Tampere University

RUSTAM PIRMAGOMEDOV

Facilitating Internet of Things on the Edge

Tampere University Dissertations 248

Tampere University Dissertations 248

RUSTAM PIRMAGOMEDOV

Facilitating Internet of Things on the Edge

ACADEMIC DISSERTATION
To be presented, with the permission of
the Faculty of Information Technology and Communication Sciences
of Tampere University,
for public discussion at Tampere University
on 12 May 2020, at 12 o’clock.

ACADEMIC DISSERTATION
Tampere University, Faculty of Information Technology and Communication Sciences
Finland

Responsible supervisor and Custos	Professor Evgeny Kucheryavy Tampere University Finland	
Pre-examiners	Professor Timo Hämäläinen University of Jyväskylä Finland	Professor Periklis Chatzimisios Alexander Technological Educational Institute of Thessaloniki Greece
Opponent	Professor Edmundo Monteiro University of Coimbra Portugal	

The originality of this thesis has been checked using the Turnitin OriginalityCheck service.

Cover design: Roihu Inc.

ISBN 978-952-03-1552-8 (print)
ISBN 978-952-03-1553-5 (pdf)
ISSN 2489-9860 (print)
ISSN 2490-0028 (pdf)
http://urn.fi/URN:ISBN:978-952-03-1553-5

PunaMusta Oy – Yliopistopaino
Tampere 2020

Dedicated to my father.

PREFACE

This work was carried out at the Electrical Engineering Unit of Tampere University (Finland) during 2018-2019. First of all, I wish to extend my appreciation to Prof. Yevgeni Koucheryavy, for his guidance and support throughout my research. I would also like to thank Dr. Dmitri Moltchanov and Assist. Prof. Sergey Andreev, for their support and expertise. Additionally, I would like to thank Dr. Alexander Pyattaev for his engaging and relevant discussions on network technologies. I would also like to give special thanks to Sari Isokääntä from Tampere University's Language Center, who provided me with a framework for the continued development of my academic writing skills. Finally, I would like to thank Stephen Wantuck for his fresh view of my publications and valuable comments towards the improvement of their presentation.

Rustam Pirmagomedov. February 14, 2020, Tampere, Finland.

ABSTRACT

The evolution of electronics and wireless technologies has entered a new era, the Internet of Things (IoT). Presently, IoT technologies influence the global market, bringing benefits in many areas, including healthcare, manufacturing, transportation, and entertainment.

Modern IoT devices serve as a thin client with data processing performed in a remote computing node, such as a cloud server or a mobile edge compute unit. These computing units own significant resources that allow prompt data processing. The user experience for such an approach relies drastically on the availability and quality of the internet connection. In this case, if the internet connection is unavailable, the resulting operations of IoT applications can be completely disrupted. It is worth noting that emerging IoT applications are even more throughput demanding and latency-sensitive which makes communication networks a practical bottleneck for the service provisioning. This thesis aims to eliminate the limitations of wireless access, via the improvement of connectivity and throughput between the devices on the edge, as well as their network identification, which is fundamentally important for IoT service management.

The introduction begins with a discussion on the emerging IoT applications and their demands. Subsequent chapters introduce scenarios of interest, describe the proposed solutions and provide selected performance evaluation results. Specifically, we start with research on the use of degraded memory chips for network identification of IoT devices as an alternative to conventional methods, such as IMEI; these methods are not vulnerable to tampering and cloning. Further, we introduce our contributions for improving connectivity and throughput among IoT devices on the edge in a case where the mobile network infrastructure is limited or totally unavailable. Finally, we conclude the introduction with a summary of the results achieved.

CONTENTS

List of Tables

ABBREVIATIONS

5G	Fifth generation mobile networks
AAA	Authentication, Authorization, Accounting
AH	Augmented human
AI	Artificial intelligence
AP	Access point
AR	Augmented reality
BAN	Body area network
CNN	Convolutional neural networks
D2D	Device-to-Device communication
e.g.	for example, from Latin *exempli gratia*
EM	Electromagnetic
EWMA	Exponentially-weighted moving average
GHz	Gigahertz
HPBW	Half-power beamwidth
ICN	Information-centric networking
ICT	Information and communications technology
IMEI	International Mobile Equipment Identity
IoT	Internet of Things
ITU	International Telecommunication Union
LoS	Line of sight
LSA	Link-state advertisement

MAC	Medium access control
mmW	Millimeter wave
MTI	The mean time required for device identification
PUF	Physical unclonable function
PWS	Passive wireless sensor
QS	Queuing system
RAT	Radio access technology
RDM	Random direction mobility
RFID	Radio frequency identification
SLS	System-level simulator
STA	Wi-Fi station (a user's device)
THz	Terahertz
UAV	Unmaned Aerial Vehichle
UGV	Unmanned ground vehicle
UHF	Ultra high frequency
VR	Virtual reality

ORIGINAL PUBLICATIONS

Publication I — R. Pirmagomedov and Y. Koucheryavy. IoT Technologies for Augmented Human: a Survey. *Internet of Things* (2019), 100120. DOI: 10.1016/j.iot.2019.100120.

Publication II — S. S. Vladimirov, R. Pirmagomedov, R. Kirichek and A. Koucheryavy. Unique Degradation of Flash Memory as an Identifier of ICT Device. *IEEE Access* 7 (2019), 107626–107634. DOI: 10.1109/ACCESS.2019.2932804.

Publication III — R. Pirmagomedov, R. Kirichek, M. Blinnikov and A. Koucheryavy. UAV-based gateways for wireless nanosensor networks deployed over large areas. *Computer Communications* 146 (2019), 55–62. DOI: 10.1016/j.comcom.2019.07.026.

Publication IV — T. D. Dinh, R. Pirmagomedov, V. D. Pham, A. A. Ahmed, R. Kirichek, R. Glushakov and A. Vladyko. Unmanned aerial system–assisted wilderness search and rescue mission. *International Journal of Distributed Sensor Networks* 15.6 (2019), 1550147719850719. DOI: 10.1177/1550147719850719.

Publication V — R. Pirmagomedov, D. Moltchanov, V. Ustinov, M. N. Saqib and S. Andreev. Performance of mmWave-Based Mesh Networks in Indoor Environments with Dynamic Blockage. *International Conference on Wired/Wireless Internet Communication*. Ed. by D. F. M., N. E., B. R. and K. A. 2019, 129–140. DOI: 10.1007/978-3-030-30523-9_11.

Publication VI — R. Pirmagomedov, D. Moltchanov, A. Ometov, K. Muhammad, S. Andreev and Y. Koucheryavy. Facilitating mmWave Mesh Reliability in PPDR Scenarios Utilizing Artificial Intelligence. *IEEE Access* (2019). ISSN: 2169-3536. DOI: `10.1109/ACCESS.2019.2958426`.

Author's contribution

This thesis relies on six main publications. Five of these are published in relevant scientific journals, and one was presented at a conference and published in its proceedings.

Publication I This paper was written entirely by the author. The idea for the paper was discussed and developed in collaboration with his supervisor – Prof. Yevgeni Koucheryavy.

Publication II In this work, the author wrote the entire text of the paper, completed a review of the identification methods, evaluated the reliability and explored the computational overheads of the discussed method.

Publication III The author made a review of state of the art technology, designed the process of data acquisition, medium access control, conducted a performance evaluation for the proposed solution, and prepared the text of the entire paper.

Publication IV For this publication, the author developed the connectivity and service quality models and then conducted performance evaluations using a computer simulation. In addition, the author wrote the introduction and the review of related works.

Publication V In this work, the author designed and developed the software for the simulator, which provided the performance evaluation. The design of research objectives, system model development, and text writing was primarily performed by the author, in collaboration with Dr. Dmitri Moltchanov.

Publication VI In this publication, the author contributed by developing the general idea for the research and conducting performance evaluations. The system model was co-developed with Dr. Dmitri

Moltchanov. The author wrote a significant share of the paper, with additional insights provided by Dr. Dmitri Moltchanov.

1 INTRODUCTION

1.1 Background information

Recent advances in electronics and wireless technologies triggered significant growth in a number of wireless connected devices entering the era of IoT. Currently, the IoT devices bringing benefits in many areas, including healthcare, manufacturing, transportation, and entertainment. In addition to the horizontal expansion, the functional capabilities and technologies behind those have also progressed notably. Early IoT devices, such as RFIDs, had significant constraints in terms of communication, computation and lifetime, and performed relatively simple tasks. Later, IoT applications have extensively utilized the potential of cloud servers and communication networks for bringing new capabilities to the devices which are nonetheless still constrained.

In such a paradigm, IoT devices act as a part of a greater system, centered around a cloud. User data processing performed in such clouds is extremely prompt due to significant computational resources available there. Such an approach relies on communication networks that allow the devices to be online continuously, constantly sending raw data to the server for processing. As a result, user experiences significantly depend on the internet connection. If an internet connection is not available, the functionality of IoT applications can be fully disabled.

