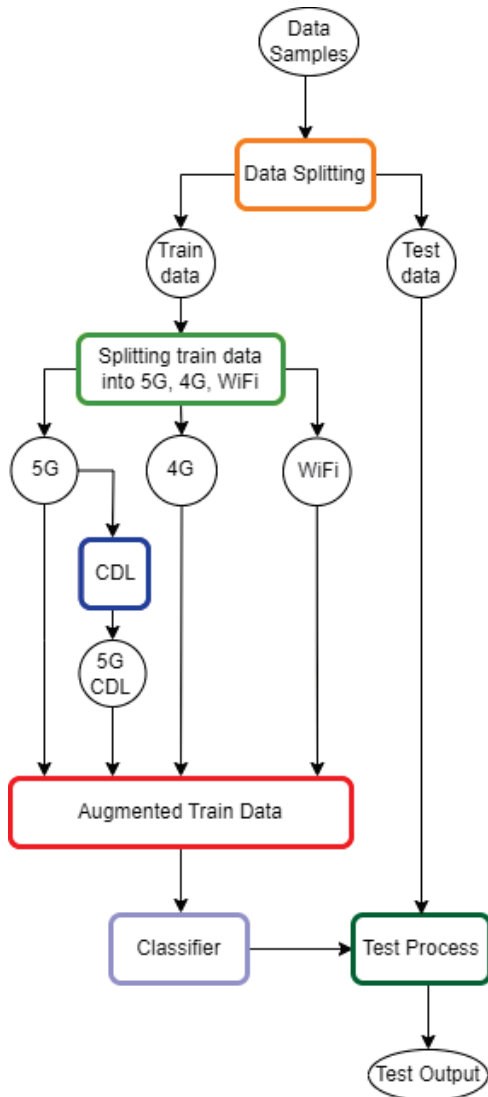


Figure 2.
The proposed 5G-only-CDL augmentation architecture.

mimic radio signal traces under varying channel and environmental conditions.

In order to train the model prior to the unobserved data with channel impairments, we propose to decouple CDL and TDL augmentation via MATLAB 5G Toolbox (available in [20]) for 5G and WiFi signals. All models are designed for full frequency ranges. It may be scaled further for observing the desired root-mean-square (RMS) delay spread. Therefore, special models for line-of-sight (LOS) and non-line-of-sight (NLOS) scenarios can be designed [21]. In addition to the Day-1 data of the dataset, we aim to add more knowledge that represents characteristics of unobserved data on Day-1. The rationale for this is to cope with the channel impairments / varying channel conditions on Day-2. To do so, more information is added to Day-1 data so as to learn and handle channel impairments during the test of Day-2 data.

The previous work in [19] introduces CDL/TDL augmentation which is quite suitable to output augmented data from

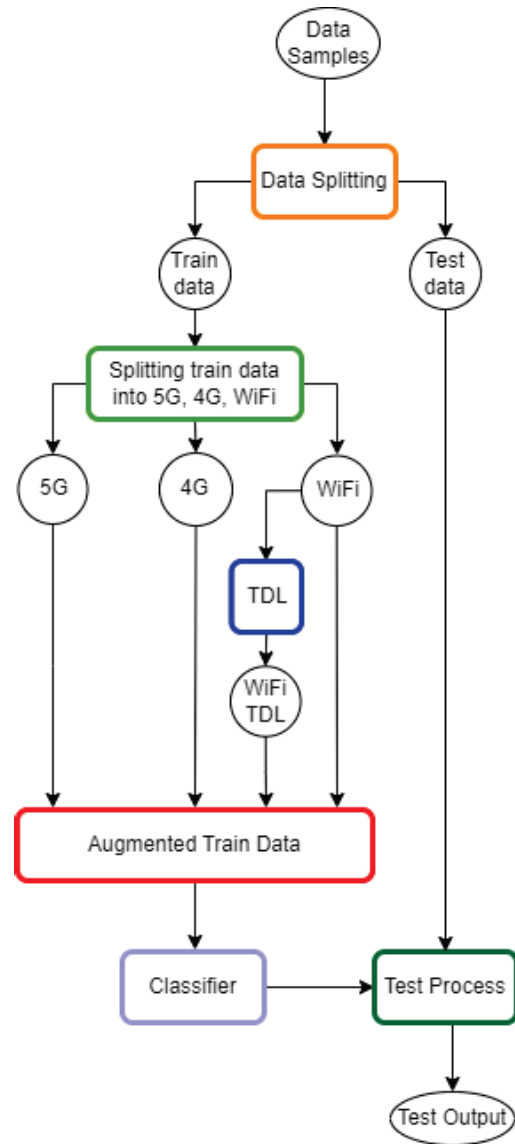


Figure 3.
The proposed WiFi-only-TDL augmentation architecture.

5G samples. Additionally, this paper shows that augmenting WiFi samples with TDL works well with WiFi samples.

This work introduces the decoupled CDL+TDL augmentation technique that augments 5G data samples by transforming them via CDL transformation, and WiFi data samples by transforming them via TDL transformation. Day-1 (train) data is augmented via this augmentation technique, and Day-2 data for testing.