Emerging IoT applications are even more demanding of throughput (utilize massive media, e.g., high-quality video, sound) and latency-sensitive (e.g., autonomous vehicles, real-time control systems). From this perspective, communication networks may become a bottleneck for future applications of IoT. To address these growing demands, vendors and operators are enabling high throughput links and moving computation resources closer to users – to the edge, which is becoming a primary focus of the further technological development of IoT and poses new challenges.

1.2 Scope of the thesis

Both industry and academia consider dynamic computational services available on the edge, in proximity to users, as an alternative to central clouds. Given the fact that the lion share of data is generated on the edge and overheads for moving data to the cloud became unacceptable, relocation of computational resources closer to users seems to be a reasonable way forward. It is worth noting that computational resources may reside not only on the edge of mobile network infrastructure (e.g., on base stations) but also on hi-end user devices (e.g., smart public vehicles, smartphones). Depending on the complexity of a user's demand, the computation can be offloaded to one of those devices.

Obviously, the performance of the systems with distributed compute resources fully relies on the users' awareness of the resources available within a certain proximity and the services which can be provisioned. The fifth-generation (5G) of mobile networks target an enhanced network experience, utilizing innovative wireless technologies. However, these improvements are expected to be available in relatively dense areas while outside of those locations, users' network experience is expected to be dramatically lower. Moreover, high throughput wireless access in 5G systems relies on millimeter-wave technologies which are not able to provide reliability comparable to microwave mobile networks due to blockages and extreme propagation losses. As a result, one may conclude that even with 5G, network infrastructure does not promise smooth session continuity for IoT devices, making wireless access level a bottleneck for emerging IoT services.

This thesis contributes to the body of work on dynamic IoT services in real-time, via the enhancement of underlying communication technologies. More specifically, the work targets better connectivity and higher throughput between the devices on the edge, as well as their network identification.

1.3 Thesis outline and main results

This thesis includes an introductory part consisting of seven chapters and six main publications on the stated topic. In order to make the thesis accessible for wider audience, Chapter 2 provides an overview of emerging IoT applications, discussing fundamental challenges and trade-offs, while the subsequent three chapters elaborate our contributions on the topic, describing the methodology and presenting the main numerical results.

In Chapters 3, we discuss challenges related to network identification of IoT devices with a focus on physically unclonable functions. First, we provide an analysis of the existing identification methods with emphasis on their disadvantages. Then, we introduce a promising proposal, based on the flash memory degradation process, evaluate its reliability and overheads. Our results demonstrate that the considered approach promises enhanced identification of devices when compared against existing methods, bringing only minor computational overheads.

Chapter 4 addresses connectivity challenges related to the provisioning of IoT services utilizing drones. More specifically, we consider cases when mobile network coverage is unavailable or very limited. This chapter consists of two parts. In the first part, we elaborate on the scenario of environmental monitoring with limited, micro-sized sensors. In the second part, we discuss the provisioning of latency-sensitive proximate communications of IoT devices via a swarm of drones.

Chapter 5 focuses on exploring the capacity of mesh networks for extending millimeter-wave access technologies. Millimeter-wave meshes are expected to enable high throughput links for hi-end IoT devices, even if these operate outside of base station coverage. In this chapter, we evaluate the performance of the millimeter-wave meshes in environments with dynamic blockages and propose a new method for improving their reliability.

Finally, Chapter 6 concludes the introductory part and discusses future trends.

2 EMERGING IOT APPLICATIONS AND NEW RESEARCH AGENDA

This Chapter reviews notable emerging IoT applications and discusses new research challenges posed by them.

2.1 IoT applications for Augmented Human

Emerging IoT applications target more extensive and deep assistance in daily human routines via augmentation of human abilities. Such kinds of augmentation rely on advanced electronic devices placed in the body or in close proximity, creating an integrated ecosystem referred to as "Human 2.0" or as Augmented Human (AH). Examples of such devices may include a leg or a hand prosthesis, artificial vision systems, augmented reality glasses, artificial organs and tissues. Recreated or extended human abilities, enabled by the IoT devices, may enhance quality of life and enable advanced abilities for their users.

Current research activities on this topic cover multiple fields of science and technology, including medicine, psychology, electrical and mechanical engineering, material science, and information technologies. From the communications perspective, the devices used in AH ecosystem are considered as wearable IoT. However, contrary to the overwhelming majority of existing wearable IoT systems, elements of the AH ecosystem provide perhaps the most critical class of services, as individuals do not exist independently, but rather as a part of, human-centric AH systems [83].

Innovative AH systems rely on cutting edge electronic devices, which are connected to a single BAN via various communication technologies. This network acts as a fundamental technological layout for high-level applications of AH, which include three directions of augmentation, as illustrated in Fig. 2.1.

Augmentation of physical abilities, targeting improved abilities to move and ma-

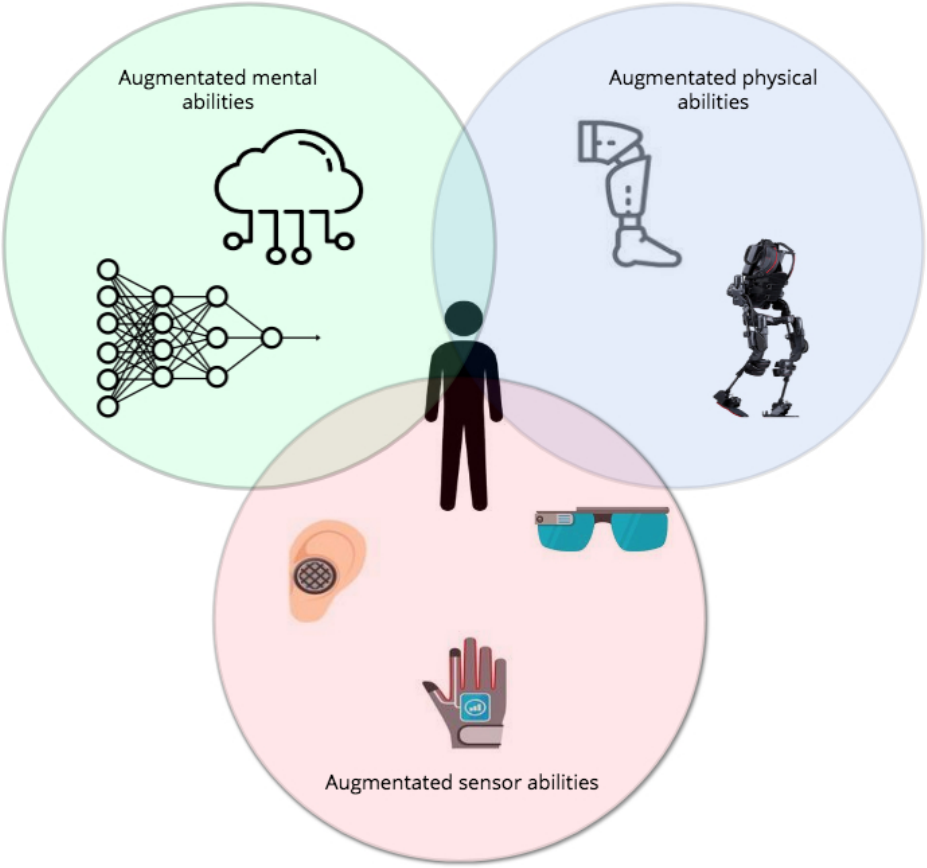

Augmentated mental abilities
Augmentated physical abilities
Augmentated sensor abilities

Figure 2.1 Human augmentation directions

nipulate objects. The physical augmentation may rely on an exoskeleton, artificial arms and legs, or a personal propulsion system.

Sensory augmentation aims at enhancing a person's awareness of the surrounding environment. Such augmentation enables advanced sensing (e.g., vision, touch, hearing, smell, and taste) via the amplification or transformation of one sensory modality into stimuli of another sensory modality [46] (e.g., visualizing sounds for people with hearing impairments).

Cognitive augmentation facilitates a person's decision making via assistance in data processing. An illustrative example of a cognitive augmentation is an electronic personal assistant, allowing one to save time and optimize resources (time and money), providing increased optimal logistics during the day. Perhaps cognitive augmentation is the most familiar to the public as a common example of human augmentation because a variety of mobile applications already provides similar types of assistance. Remarkably, such applications are often computation hungry [65], and thus organized as a thin client, making them heavily dependent on the internet connection.

The AH applications can be classified onto the three classes, depending on the augmentation goals:

- assisted living;
- enhanced professional performance;
- entertainment and resource optimization.

The applications for assisted living allow users to support their basic daily needs without other peoples' assistance. Additionally, such applications may monitor health conditions in real-time and facilitate a user's safety (e.g., avoiding hazards and protecting from occasional falls). The utilization of assisted living applications reduces social security costs (e.g., nursing) and improves the quality of life for users.

Entertainment applications deliver immersive experiences (e.g., virtual reality gaming), including extreme situations without actual risks.

Applications that target higher professional performance augment the abilities required in specific professional areas. For example, exoskeleton-based solutions allow for moving heavy weights while reducing stress for the spine, which can be relevant for certain kinds of workers. In an emergency response scenario, such advanced IoT applications may significantly enhance the efficiency of rescue crews via augmented sensing (e.g., gas detecting, thermal vision), facilitated physical abilities (e.g., exoskeleton-based solutions), and more efficient and prompt decision making (e.g., AI-aided assistance).

All the devices used by an individual constitute an integrated IoT ecosystem and should work synchronously, enabled by underlying network technologies. The communication technologies may vary from conventional radio interfaces such as Bluetooth to highly specific technologies based on THz frequency range communication or molecular nanonetworks. The variety of technologies allows one to consider the interconnected IoT devices used in AH systems as a highly heterogeneous network.

Remarkably, AH systems may interact with the proximate objects, including city

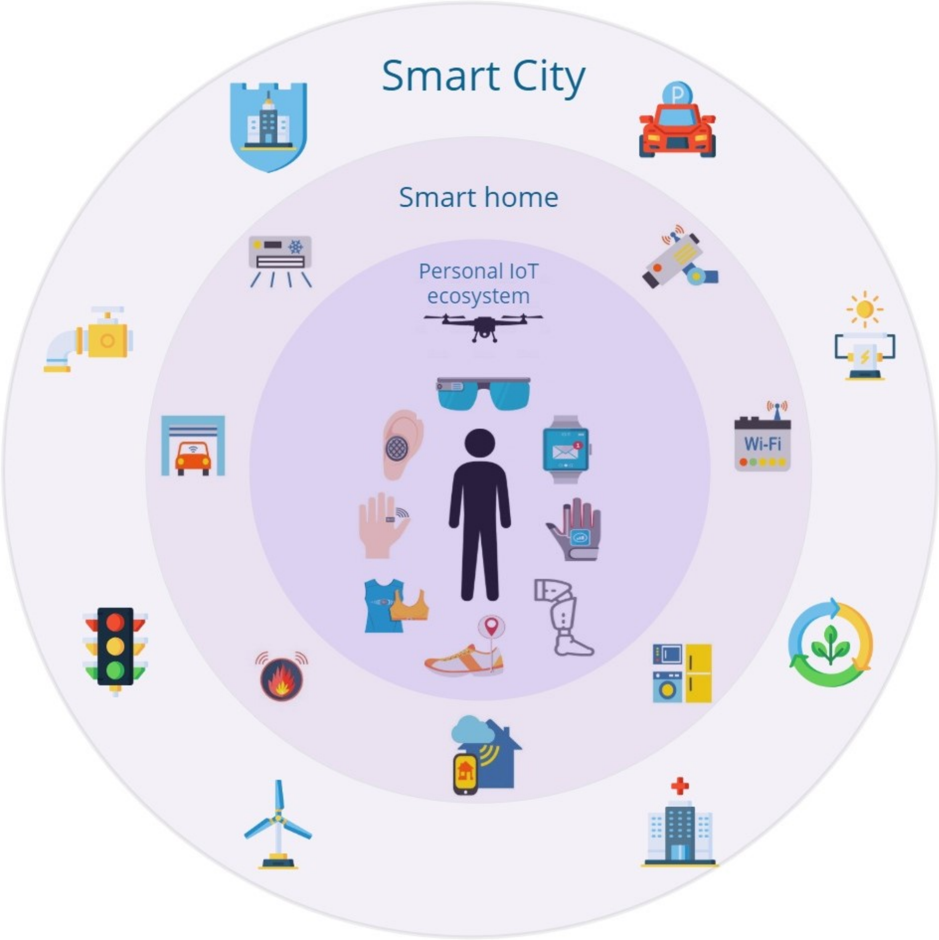


Figure 2.2 Multi-tier smart environment

infrastructure [13] and the electronic devices of other people, forming an integrated smart environment (Fig.2.2), and internet connection (e.g., for maintaining context-awareness, upgrading software). As a result, the IoT-based human augmentation systems open a number of communication challenges, since the reliable operation of communication technologies in such systems is vitally important to the users' well-being.

2.2 Environmental monitoring utilizing micro-sized devices

Another area where IoT applications are rapidly developing is environmental monitoring. Gathering information about ongoing environmental processes using wireless sensor networks allows for better resource management (e.g., soil nutrition in agriculture) and earlier detection of pollution. Therefore, wireless sensor networks are considered of primary importance for emerging industry 4.0. Commonly a wireless sensor device consists of a power element, communication block, microprocessor, and sensory element. Multiple industrial cases require the autonomous operation of the devices over a long duration. The devices in such deployments typically utilize energy harvesting technologies (e.g., solar power, the energy of electromagnetic waves). Such devices are referred to as passive wireless sensors (PWS). The benefits of PWS (low manufacturing and maintenance costs, long exploitation time) make them highly suited to many sectors, including manufacturing [75], healthcare [85], logistics [12], and environmental monitoring [48].

Recent advances in nanotechnologies have enabled micro-sized sensor devices, referred to as wireless nanosensor networks. A nanodevice is not necessarily limited to nanometers, but rather a device that utilizes unique properties of nanomaterials for the detection and measurement in the nanoscale [5]. Presently, wireless nanosensors have micro dimensions (e.g., passive acoustic nanosensor [8, 9]) and utilize the unique properties of graphene to transmit data in the THz frequency range [6].

Nanodevices may significantly advance environmental applications via more precise and inexpensive solutions. Moreover, nanosensors, due to their reduced size, can be easily integrated into a biological object. In such hybrid bio-electronic sensors, electronic nanosensors serve as a proxy, measuring the natural reaction of biological objects (e.g., bacteria, plants, animals) to the changing environment. There are a number of current examples of biosensors facilitating industrial processes. For instance, crayfish are used in water treatment plants to indicate water quality. Cutting-edge developments in this area enable minimally invasive electronic devices embedded directly onto biological objects. Such hybrid systems may enable high-quality, low-cost sensors as an alternative to expensive fully electronic devices. For example, an electronic device integrated with a plant, capable of detecting soil pollution via changes in the plant's metabolism and signaling pathways.

It is worth noting that nanosensor devices are significantly limited in energy and commonly designed as passive devices with energy harvesting capabilities [5]. Because of the energy constraints, the devices are unable to maintain a permanent communication channel with other network elements. Recent works have been mostly restricted to considering the communication aspects of body area deployments of nanonetworks. Particularly, these publications considered powering nanosensors via wireless energy transfer from on-body gateways [24, 43]. These methods facilitate personal applications of nanonetworks (e.g., medical applications), but remains almost non-applicable when nanosensors are deployed over a large area because the efficient distance of wireless energy transfer is limited by several meters as well as the distance of communication in THz range. Therefore, to utilize the potential of nanodevices in scenarios that require deployments of sensors over large areas, new methods of communicating with and powering of the devices are of essential importance.

2.3 Gaming and heavy media applications

On-line gaming and media-heavy applications can be listed as drivers of Hi-End IoT developments. Such applications often rely on augmented (AR) and virtual reality (VR) devices and high-quality sound systems to provide an immersive experience for users.

VR-based games visualize a new environment for players, one which is fully capable of reacting to their actions in the game. Such applications require extremely high throughput on the downlink [23]; however, gaming sessions commonly arise in the predefined locations (the safety of the gamers must be ensured during these sessions). Since the demands of VR-based gaming are spatially and timely predefined, these can be addressed using conventional network planning methods.

In a case of AR-based gaming, the user interface (e.g., AR glasses) visualize virtual entities over the real layout. Such a virtual entity may react in accordance with changing user behavior and changing context. The development of AR-based games is different from conventional video games because developers do not need to design the environment (e.g., grass, sky, trees) and focus their efforts only on virtual objects and their behavior. Such games are less demanding for downlink throughput while requiring higher throughput at uplink. Moreover, such applications are latency-sensitive, because notable delays disable prompt reaction of the application to user behavior or changes in the environment. Thus, users' experience in AR gaming significantly depends on the quality of the Internet connection. Moreover, AR games can be context-specific, implying that gaming sessions may appear in different locations, and actual users' network demands will present with a high degree of temporal and spatial variation. As a result, this class of applications requires new methods for provisioning reliable wireless links.

2.4 Summary of major communication challenges

At present, IoT services have advanced far beyond initial sensor networks and now utilize high-quality media. This has resulted in rapidly *increased throughput demands*, which are continuously growing. It is commonly expected that the fifth-generation (5G) networks will accommodate emerging IoT scenarios, which are not handled efficiently by 4G+ deployments [47]. As an alternative to an extensively employed microwave spectrum, the use of millimeter-wave (mmW) links is proposed [77]. Due to the higher spectrum and less interference, mmW links are considered to be a solution for the mitigation of interference and throughput concerns in emerging wearable networks [76]. However, as it was elaborated in this chapter, emerging IoT applications present with a high degree of temporal and spatial variation, while a mmW connection is expected to be limited due to high propagation losses and blockages. As a result, emerging IoT applications call for innovations of high throughput wireless communication provisioning on the access level. Further on, this work proposes and evaluates new technologies based on mmW multi-hop meshes, which are expected to address the growing throughput demands. These technologies are expected to be used for enabling high-speed Internet access to users who are out of the mmW base station coverage, as well as for offloading traffic onto D2D links if there are users who are interacting located in the proximity.