It should be noted that CDL/TDL augmentation technique applies the same transform to all waveform types of 5G, LTE, and WiFi data. In other words, it applies CDL transform to all or TDL transform to all. Both transforms are highly correlated with each other so their performance figures are alike. On the other hand, the CDL+TDL augmentation technique decouples TDL and CDL transforms by applying CDL transform just on 5G-waveform data

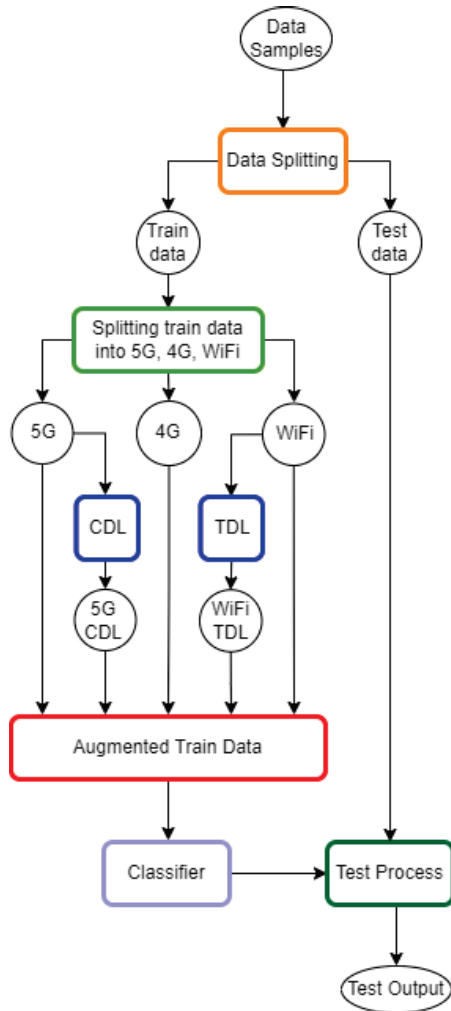


Figure 4.
The proposed decoupled CDL+TDL augmentation architecture.

and TDL transform just on WiFi data for data augmentation (LTE-waveform data is not augmented in CDL+TDL augmentation approach).

A minimalist illustration of decoupled CDL+TDL augmentation is given in Figure 4.

Figure 5 shows the classifier structure which is adopted from the study in [17].

6. Numerical Results

As mentioned in Section III, the POWDER dataset [17] is used in this work to test the decoupled CDL+TDL, 5G-only-CDL, and WiFi-only-TDL augmentation approaches. The public dataset used in this work includes raw I/Q signal samples transmitted from four base stations. The data instances of 5G, WiFi, and LTE signals in the dataset are collected on two different days, i.e., Day-1 and Day-2.

To cope with the channel impairments, we use CDL transformation for 5G samples and TDL transformation for WiFi samples

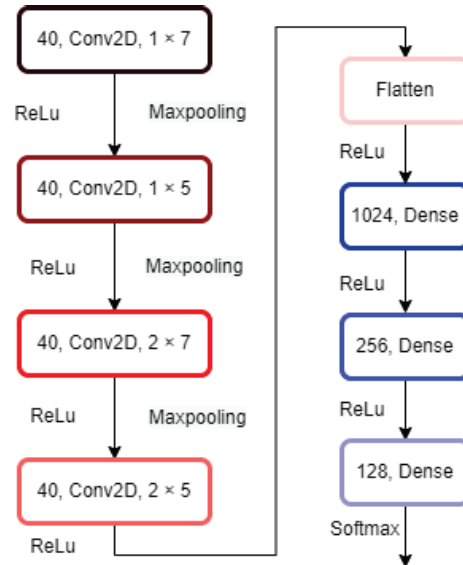


Figure 5.
The classifier used in CDL+TDL augmentation architecture.

Table 2.

Comparison of accuracy levels obtained on Day-1 and Day-2		
Augmentation Method	Accuracy on Day-1 (%)	Accuracy on Day-2 (%)
Without Augmentation	99.72	74.84
TDL/CDL-based Augmentation	99.30	77.81
5G-only-CDL Augmentation	99.20	81.63
WiFi-only-TDL Augmentation	99.23	79.21
CDL+TDL Augmentation	97.35	87.59

for data augmentation. Moreover, for a better comparison of CDL+TDL, 5G-only-CDL, and WiFi-only-TDL augmentation approaches with CDL/TDL augmentation introduced in [19], we use the same settings in [19] and the POWDER dataset in [17].

Table 2 exhibits the accuracy levels under the CDL+TDL augmentation technique, 5G-only-CDL, WiFi-only-TDL, TDL/CDL augmentation technique, and no augmentation in the four base station scenario under the POWDER dataset. The training process lasts 16 epochs for all cases.

Numerical results show that the decoupled CDL+TDL augmentation achieves significantly better performance than the TDL/CDL-based augmentation and no augmentation approaches for the scenario where the model is trained with Day-1 data and tested under Day-2 data. According to the numerical results, the decoupled CDL+TDL, 5G-only-CDL, WiFi-only-TDL augmentation techniques achieve an accuracy level of 87.59%, 81.63%, 79.21%, respectively on previously unobserved data whereas TDL/CDL augmentation technique and no augmentation approach result in an accuracy level of 77.81% and 74.84%, respectively on unobserved data. Figures 6–8 demonstrate the t-SNE plots for four base stations to visualize the impact of the augmentation methods.