In addition to the throughput demands, emerging IoT applications *require enhanced connectivity* for peer-to-peer users interaction and for extensive sensor deployments, which are expected to become fundamental elements of smart cities and communities [3, 37, 59, 69, 86]. Moreover, due to size constraints, sensor devices are often not able to connect directly to microwave mobile networks which will notably change the existing paradigm. The connectivity challenges can be addressed using mesh networks and on-demand access points, e.g., access points installed on drones.

Both of the proposed approaches imply the support of local networking among devices even if an Internet connection is unavailable. Conventional networking technologies utilize device identifiers for routing and switching. The most common of those is the Internet Protocol (IP). However, in the case of mesh networking on the edge, it is challenging to assign IP addresses among the devices due to the continuous dynamics at the channel level (topology of the network continuously changing). Moreover, to enable protection from the malicious intent of some users, it is fundamentally important to enable network identification of the devices. Currently, identification relies on virtual identifiers (e.g., IMEI, MAC address) which do not address reliability concerns in the era of IoT [1]. Therefore, it is essential to develop efficient and reliable methods addressing this challenge; otherwise, the potential of multi-hop mesh networking may become useless.

[1] https://www.itu.int/dms_pub/itu-t/opb/tut/T-TUT-CCICT-2015-PDF-E.pdf

3 ROBUST NETWORK IDENTIFICATION SYSTEM FOR IOT ELECTRONICS

This chapter considers a reliable method for the identification of IoT devices. This method may serve as a technological enabler for advanced IoT service management and lead to the elimination of concerns related to counterfeit and stolen IoT devices.

3.1 Analysis of existing identification systems

The taxonomy of existing identification methods includes two major classes:(i) virtual and (ii) physical. The first class includes software properties capturing (fingerprinting) and identifiers recorded in the memory of devices. Fingerprinting is widely utilized for capturing users in context advertisement applications. Such methods may employ an IP address, GPS data, OS version, battery status, display resolution, used languages. These characteristics identify devices reliable; however, the reliability of such a method further increases with a higher number of characteristics included in the fingerprinting. In a case of simple IoT electronics, the number of parameters which can be used for fingerprinting is very limited, thus such an approach is not appropriate.

The identifiers prerecorded in the memory of devices allow identification for most of the devices, including primitive IoT electronics. Such identifiers are stored in the memory of devices by their manufacturer [31, 78]. For instance, MAC-address or IMEI. Besides the widespread use of such identifiers, they do not provide reliable identification due to their vulnerability to tampering and copying [20, 27, 60, 73].

These physical identification methods utilize the uniqueness of the hardware's properties. Particularly these methods widely rely on integrated clock skew estimation, radio signal individuality, and flash memory degradation. The clock skew estimation employs timestamps of network packets (e.g., ICMP) [44]. However, as it was demonstrated in [16] the timestamps could be easily altered or disabled, which makes this method inefficient.

The utilization of radio signal individuality for device identification was derived from military technologies, where signal individuality was used to distinguish both friendly and enemy devices [16]. Further, this technology was implemented in mobile networks for blocking unauthorized devices [68]. In [62], the authors demonstrated differences among signals of devices operated in accordance with 802.11 specifications. The provided experiments confirmed that radio signal properties allow for the identification of devices in a laboratory environment. However, in real deployments, radio signals were affected by the external environment which significantly limits the applicability of such methods for robust identification of IoT electronics.

We can conclude this overview of identification systems by positing that the development of identification methods for IoT electronics is still an open challenge.

3.2 Utilizing flash memory degradation process for identification of an electronic device

Hardware-based identification methods are significantly enhanced by the theory of physical unclonable functions (PUF) [66]. More specifically, PUF theory provided a rationale for the new identification methods which rely on the unique properties of a memory chip. A notable application of this technology is introduced in [81] and [38], where authors used unique variations of electric current in NAND chips in distinguishing devices. However, the electric properties of memory chips are very sensitive to external conditions (e.g., temperature, humidity) and may significantly change during the exploitation period. This proposed method only allows for the identification of devices within a limited timeframe.

Another identification method enabled by the PUF concept employs the memory chip degradation process due to the degradation of minor segments of a NAND chip as they stop operating properly [36]. The distribution of the degraded segments is rather unique among devices and promises a technological foundation for reliable identification of devices. However, the repeatability of such an identification method has yet to be evaluated. Additionally, NAND chips include presets to avoid using broken segments [67] and not widespread in simple devices (the majority of

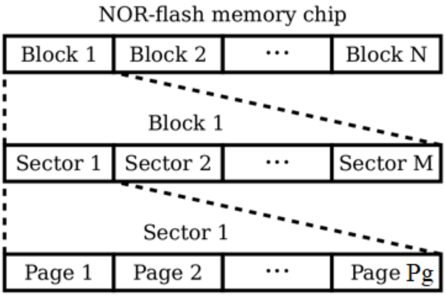


Figure 3.1 Architecture of a NOR-flash memory

IoT devices). Thus the efforts were focused on considering NOR chips. The stated exploitation period of NOR chips is approximately ten times lower, which leads to faster degradation. Moreover, this type of chip is widely used in IoT electronics (e.g., for storing firmware).

Architecturally, a NOR chip is an array of memory cells (Fig. 3.1.), which is capable of storing up to several bits of data. For recording (or erasing) information, a cell's charge should be alternated [71].

Every cycle causes irreversible changes in the chip, which results in the formation of bad-cells. Such a process also is also referred to as the degradation of a chip. The bad-cells lose the capability to alternate their charge during erasing or recording procedures [11, 67].

The pattern of bad-cells distribution (S) is relatively unique due to the significant number of possible cell combinations. Thus it can represent a unique identifier for the chip. If one sector in a memory chip of IoT device can be forcibly degraded and allocated for identification purposes, a single device can be distinguished among a vast quantity.

To enable device identification, a memory chip first should contain degraded sectors with a sufficient number of bad-cells. This can be achieved via forcible degradation of certain areas in chips by multiple rounds of overwriting. After the manufacturer releases the device, the pattern of bad-cells should be stored in a database of produced devices.

During a device exploitation period, the sector of the memory chip used for identification should be directly available for stakeholder access (e.g., telecommunication providers or local authorities). To identify a device, its pattern of bad-cells should be verified in the database of authorized devices. If the entry for the device does not exist in the database, network services will be deemed unavailable.

3.3 Evaluating repeatability among the identifiers

Memory chips degradation is a stochastic process; thus, two or more devices may have an identical distribution of bad cells and consequently not unique identifiers. The probability of such an event depends on:

- volume of the memory segment utilized for identification;
- total number of devices needs to be identified;
- fraction of bad cells in memory sector used for identification.

Let assume that the number of bad-cells used for identification is in a range from m_1 to m_2 and the probability of bad cells forming is uniformly distributed among the sector of the chip. Then, the maximum possible number of unique bad cells patterns given as

$$K = \sum_{m=m_1}^{m_2} C_T^m, \tag{3.1}$$

where T is the total number of memory cells in the sector utilized for identification, m is a number of bad-cells in this sector, and C is the number of unique combinations of the bad-cells. Two absolutely equal identifiers may appear among devices (d), with the probability

$$\delta = 1 - \frac{K!}{K^d (K-d)!}, \tag{3.2}$$

where d is a number of devices which should be identified uniquely.

The degradation process is aimed at reaching a certain number of bad-cells in a chip. As demonstrated in Fig. 3.2, the maximum number of unique bad-cells combinations achievable if the total portion of bad-cells in a sector is about 50 %.

To reduce computational complexity, the probability that two devices have the same patterns of bad-cells followed from 3.2 can be approached utilizing the expan-

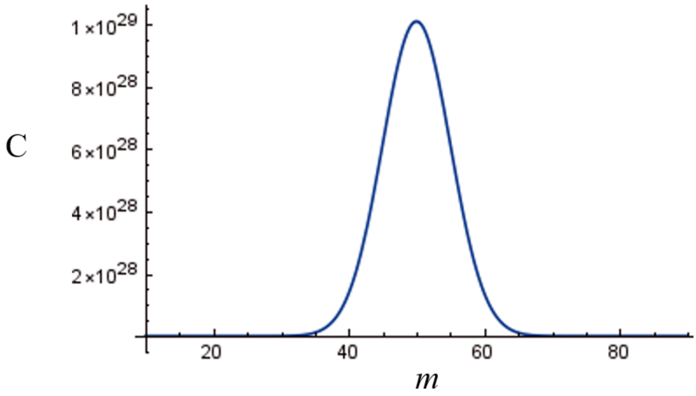


Figure 3.2 Numerical dependence of unique bad-cells combinations on their share in a memory segment.