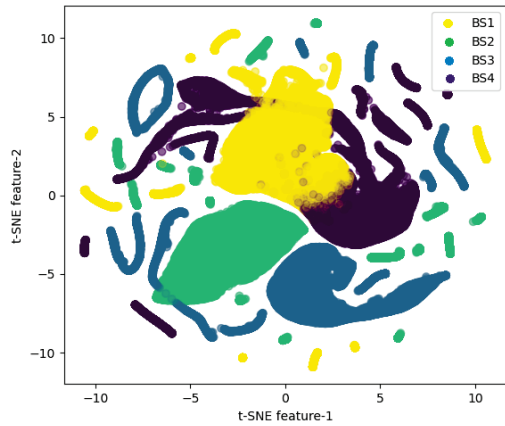


Figure 6. t-SNE visualization for Train Data with no augmentation.

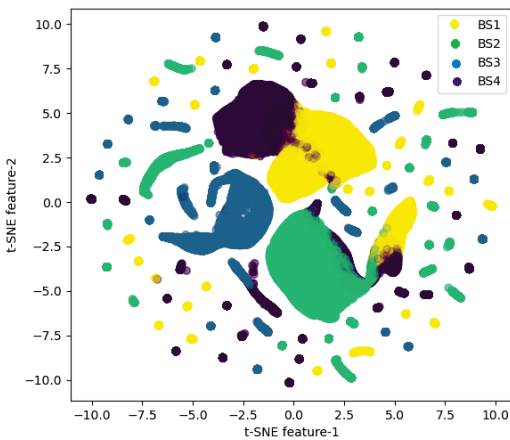


Figure 7. t-SNE visualization for Train Data with TDL/CDL Augmentation.

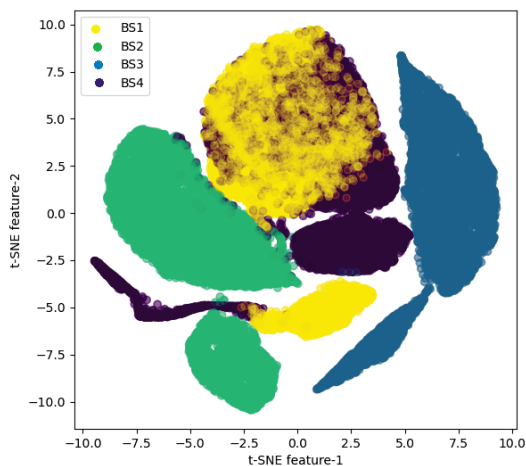


Figure 8. t-SNE visualization for Test data.

7. Future Research Challenges

This paper investigates several potential wireless security problems and proposes wireless device fingerprinting methods. In this section, we identify a few significant open research difficulties and possibilities to construct a robust radio frequency fingerprinting system (RFFS) for completeness. These issues include simulation-reality gap, the impact of receiver hardware, RFFS vulnerabilities [8].

7.1. Simulation-reality Gap

Synthetic data realism is also crucial. Deep learning models trained on synthetic data are hard to apply to real radio transmissions. The synthetic dataset’s assumptions about transmitter hardware defects and fading channel vs actual hardware and environmental impacts create a capability gap. Real-world data from IoT sensors and radios are scarce, which leads to synthetic data. In natural language processing (NLP) and computer vision (CV), large-scale datasets like MNIST [24], Stanford sentiment [25], IMDb [26], Sentiment140 [27], etc., are easily available. Existing datasets cannot be used in different machine learning frameworks since there is no standard dataset structure and organization. To achieve universal performance, neural networks must be trained with more data. Generalization is the first step to deployment-ready fingerprinting.

7.2. Impact of Receiver Hardware

Like transmitter hardware, receiver hardware captures and processes emissions for fingerprinting. The receiver hardware’s phase noise, clock offsets, filter distortions, IQ imbalance, etc., could alter the transmitter’s measured fingerprint to look like a rogue or unrecognized emitter. The ADC sampling rate and low pass filter (LPF) bandwidth are equally crucial in keeping fingerprint features in the PSD side lobes. Using MicaZ sensors, higher sampling rates retained fingerprint features but increased noise [28].

The transmitter and receiver antenna polarization and orientation can also affect the radiation pattern and fingerprint extraction performance. Emitter antenna hardware imperfections can contribute to the fingerprint feature set for wireless emitter identification [29]. The quantity, type, direction, and polarization of receiving antennas can affect fingerprinting system classification performance.

In supervised learning, several receiver hardware captures for an emitter might be used. The model could generalize and distinguish the emitter fingerprint from recorded waveforms with a bigger training data set. Training on samples from certain receiver hardware and analyzing the learned emitter function can determine fingerprinting algorithm independence.

7.3. RFFS Vulnerabilities

Broadcast wireless emitters are vulnerable to identity spoofing, Impersonation, DoS, bandwidth theft, etc. It’s often forgotten that passive receiver threats can accumulate cognitive RFFS from

transmitter emissions. Another intriguing research challenge to improve wireless security is developing or perturbing the emitter fingerprint so that only legitimate receivers can extract or identify it. RF fingerprinting is expected to be robust to impersonation attacks due to the difficulty of duplicating frontend impairments with replay assaults due to the replaying device's hardware faults. This area is still mostly unexplored. Modulation-based RF fingerprinting is more vulnerable to impersonation attempts than transient-based [30]. Another study in [31] examined how numerous low-end receivers made with cheap analog components affected modulation-based RF fingerprinting's impersonation resistance. The receivers' RF fingerprints differed from the transmitter's. As mentioned above, the receiver's hardware defects enhance the fingerprint feature set. They use this information to combat impersonation attacks by claiming that the impersonator would not be able to extract the receiver hardware's fingerprint features, making RFFS even more secure. Jamming DoS assaults, when the intruder transmits continuously on the working frequency, can also affect the RFFS. RFFS's resilience to DoS attacks will need further investigation. Jamming can also be utilized for clandestine and confidential activities to hide transmitters' RF fingerprints. In [32], WiFi transmissions were experimentally analyzed for RF fingerprint obfuscation that only the authentic receiver can extract. Random phase faults allow only authentic receivers with a preshared key and randomization index to decode the message and fingerprint.

8. Conclusion

Deep Learning (DL) has been shown to strengthen RF fingerprinting in identifying a transmitter; however, identification accuracy may fail in the presence of unobserved data due to changing channel conditions. Since collecting data under all circumstances is not feasible, data augmentation emerges as a viable approach. Previously, Day-1 data was augmented by using TDL/CDL transformation along with AWGN added on the transformed data. In a four-base station scenario with POWDER dataset, TDL/CDL-based augmentation can result in an accuracy of 77.81% on Day-2. In this study, we introduce the 5G-only-CDL and WiFi-only-TDL approaches along with the decoupled CDL+TDL augmentation where 5G-waveform data is augmented with CDL transformation, and WiFi-waveform data is augmented with TDL transformation. CDL+TDL, 5G-only-CDL, WiFi-only-TDL augmentation approaches have been shown to increase the accuracy level to 87.59%, 81.63%, 79.21% on Day-2 leading to a remarkable improvement in the performance of the TDL/CDL-based augmentation.

References

- [1] J.P. Yaacoub, H. Noura, O. Salman, and A. Chehab, "Security analysis of drones systems: attacks, limitations, and recommendations", *Internet of Things*. 2020 Sep; 11: 100218.
- [2] T. Alladi, Naren, G. Bansal, V. Chamola, M. Guizan, "SecAuthUAV: a novel authentication scheme for UAV-ground station and UAV-UAV communication", *IEEE Transactions on Vehicular Technology*, vol. 69, no. 12, December 2020.
- [3] J. Basse, X. Li, L. Qian, "Device authentication codes based on RF fingerprinting using deep learning", arxiv: 2004.08742v1, 19 April 2020.
- [4] Y. Shi, K. Davaslioglu, Y. E. Sagduyu, "Generative adversarial network in the air: Deep adversarial learning for wireless signal spoofing", *IEEE Transactions on Cognitive Communications and Networking*, Mar. 2021.
- [5] Y. E. Sagduyu, Y. Shi, and T. Erpek, "Adversarial deep learning for over-the-air spectrum poisoning attacks", *IEEE Transactions on Mobile Computing*, Feb. 2021.
- [6] Y.S. Shiu, S. Y. Chang, H. C. Wu, S-CH Huang and H. H. Chen, "Physical layer security in wireless networks: a tutorial", *IEEE Wireless Communications*, April 2011.
- [7] S. Karunaratne, E. Krijestorac and D. Cabric, "Penetrating RF Fingerprinting-based Authentication with a Generative Adversarial Attack", ICC 2021 - IEEE International Conference on Communications, Montreal, QC, Canada, 2021, pp. 1–6, doi: 10.1109/ICC42927.2021.9500893.
- [8] Jagannath, A., Jagannath, J., and Kumar, P.S.P.V.: A Comprehensive Survey on Radio Frequency (RF) Fingerprinting Traditional Approaches, Deep Learning, and Open Challenges. *Computer Networks*, vol. 219, 2022, 109455, ISSN 1389-1286, <https://doi.org/10.1016/j.comnet.2022.109455>.
- [9] C. Comert, O. M. Gul, M. Kulhandjian, A. Touazi, C. Ellement, B. Kantarci, C. D'Amours, "Secure Design of Cyber-Physical Systems at the Radio Frequency Level: Machine and Deep Learning-Driven Approaches, Challenges and Opportunities", *the book chapter (accepted on 20 April 2022) for a book publication by Springer*, pp. 1–32.
- [10] J. Liu, M. Nogueira, J. Fernandes and B. Kantarci, "Adversarial machine learning: A multilayer review of the state-of-the-art and challenges for wireless and mobile systems", in *IEEE Communications Surveys & Tutorials*, vol. 24, no. 1, pp. 123–159, Firstquarter 2022.
- [11] M. Cekic, S. Gopalakrishnan, U. Madhoo, "Wireless fingerprinting via deep learning the impact of confounding factors", *55th Asilomar Conference on Signals, Systems, and Computers*, 2021, pp. 677–684
- [12] T. J. O'Shea, T. Roy, T. C. Clansy, "Over the air deep learning based radio signal classification", *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 1, pp. 168–179, Feb. 2018.
- [13] K. Sankhe, M. Belgiovine, F. Zhou, S. Riyaz, S. Ioannidis and K. Chowdhury, "ORACLE: Optimized Radio Classification through convolutional neural networks", *IEEE Conference on Computer Communications (Infocom)*, 2019, pp. 370–378.
- [14] E. Ozturk, F. Erden, and I. Guvenc "RF-based low-SNR classification of UAVs using convolutional neural networks", arXiv online, arXiv:2009.05519v2
- [15] S. Mohanti, N. Soltani, K. Sankhe, D. Jaisinghani, M. Di Felice and K. Chowdhury, "AirID: Injecting a custom RF fingerprint for enhanced UAV identification using deep learning", *GLOBECOM 2020 – 2020 IEEE Global Communications Conference*, 2020, pp. 1–6.
- [16] N. Soltani, K. Sankhe, J. Dy, S. Ioannidis, and K. Chowdhury, "More is better: Data augmentation for channel-resilient RF fingerprinting", *IEEE Communications Magazine*, September 2020.
- [17] G. Reus-Muns, D. Jaisinghani, K. Sankhe, K. R. Chowdhury, "Trust in 5G Open RANs through Machine Learning: RF Fingerprinting