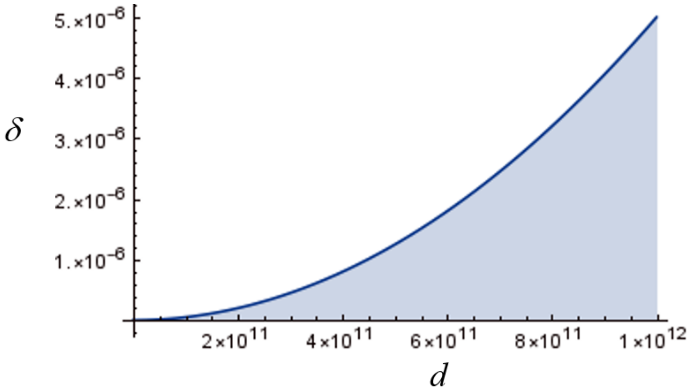


Figure 3.3 The probability that at least two devices have equal identifiers.

sion of an exponential function in a Taylor series

$$\delta \approx 1 - e^{-\frac{d^2}{2\times K}}. \tag{3.3}$$

Let us consider a simple example, which allows us to develop some insights about the probability of such an event. Assume that there are only 100 memory cells in a memory segment allocated for identification, and the number of bad-cells is 50. Thus, following (3.3), the total number of unique combinations is $K = 10^{29}$. Then, the probability that two devices have the same identifiers depend on the total number of devices d needing to be identified, as shown in Fig. 3.3. As one can conclude from the plot, the repeatability of identifiers is extremely low even if the total number of devices to be identified is significant.

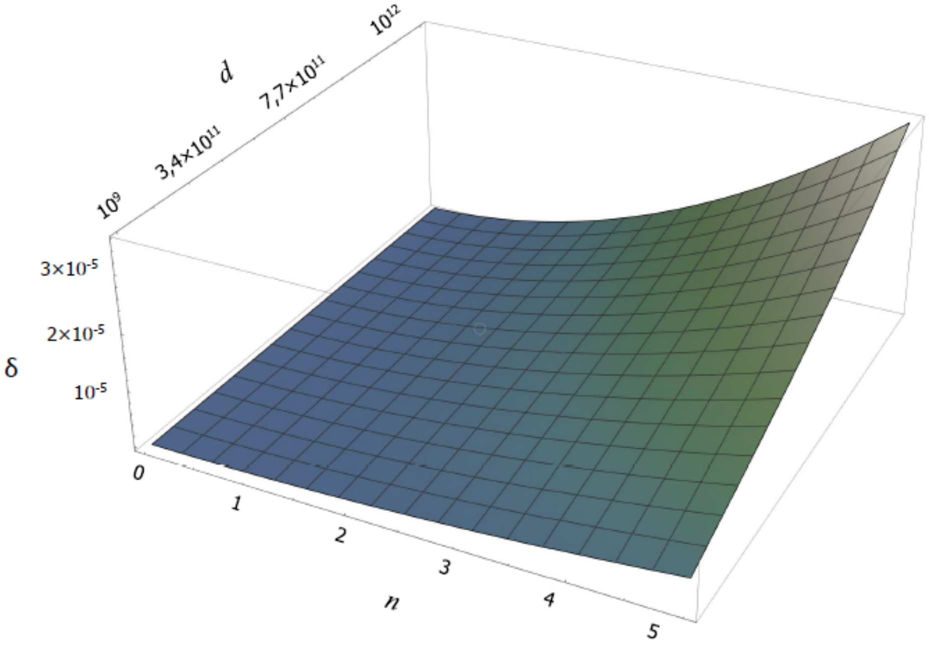


Figure 3.4 The probability of two equal identifiers due to bad-cells developed during the exploitation period.

During a device exploitation phase, its memory chip is subjected to read/write cycles, which may develop new bad-cells. As it was demonstrated in [79], the development of new bad-cells may require up to 10 years of the device exploitation with daily identification. However, it is worth noting that the development of new bad-cells in a memory chip may allow two or more devices with the same bad-cells patterns. The possibility of such an event can be expressed as

$$\delta \approx 1-(e^{-\frac{d^2}{2\times C_T^m}} \times e^{-\frac{d^2}{2\times C_T^{m-1}}}) \tag{3.4}$$

The (3.4) can be generalized in the following form:

$$\delta \approx 1-\prod_{i=1}^{n} e^{-\frac{d^2}{2\times C_T^{m-i}}} \tag{3.5}$$

where n is a number of new bad-cells developed during the exploitation period.

Fig. 3.4 numerically illustrates (3.5) for $T = 100$.

The probability of two equal identifiers is negligibly low even if several new bad-cells developed during the exploitation of a device. It should be noted that the demonstrated results obtained for $T = 100$ are far below the real volume of the memory sector (as the total number of cells in a memory sector is about 32 thousand). Therefore, we may conclude that the appearance of two equal identifiers is statistically infinitesimal even if a total number of devices needs to be identified is in order of trillions.

3.4 Mitigating computational overheads

Computational complexity is one of the important metrics that should be discussed when considering identification methods. Emerging IoT applications require "on a fly" connectivity, and identification should be performed instantly. The method based on PUF is naturally computationally hungry. This section evaluates computational overheads introduced by the proposed identification method (mean time required for the identification) and discusses approaches for their mitigation.

The mean time required for device identification (MTI) depends on many factors, including the size of the registered identifiers database, the size of the identifier used and search optimizations. To reduce the MTI, it can be organized in a hierarchical distributed way. Once a device is identified in a global register, the record related to the devices will be copied in a local register, which resides closer to the edge, and includes a limited number of records corresponding to the devices operating in that area. Moreover, to reduce the mean search time, the entries in a database can be indexed or use conventional identifiers, such as IMEI, as a key. After the entry is found, the device's identity can be verified using the distribution of bad-cells in the allocated sector.

To evaluate the computational overheads for the use of degraded flash memory as an identifier of a device, we executed an experimental campaign utilizing GNU/Octave. The primary purpose of the experiment is to capture the difference between conventional identifiers and the proposed one. More specifically, we considered the time required for a search of the following entries:

- full identifier S;
- searching of S using index P;
- IMEI;

- S using IMEI as a pointer.

For each of the above-listed options, the experiment was performed using a database consisting of 10,000 entries of the appropriate type. For each entry type, we performed 1,000 searches in a database. The summary of the experiment presented in Table 3.1.

Table 3.1 Mean time required of an ID search in a database consist of 10,000 entries

	S	$P+S$	IMEI	IMEI $+S$
Search time t_s, sec	0.3707	0.2721	0.2428	0.2653

The results of the experiment summarized in Table 3.1 indicate relative differences in average identification time. Absolute values can be reduced notably via search engine optimization and implementing acceleration methods. However, the ratio between the time required to find an entry of each type is expected to remain the same.

As it follows from Table 3.1, the average time required for device identification utilizing a unique degradation pattern of NOR flash-memory chip is about 50 percent longer than the conventional approach. However, if an entry search is performed utilizing a pointer or index, the time difference is less than 15 percent.

4 FACILITATING CONNECTIVITY OF WIRELESS IOT DEVICES ON THE EDGE USING DRONES

Unmanned aerial vehicles allow for the physical flexibility of network infrastructure. International standardization bodies considering a UAV equipped with wireless access gateway as a primary enabler for on-demand coverage in millimeter-wave range. The access points installed on drones may also notably enhance the performance of IoT applications on the edge by providing additional connectivity capabilities.

4.1 UAV-based gateway for passive sensor networks deployed over large area

Wireless sensor networks popular for environmental monitoring, as a part of emerging IoT concepts, e.g., Smart Cities and Communities. Technologically, such systems often rely on passive sensors that extensively utilize modern energy harvesting solutions.

The recent achievements in nanomaterials allow for the development of very small passive sensor devices, which significantly improve their applicability. It is worth noting that such devices are not necessarily a nano-sized device but rely on unique properties of novel materials [5]. The maintenance of a permanent communication channel is a challenging task for such devices, primarily due to energy constraints. Recent scientific publications considered communication problems, from the perspective of BAN. Meanwhile, other deployments are rarely discussed.