- on the POWDER PAWR Platform”, *GLOBECOM 2020 IEEE Global Communications Conference*.
- [18] C. Zhao, C. Chen, Z. Cai, M. Shi, X. Du, and M. Guizani, “Classification of small uavs based on auxiliary classifier wasserstein gans”, in *IEEE Global Communications Conference (GLOBECOM) 2018*, pp. 206–212.
 - [19] C. Cömert, M. Kulhandjian, O. M. Gul, A. Touazi, C. Ellement, B. Kantarci, C. D’Amours, “Analysis of Augmentation Methods for RF Fingerprinting under Impaired Channels”, *ACM WiseML 2022*, pp. 3–8.
 - [20] End to end Simulations, Mathworks, available at: <https://www.mathworks.com/help/5g/end-to-end-simulation.html>.
 - [21] Study on channel model for frequencies from 0.5 to 100 GHz, ETSI Technical Report, version 16.1.0 Release 16, available at: https://www.etsi.org/deliver/etsi_tr/138900_138999/138901/16.01.00_60/tr_138901v160100p.pdf.
 - [22] N. Moseley, C.H. Slump, “A low-complexity feed-forward I/Q imbalance compensation algorithm”, *Computational Statistics & Data Analysis*, January 2006.
 - [23] O. M. Gul, M. Kulhandjian, B. Kantarci, A. Touazi, C. Ellement, C. D’Amours, “On the Impact of CDL and TDL Augmentation for RF Fingerprinting under Impaired Channels”, *48th Wireless World Research Forum (WWRF 2022)*, 07–09 November 2022, UAE, pp.1–6.
 - [24] Y. LeCun and C. Cortes, “MNIST handwritten digit database”, 2010. [Online]. Available: <http://yann.lecun.com/exdb/mnist/>.
 - [25] R. Socher, A. Perelygin, J. Wu, J. Chuang, C. D. Manning, A. Ng, and C. Potts, “Recursive deep models for semantic compositionality over a sentiment treebank”, in *Proc. of Empirical Methods in Natural Language Processing*, 2013.
 - [26] A. L. Maas, R. E. Daly, P. T. Pham, D. Huang, A. Y. Ng, and C. Potts, “Learning word vectors for sentiment analysis”, in *Proc. of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*. Portland, Oregon, USA: Association for Computational Linguistics, June 2011, pp. 142–150. [Online]. Available: <http://www.aclweb.org/anthology/P11-1015>.
 - [27] A. Go, R. Bhayani, and L. Huang, “Twitter sentiment classification using distant supervision”, pp. 1-6, 2009. [Online]. Available: <http://www.stanford.edu/~alecmgo/papers/TwitterDistantSupervision09.pdf>.
 - [28] W. Wang, Z. Sun, S. Piao, B. Zhu, and K. Ren, “Wireless physical layer identification: Modeling and validation,” *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 9, pp. 2091–2106, 2016.
 - [29] B. Danev, T. S. Heydt-Benjamin, and S. Capkun, “Physical-layer identification of rfid devices”, in *USENIX security symposium*, 2009, pp. 199–214.
 - [30] B. Danev, H. Luecken, S. Capkun, and K. El Defrawy, “Attacks on physical-layer identification”, in *Proc. of the Third ACM Conference on Wireless Network Security*, ser. *WiSec ’10*. New York, NY, USA: Association for Computing Machinery, 2010, pp. 89–98. [Online]. Available: <https://doi.org/10.1145/1741866.1741882>.
 - [31] S. U. Rehman, K. W. Sowerby, and C. Coghill, “Analysis of impersonation attacks on systems using rf fingerprinting and low-end receivers”, *Journal of Computer and System Sciences*, vol. 80, no. 3, pp. 591–601, 2014, special Issue on Wireless Network Intrusion. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0022000013001220>.
 - [32] L. F. Abanto-Leon, A. Bauml, G. H. A. Sim, M. Hollick, and A. Asadi, “Stay connected, leave no trace: Enhancing security and privacy in wifi via obfuscating radiometric fingerprints”, *Proc. ACM Meas. Anal. Comput. Syst.*, vol. 4, no. 3, Nov 2020. [Online]. Available: <https://doi.org/10.1145/3428329>.

Biographies



Omer Melih Gul (S’17, M’21) received BSc., MSc., and PhD. degrees from the Department of Electrical and Electronics Engineering at Middle East Technical University (METU), Ankara, Türkiye, in 2012, 2014, and 2020, respectively by working as a research assistant at the same department. His research interests include machine learning applications, wireless security, networking, scheduling, IoT, UAV, robotics, blockchain and edge/fog computing. He has co-authored over 30 papers and 4 book chapters. He was awarded third place in the 2019 Lance Stafford Larson Outstanding Student Paper Award by the IEEE Computer Society. He was also awarded third place in the poster competition at the 2021 IEEE Rising Stars Global Conference. In 2022, he worked as a postdoctoral fellow at the School of Electrical Engineering and Computer Science at the University of Ottawa, Canada. He is a recipient of the best paper award at the 48th Wireless World Research Forum (WWRF) in 2022. Currently, he is an assistant professor in the Department of Computer Engineering at Bahcesehir University in Istanbul, Türkiye. He is also an Editor at (Springer) Telecommunications Systems and Wireless Networks journals.