In this section, we consider data gathering from wireless nanosensor networks deployed over large areas, e.g., agricultural fields, oil pipelines, or construction sites,

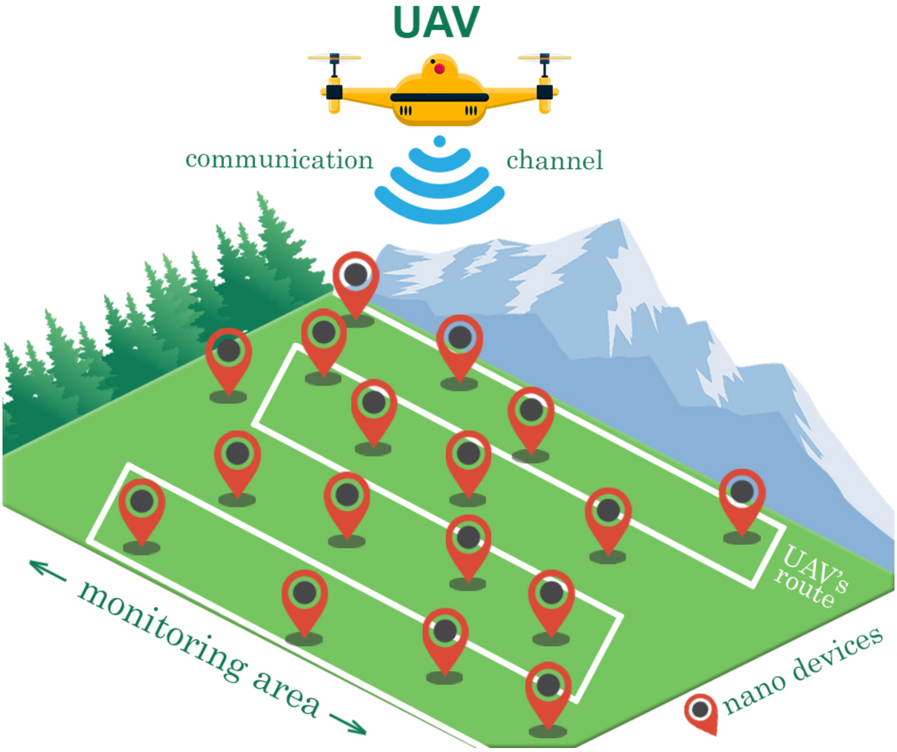


Figure 4.1 Acquiring data from sensor network

utilizing a wireless gateway integrated into a UAV. As an alternative to the terrestrial machines, UAVs allow faster delivery of the gateway to the area of interest, as well as more flexible routes. More specifically, in this section, we are conceptualizing both the wireless energy transfer and data transmission when utilizing a UAV-based THz-frequency wireless gateway.

4.1.1 Acquiring the data from passive sensors using UAVs

UAVs may notably improve connectivity with multiple distributed IoT devices, providing on-demand coverage. This option is specifically relevant for nanonetworks, where the communication distance among devices is up to several meters. The UAV, which flies over the sensor network deployment, may gather sensory data as illustrated in Fig. 4.1.

During the flight, the UAV radiates EM waves, which can be used by sensor nodes

Pre bits	Start byte	Sensor model	Data	Ending byte
2 bytes	1 byte	2 byte	8 bytes	1 byte

Table 4.1 Structure of a dataframe utilized by the nanosensor

Data received from a sensor			Data added by a gateway (UAV)		
Type of sensor	Sensor model	Data	Location	Time	Date
Temperature	XFD3112	34.211	59.903176, 30.491099	12:32:03	22.09.2017

Table 4.2 Structure of the dataframe sent by the UAV to the remote server

for energy harvesting. When a sensor harvests energy, it starts to operate and sends the dataframe back to UAV. An example of such a dataframe is shown in Table 4.1. Further data from a sensor can be supplemented by the UAV, e.g., by adding GPS coordinates or timestamps (Table 4.2).

Sensors are capable of accumulating the required energy if a UAV radiates them for a certain period. Thus it will depend on the velocity of UAV, distance to and transmit power of the EM wave source, the frequency used, and obstacles.

Sensors that are in the coverage area of the UAV may communicate simultaneously, which may cause interference and collisions. To avoid those, we assume that the MAC relies on frequency division, with the carrier frequency of each sensor predefined during the manufacturing process. On the other hand, the gateway installed on the UAV operates in a wide frequency range. The notable drawback of such medium access is interference between sensors. However, if the frequency range is wide enough, the probability of interference is low. More specifically, the probability of two sensors having the same carrier frequency can be assessed using 4.1.

$$\delta = 1 - \frac{N_f!}{{N_f}^d (N_f - d)!}, \tag{4.1}$$

where N_f is the number of central frequencies used in the system, d is the top border on the number of the sensors simultaneously communicating with the UAV.

Assuming $d = 10$ and $N_f = 1000$, the probability of two sensors having the same carrier frequency will be 0.045, which proves high reliability for the considered MAC mechanism.

4.1.2 System model

The total time (t_{tot}) between the moment when a sensor enters the coverage of an aerial gateway and the moment the sensory report is sent, consists of the time required for energy harvesting t_{ch}, and time for making measurements and sending a sensor report back to the gateway t_{sg}:

$$t_{tot} = t_{ch} + t_{sg} \tag{4.2}$$

For further calculations we assumed that $t_{sg} = 10ms$, which corresponds to mean maximal time for such systems, while the time required for energy harvesting is following from (4.3):

$$t_{ch} = \frac{E_{tot}}{E_{rx} r_c} \tag{4.3}$$

where E_{tot} — total amount of energy required for sending a sensor report; E_{rx} — energy, harvested by sensor per second; r_c — coefficient representing the efficiency of energy harvesting (transformation of electromagnetic energy to electric current). For further calculations we assumed $r_c = 0.5$.

Energy costs of a passive sensor include (E_S) energy costs on maintaining the operation of a sensor until it finished measuring the required parameter; (E_p) measurement costs, and ($E_{packet-tx}$) communication expands:

$$E_{tot} = E_S + E_p + E_{packet-tx} \tag{4.4}$$

To specify the energy consumption of sensors, we utilized specification described in [64], where $E_p = 0.73\ \mu J$ and $E_S = 1.06\ \mu J$.

Following [42], the energy required for sending a data packet can be expressed as:

$$E_{packet-tx} = N_{bits} W E_{pulse-tx} \tag{4.5}$$

where N_{bits} — the number of bits in the packet; W — the code weight (the mean expected ratio of "1" and "0" bits), $E_{pulse-tx}$ — the energy required for sending of "1"bit. Following the coding scheme from [41], we take $W = 0.5$. Therefore, energy for transmission of "1"bit to a distance of 10 mm is $E_{pulse-tx} = 1\ pJ$ [40].

The signal power on the receiver given as:

$$P_{rx} = \frac{P_{tx}G(f)}{A(f)} + N_{mol}(f) \tag{4.6}$$

where P_{tx} — power of the transmitted signal; $G(f)$ — antenna gain parameter; $A(f)$ — total attenuation ratio; $N_{mol}(f)$ — molecular-based absorption noise.

The signal attenuation constitutes of free-space propagation losses A_{fspl}, and molecular absorption A_{mol} [15, 45, 72].

The free-space propagation losses can be expressed as:

$$A_{fspl}(f) = (\frac{4\pi f d}{c})^2 \tag{4.7}$$

where d — distance of communication; f — carrier frequency; c — speed of light.

The THz frequency range is featured by molecular absorption, which is caused by vibrations and the rotation of particles in the medium. The molecular absorption phenomena can be observed if the EM signal frequency is close to the resonance frequencies of molecules. In this case, molecules absorb the energy of the signal and produce molecular noise $N_{mol}(f)$ of the same frequency as signal [4].

$$A_{mol}f = e^{k(f)d} \tag{4.8}$$

where k is an absorption coefficient, which determines the ability of a molecule for energy absorbing [39], and does not depend on the communication distance. In this work the molecular absorption coefficient estimated utilizing the HITRAN database [1, 63] for the following conditions: $H_2O = 1.860000$ %, $CO_2 = 0.033000$ %, $N_2O = 0.000032$ %, $O_3 = 0.000003$ %, $CO = 0.000015$ %, $CH_4 = 0.000170$ %, $N_2 = 77.206000$ %, $O_2 = 20.900001$ % at a temperature of 296 K and a pressure of 1 atm.