On the IEEE volunteering side, after holding several IEEE volunteer roles, he has currently been a Member-at-Large (MGA Board) and the Region 8 (EMEA) Coordinator at the IEEE Computer Society since April 2021. He was awarded the 2022 IEEE MGA Young Professionals Achievement Award. He was also the chair of IEEE Turkey Young Professionals Affinity Group which was awarded the 2021 IEEE Region 8 Outstanding Young Professionals Affinity Group Award and the 2022 IEEE MGA Young Professionals Hall of Fame Honorable Mention.



Michel Kulhandjian (M’18-SM’20) received his B.S. degree in Electronics Engineering and Computer Science (Minor), with “Summa Cum Laude” from the American University in Cairo (AUC) in 2005, and the M.S. and Ph.D. degrees in Electrical Engineering from the State University of New York at Buffalo in 2007 and

2012, respectively. He was employed at Alcatel-Lucent, in Ottawa, Ontario, in 2012. In the same year he was appointed as a Research Associate at EION Inc. In 2016 has been appointed as a Research Scientist at the School of Electrical Engineering and Computer Science at the University of Ottawa. He was also employed as a senior embedded software engineer at L3Harris Technologies from 2016 to 2021. Currently, he is a Research Scientist at the Electrical and Computer Engineering Department at Rice University. He received the Natural Science and Engineering Research Council of Canada (NSERC) Industrial R&D Fellowship (IRDF).

His research interests include wireless multiple access communications, adaptive coded modulation, waveform design for overloaded code-division multiplexing applications, RF and audio fingerprinting, channel coding, space-time coding, adaptive multiuser detection, statistical signal processing, machine learning, covert communications, spread-spectrum steganography and steganalysis. He has served as a guest editor for the Journal of Sensor and Actuator Networks (JSAN). He actively serves as a member of Technical Program Committee (TPC) of IEEE WCNC, IEEE GLOBECOM, IEEE ICC, and IEEE VTC. He has served as a Guest Editor for Journal of Sensor and Actuator Networks. He is a recipient of the best paper award at the 48th Wireless World Research Forum (WWRF) in 2022.



Burak Kantarci (S'05, M'09, SM'12) received the Ph.D degree in computer engineering in 2009. He is a Full Professor and the Founding Director of Smart Connected Vehicles Innovation Centre (SCVIC), and the Founding Director of the Next Generation Communications and Computing Networks (NEXTCON) Research Lab, University of Ottawa. He has coauthored over 250 publications in established journals and conferences, and 15 book chapters. He is well known for his contributions to the quantification of data trustworthiness in mobile crowd-sensing (MCS) systems, and game theoretic incentives to promote user participation in MCS campaigns with high value data; as well as AI-backed access control, authentication and machine learning-backed intrusion detection solutions in sensing environments. In 2022, he has been awarded a Minister's Award of Excellence in Innovation and Entrepreneurship from Ontario Ministry of Colleges and Universities. He served as the Chair of IEEE Communications Systems Integration and Modeling Technical Committee, and has served as the Technical Program Co-Chair/Symposium Co-chair of more than twenty international conferences/symposia/workshops, including IEEE Global Communications Conference (GLOBECOM)—Communications Systems QoS, Reliability and Modeling (CQRM) symposium. In 2021, he has been elected as the new Secretary of IEEE Social Networks Technical Committee. He is

an Editor of the IEEE Communications Surveys & Tutorials, IEEE Internet of Things Journal, Vehicular Communications (Elsevier), and an Associate Editor for IEEE Networking Letters, and Journal of Cybersecurity and Privacy. He is Editor for several IEEE and Elsevier journals. He was an ACM Distinguished Speaker in 2019-2021, currently an ACM Senior Member. He is IEEE Systems Council and IEEE Communications Society Distinguished Lecturer.



Claude D'Amours (Member, IEEE) received the B.A.Sc., M.A.Sc., and Ph.D. degrees in electrical engineering from the University of Ottawa, in 1990, 1992, and 1995, respectively. In 1992, he was employed as a Systems Engineer at Calian Communications Ltd. In 1995, he joined the Communications Research Centre, Ottawa, ON, Canada, as a Systems Engineer. Later in 1995, he joined the Department of Electrical and Computer Engineering, Royal Military College of Canada, Kingston, ON, Canada, as an Assistant Professor. He joined the School of Information Technology and Engineering (SITE), which has since been renamed as the School of Electrical Engineering and Computer Science (EECS), University of Ottawa, as an Assistant Professor, in 1999. From 2007 to 2011, he has worked as the Vice Dean of undergraduate studies with the Faculty of Engineering and has been working as the Director of the School of EECS, University of Ottawa, since 2013. His research interests include physical layer technologies for wireless communications systems, notably in multiple access techniques and interference cancellation. He is a recipient of the best paper award at 48th Wireless World Research Forum (WWRF) in 2022.



Azzedine Touazi received BSc., MSc., and PhD. degrees from Ecole Nationale Polytechnique (ENP) d'Alger and USTHB university, Algeria, all in electrical engineering, in 2003, 2007, and 2017, respectively. His research interests include signal and image processing, machine learning, statistical data analysis, anomaly detection, and 4G/5G networks. He is currently ML and 4G/5G engineer at ThinkRF Corp. Prior to this he held engineering

positions at Alcatel-Lucent and Aerosystems International (ASI). Also, he was a researcher at the Center for Development of Advanced Technologies (CDTA). He is a recipient of the best paper award at 48th Wireless World Research Forum (WWRF) in 2022.