As it following from [39], when the absorption coefficient is higher than 5.5 %, the molecular absorption noise N_{mol} reaching the maximum value of —203.89 dB/Hz ($\approx 10^{-20}$ W/Hz). Considering a bandwidth of 100 kHz per one sensor and frequency range of 0.1-0.15 THz, we can calculate that molecular noise value will be

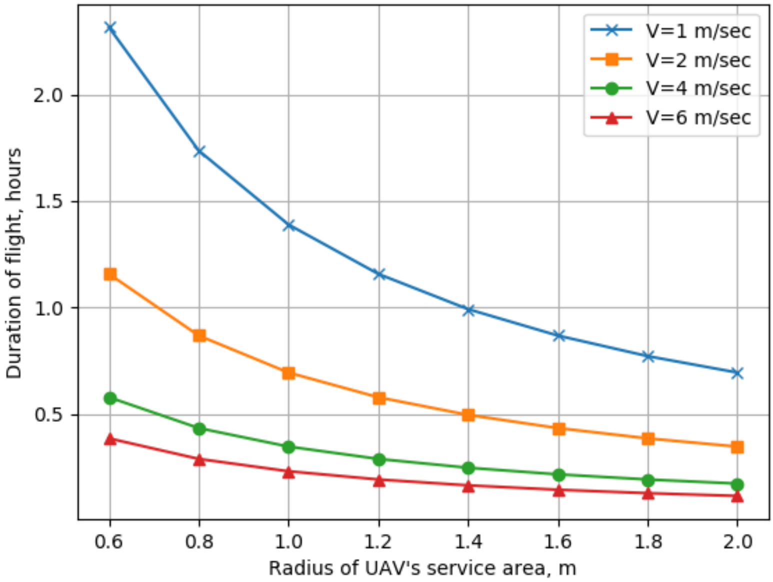


Figure 4.2 Time required for serving a hectare by one UAV

about $1fW$ which is negligibly small. Therefore, 4.6 can be simplified as follows:

$$P_{rx} = \frac{T_{tx}G(f)c^2}{(4\pi f d)^2 e^{k(f)d}} \tag{4.9}$$

For convenience, we summarized the main parameters used in our model in Table 4.3.

To collect data from sensors, the UAV will fly with a constant velocity v over the area of interests. Assume the flight altitude is 2 meters. Such an altitude allows for the elimination of the effect of flight stability issues on our results. To enable wireless energy transfer and communication, the UAV is equipped with an antenna array, whose beamwidth defines the ground service area. For simplification, the coverage area can be considered as a circle with radius R. The broader the coverage area, the faster the area is served, as it is demonstrated in Fig. 4.2. Remarkable, it also can cause additional losses because EM power density over the coverage area will be reduced. To better highlight this trade-off, we further consider varying R, with a constant transmitter power.

The number of packets received by the UAV from a sensor depends on the time t_{ex} when the sensor is in the coverage area:

$$t_{ex} = \frac{D}{v} \tag{4.10}$$

where D ($D \leq 2R$) — communication range.

The measurement will be performed if $t_{ex} \geq (t_{ch} + t_{sg})$. While, if $t_{ex} \geq n(t_{ch} + t_{sg})$, then one sensor becomes capable of performing multiple measurements ($'n'$ repetitions) because the energy harvesting cycle can be completed more than once.

Following the [82] the distance D can be determined as follows:

$$D = 2R(\frac{\pi}{2}F(S)) \tag{4.11}$$

$$F(S) = \frac{2}{\pi}\arcsin(\frac{S}{2}) + C \tag{4.12}$$

where $F(S)$ — unit circle probability density function; S — the distance between two points on the circle which outlined UAV's coverage area; C — constant.

Parameter	Identification	Value
Transmission power of a THz-reader	$P_{tx} \cdot G$	1 W
Antenna of the THz-reader		Directional [32, 84]
Antenna of a sensor		Isotropic
Frequency range	f_1 - f_2	0.1 - 0.15 THz
Bandwidth per sensor	Δf	100 kHz
Altitude of a UAV	h	2 m
UAV's coverage area radius	R	[0.8, 1.2, 1.6, 2] m
Velocity of a UAV	V	[1, 2, 4, 6, 8, 10, 12] m/sec
Absorption factor (0.1 THz)	k_1	2.58×10^{-5} m^{-1}
Absorption factor (0.15 THz)	k_2	1.01×10^{-4} m^{-1}
Mean energy required for transmission of one data packet	$E_{tot,min}$	2,27 μJ
Time required for creating and sending a packet with sensor's report	t_{sg}	0.01 s

Table 4.3 Parameters used in the model

4.1.3 Performance evaluation

To evaluate the performance of the considered system, we executed a simulation campaign utilizing a system-level network simulator (WinterSIM) developed at Tampere University. Our interest focused on the dependency of losses (when measurement from sensors was not received) on the velocity of a UAV. We evaluated this metric for two frequencies. For the simulation, we assumed that the UAV flight path did not overlap with the previously served territory.

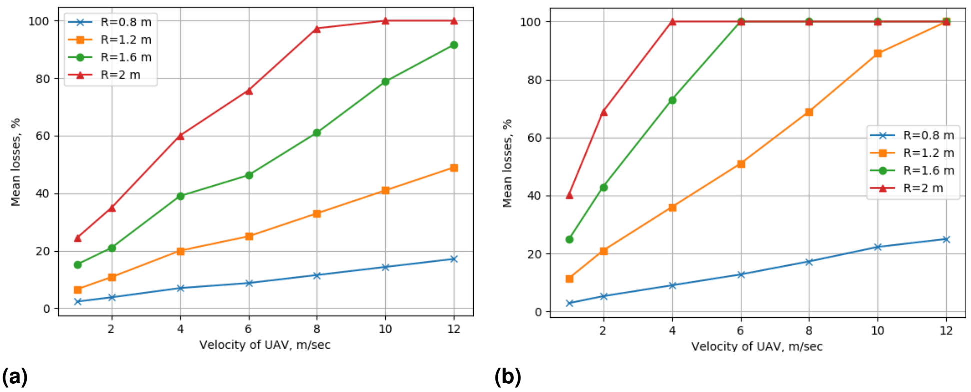
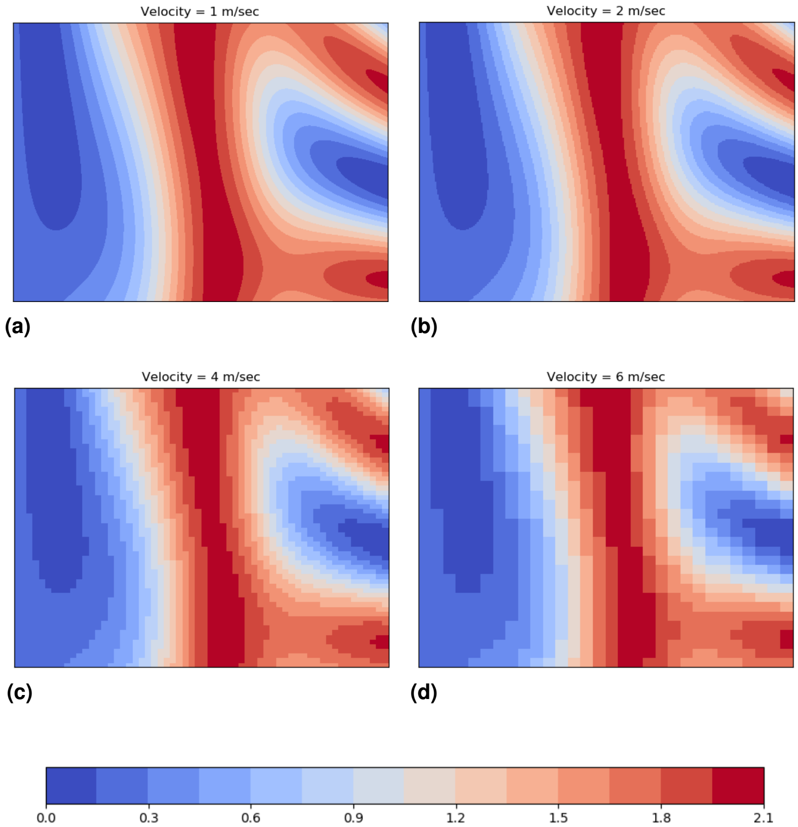


Figure 4.3 Packet losses a) f = 0.1 THz, b) f = 0.15 THz

The numerical results presented in Fig. 4.3, demonstrate that an acceptable level of losses can be reached with low flight speeds and narrow beamwidth. Increased velocity causes higher losses because sensors on the edge of the coverage area are not able to harvest the required energy. A greater beamwidth leads to reduced power density on the area of coverage, which also leads to higher losses. In addition, our results demonstrate that higher carrier frequencies have notably higher losses.

To better illustrate the effect of losses on the accuracy of the field monitoring system, we compiled measurement maps (Fig. 4.4) utilizing data about losses obtained during the simulation campaign. The maps visualize the decrease of density due to higher losses.


Velocity = 1 m/sec

Velocity = 2 m/sec

Velocity = 4 m/sec

Velocity = 6 m/sec

Figure 4.4 Compilation of measurements maps for different UAV velocity (f = 0.1 THz, R =2 m)

4.2 Enabling latency-sensitive services via a flying network

One of the relevant scenarios for IoT on the edge is the lack of communication due to the unavailability of the mobile network coverage. This challenge can be addressed utilizing a fleet of interconnected UAVs – a flying network.

4.2.1 The flying network architecture

The flying network consists of UAVs that are participating in multi-hop communication and provide wireless access (coverage) for the terrestrial segment. Optionally, some of the UAVs may also have backhaul. The UAVs are capable of prompt on-demand network deployment. If the number of UAVs constituting the flying network is significant, the communication will require multiple relaying hops, which notably increasing the latency of communication. To address this issue, we consider the two-level architecture of a flying network (Fig. 4.5), which enables lower latency for end-to-end communication for latency-sensitive services.