Cliff Ellement, Head AI Solutions at ThinkRF, has 25 years of Product and Business development experience. He holds a Masters of Electrical Engineering at Concordia University researching Spread-Spectrum and Mobile communications relating anti-jamming and network optimization techniques. Developed a number of products in the Optical, Microwave, and Data networking businesses at Nortel. Later joining Mitel Networks leading the Unified Communications group developing secure Intelligent Enterprise collaboration applications. Cliff has been involved at the intersection of RF Communications, Security and Artificial Intelligence initiatives to help drive the next generation of wireless security solutions for Critical Infrastructures. He is a recipient of the best paper award at 48th Wireless World Research Forum (WWRF) in 2022.



Call for Papers

Future Global Intelligent and Sustainable Communications

51st Wireless World Research Forum

Rowan University, New Jersey, USA

18-20 June 2024

Future wireless networks for global communication stand at the confluence of a multifaceted evolution, innovations in artificial intelligence and machine learning, and evolving techniques for ensuring sustainability. The demand for new services has given rise to explosive growth in wireless traffic and the resulting impacts on the environment have raised concerns about sustainability. In order to ensure sustainability in future global wireless communication, intelligence in networks emerges as a critical tool.

Artificial intelligence (AI) and machine learning (ML) will serve as key enablers on multiple aspects of future wireless networks. On the air interface, AI/ML algorithms will ensure optimal and efficient allocation of resources, increase spectral efficiency, and maximize network capacity. Intelligence in base stations will minimize power consumption based on real-time demand. Network optimization algorithms based on AI/ML can analyze vast amounts of data, predict traffic patterns, optimally allocate network resources, and minimize energy consumption during low-traffic periods. Intelligent network management techniques using AI/ML algorithms can predict cyber threats and proactively thwart such threats and improve network security.

AI/ML algorithms will also impart sustainability to future wireless networks. For example, AI/ML-based optimization algorithms can facilitate the integration of renewable energy sources into the network infrastructure, minimize the environmental impact by the use of biodegradable resources, and extend the lifespan of network resources. By exploiting the synergy between intelligence and sustainability, optimized AI/ML algorithms can pave the way for reducing greenhouse emissions, and minimizing energy consumption in the edge, network, and cloud.

Under the theme “Future Global Intelligent and Sustainable Communications”, the 51st Wireless World Research Forum (WWRF) will take place at Rowan University, Glassboro, New Jersey, USA from 18-20 June 2024. You are invited to be part of designing the wireless future by joining us for three days of insightful discussions, presentations, innovative brainstorming, and expert-level networking.

Topics of Interest

Authors are invited to submit original manuscripts aligned with the theme of the event, on one or the more of the following topics, or on any relevant subject relating to intelligent and sustainable communications for a better world:

- Autonomous communications
- Towards automated wireless communications
- Resource sustainability in future communications
- Cutting-edge solutions for sustainable communications
- Openness, disaggregation, modularity and programmability
- Intelligent applications for vertical industries
- Data analytics, AI, and machine learning for sustainability and network automation
- Software-defined infrastructures
- Advanced radio technologies
- Tactile Internet
- Green communications and networking
- THz communications
- Bridging the digital divide beyond 5G
- Beyond-5G technologies
- Innovations in business models for wireless networks
- Cyber-physical systems and networks
- Privacy and security
- Connected vehicles
- Holographic MIMO & reconfigurable intelligent surfaces
- Applications and impact of quantum-based technologies
- Semantic communications
- Spectrum issues and regulatory principles
- Social network-aware wireless
- Internet of Things and wearable technologies
- Quantum communications

Authors are expected to be physically present at the meeting to present their contribution.

Submission Instructions

Contributors should submit an extended abstract by 30th April 2024 to contributions@wwrf.ch for review. Extended abstracts should be preferably at least two pages in length, either in plain ASCII text, MS Word or Adobe PDF A template for abstracts or papers is available. Full papers must be prepared using the WWRF template.

Further details, including information on WWRF working groups, copyright licences, templates and availability of student travel grants are available on the WWRF website www.wwrf.ch

WIRELESS WORLD
RESEARCH FORUM

 **RowanUniversity**
HENRY M. ROWAN
COLLEGE OF ENGINEERING

WIRELESS WORLD RESEARCH FORUM

HUDDLE 2024

Implementing the 6G Framework | Terrestrial and Beyond

17 - 18 APRIL 2024

BERLIN

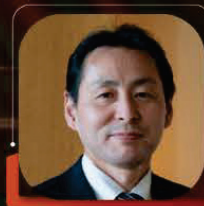
Confirmed Speakers Include



**Dr. Wilhelm
Eschweiler**
Vice President
BNetzA



Thabisa Faye
Councillor
Independent
Communication
Authority of South
Africa



**Takehiro
Nakamura**
Director
NTT Docomo



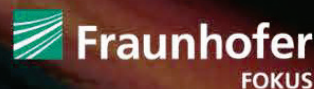
Martha Suárez
President
Dynamic
Spectrum
Alliance
(DSA)

Register Here:



wwrfhuddle.com

In partnership with



Bundesnetzagentur

Upcoming Events of Interest (Provided by WWRF Members)

No	Event	Date/Venue	Member/Host	Website
1	WWRF Huddle 2024	17–18 April, 2024, Berlin, Germany	Fraunhofer-Institut fuer Offene Kommunikationssysteme FOKUS, Germany	https://wp.wwrfhuddle.com/
2	WWRF Internal Workshop ‘6G Strategy for WWRF’	16th April, Berlin, Germany	Bundesnetzagentur (BNetzA), Germany	https://wp.wwrfhuddle.com/
3	WWRF Workshop ‘ehealth’	16th April, Berlin, Germany	Bundesnetzagentur (BNetzA), Germany	https://wp.wwrfhuddle.com/
4	WWRF Workshop ‘Integrated Sensing and Communication’	19th April, Berlin, Germany	Bundesnetzagentur (BNetzA), Germany	https://wp.wwrfhuddle.com/
5	WWRF Workshop ‘Connected Vehicles and AI for 6G’	19th April, Berlin, Germany	Bundesnetzagentur (BNetzA), Germany	https://wp.wwrfhuddle.com/
6	WWRF Meeting #51	18–20 June, 2024, Glassboro, NJ, USA	Rowan University, USA	www.wwrf.ch
7	WWRF Meeting #52	10–12 September, 2024, London, U.K.	Kings College London, UK	www.wwrf.ch
8	Second RF Summit Finland	March 12 2024 Dipoli, Espoo, Finland	University of Oulu, Finland	https://www.6gflagship.com/event/second-rf-summit-finland/
9	Future mobile radio frequency technologies and solutions	March 16 2023, University of Oulu	University of Oulu, Finland	https://rf-sampo.rf-hub.org/future-mobile-radio-frequency-technologies-and-solutions-seminar/
10	IEEE WCNC workshop on “WS-06: Beyond Connectivity: When Wireless Communications Meet Generative AI”	21–24 April 2024, Conrad Hotel in Dubai, United Arab Emirates	Merouane Debbah, Khalifa University	https://wcnc2024.ieee-wcnc.org/workshop/ws-06-beyond-connectivity-when-wireless-communications-meet-generative-ai
11	11th Nordic Workshop on System and Network Optimization for Wireless	8 April 2024, Levi Summit, Levi, Finland	University of Oulu, Finland	https://www.6gflagship.com/snow-2024/
12	6G Symposium, Crucial Conversations 2024	9–11, April 2024, Levi Summit	University of Oulu, Finland	https://www.6gworld.com/6gsymposium-spring-2024/
13	IEEE WCNC Tutorial – “Security, Privacy and Anonymity in Emerging Applications in 6G and Beyond: A Physical Layer Perspective”	Apr 21, 2024, Conrad Hotel in Dubai, United Arab Emirates	University College London (UCL)	https://wcnc2024.ieee-wcnc.org/tutorials#tut-03

No	Event	Date/Venue	Member/Host	Website
14	The 2nd International Workshop on Digital Twins for Next Generation Networks & Real-Time IoT Systems Beyond 5G (DT4NGN-IoT)	May 22–25, 2024, Carthage, Tunis, Tunisia	University College London (UCL)	https://isorc.github.io/2024/html/workshop.html
15	IEEE Radar Conference Tutorial – “Multi-Function RF Systems for Radar and Communications: Signal Processing, Prototyping and Experiments”	May 6–10, 2024, Denver Colorado	University College London (UCL)	https://2024.ieee-radarconf.org/program/tutorials
16	Nordic Conference on Digital Health and Wireless Solutions	May 7–8, 2024, Hotel Lasaretti, Oulu, Finland	University of Oulu, Finland	https://nordic-digihealth.com/
17	IEEE International Communications Conference Workshop – GreenNet KEYNOTE	9–13 June 2024, Denver, CO, USA	John M Cioffi, Advisory Board, TII	https://icc2024.ieee-icc.org/workshop/ws-11-third-international-workshop-green-and-sustainable-networking-greennet
18	IEEE ICC Panel on “Towards an Open 6G for the inclusion of all – the role of open 6G research infrastructures and toolkits”	9–13 June 2024, Denver, CO, USA	Nigel Jefferies, WWRF	https://icc2024.ieee-icc.org/
19	IEEE ICC workshop on “Empowering Telecom Networks with Large-Scale Language Models”	9–13 June 2024, Denver, CO, USA	Merouane Debbah, Khalifa University	https://sites.google.com/view/ieee-icc-2024-llmtelecom/home
20	UBISS 2024 12th International UBI Summer School	June 10–15, 2024, Center for Ubiquitous Computing at University of Oululand,	University of Oulu, Finland	https://ubicomp.oulu.fi/UBISS2024
21	2024 EuCNC & 6G Summit	June 3–6, 2024, Anrwerp Belgium	University of Oulu, Finland	https://www.eucnc.eu/
22	Certified IPv6 Migration Strategy for Regulators and Governing Authorities	June 20–21, 2024, Kuala Lumpur, Malaysia	Sures Ramadass, Malaysia	https://forms.gle/Rp2G6fr31V1XLWg9A
23	SPAWC Special Session on “AI-Driven Signal Processing for Advanced Wireless Communications”	September 10–13, 2024, Lucca Italy	Merouane Debbah, Khalifa University	https://spawc2024.org/special-sessions-cs/
24	IEEE Veh Tech Magazine Special Issue on Integrated Sensing and Communication		University College London (UCL)	http://www.ieeevtc.org/vtmagazine/specisu--ISAC-SI-2023.php