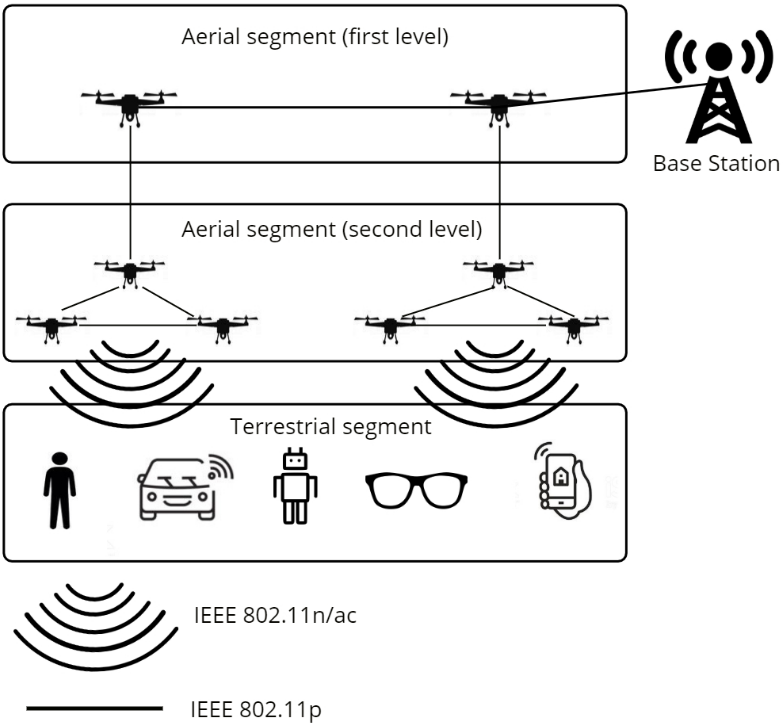


Figure 4.5 Two level architecture of flying network for latency sensitive services.

The considered architecture utilizes clustering among lower-level UAVs and cluster heads at a higher level. The communication beyond a cluster is performed via these cluster heads, which promises a lower number of hops and consequently lower latency.

In this work, we assumed that communication between the terrestrial segment (IoT devices) and UAVs (the devices may communicate only with the lower level UAVs) is performed using the IEEE 802.11n/ac standard which is widely supported by existing wireless devices. The communication among UAVs relies on the IEEE

802.11p standard. The communication among the cluster heads utilize an advanced mode of IEEE 802.11p technology with a distance increased up to 750 m [7, 18, 61].

4.2.2 Continuous connectivity method for ensuring quality of communication

The major challenge one faces when utilizing a dynamic multi-hop network for latency-sensitive applications are slow handovers. To enable seamless connectivity, the handover process needs to be enhanced. A conventional handover performs the following algorithm:

1) exploring access points (APs) available for association;

2) selecting one of the available APs and associating with it;

3) (re) establishing a communication session.

The user authentication mechanisms used in WiFi are relatively slow, which significantly contributes to high latency during handover. For the considered scenario authentication process executes as shown in Fig. 4.5.

The handover process starts with exploring the new AP and deciding if the UE needs to be reassociated. This step may take several seconds; however, it can be performed in a background mode and does not require termination of the current session. Furthermore, authentication with a new AP should be performed. After that device, the UE can establish a logical connection with the AP. The most time-demanding phase starts after the association is done. In this phase, the AAA server authenticates the UE, which allows for establishing a session between UE and AP. The handover process is finalized by four-way handshaking between the UE and the AP.

The total time required for the handover process consists of the exploring phase time (t_{scan}), probing and initiating authentication time (t_{auth}), association time (t_{asso}), AAA delay (t_{1x}) [34], and four-sided handshake (t_{4way}).

To accelerate the most time demanding phase in the handover, [14] proposes a modification of the EAP-SIM protocol (RFC 4186) in order to reduce time costs. According to the EAP-SIM, when a UE moves between APs, it should re-identify itself. The basic idea of the proposed modification implies that the new access point is starting to prepare for the handover in advance, avoiding the necessity to communicate with the remote AAA server during actual association. Al required credentials

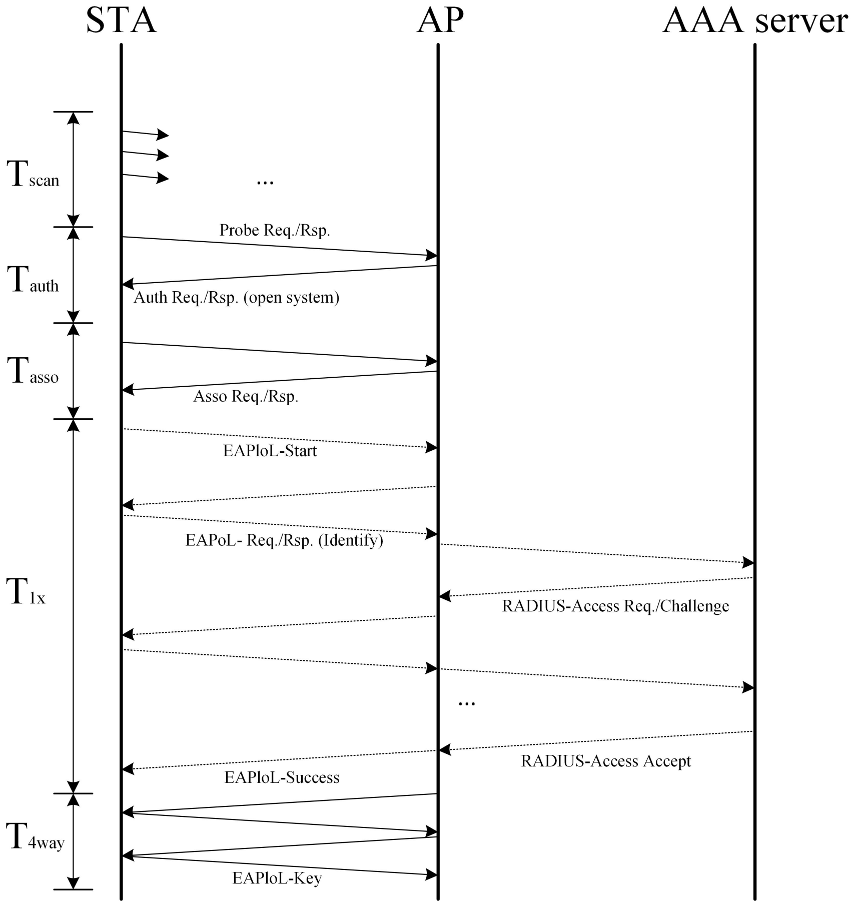


Figure 4.6 Handover process [14]

can be loaded to the new AP while the UE is still connected to the previous AP. The experimental results demonstrated that such an approach allows reducing handover time on about 55 ms.

4.2.3 Service quality assurance and selected numerical results

The latency-sensitive service can be represented by two UE interacting with each other via the chain of UAVs. Assume that the threshold of the acceptable latency is 100 ms. The service provisioning quality can be modeled using the Queuing Theory model presented in Fig. 4.7.

The average packet transmission delay measured on the MAC level can be ex-

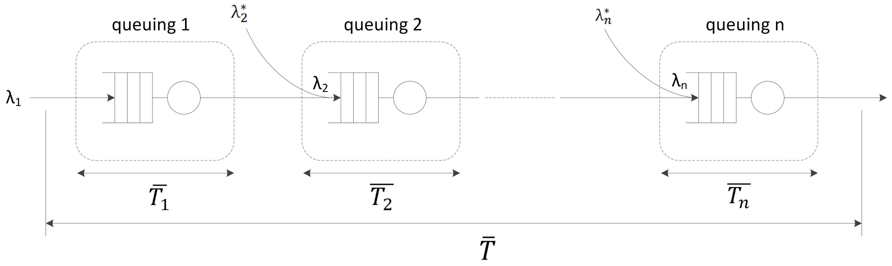


Figure 4.7 Queuing model for the flying.

pressed as:

$$\Delta t_{net} = 2 \cdot t_a + 2 \cdot t_c + (n-1) \cdot t_h \tag{4.13}$$

Where t_a – delay between UE and UAV; t_c – delay between UAVs of the same cluster; t_h – delay between the cluster heads; n - number of cluster heads in the route.

Given an M/M/1 service model for each UAV, and assumed that the incoming flows to each UAV have the same properties, the service time follows an exponential distribution, and the intensity of the load is determined as:

$$\gamma_i = \rho_i = \frac{\lambda_i}{\mu_i}(Erl) \tag{4.14}$$

Where: λ_i is a rate of incoming requests (1 / ms) and μ_i is a service rate (1 / ms).

The time spent for processing of requests in the system can be calculated using equation (4.15):

$$\Delta t_{proc} = w_i + t_i = \frac{t_i}{1-\rho_i} \tag{4.15}$$

Where: w_i is the average waiting time in the queue; t_i is an average duration of service session.

Consequently, it is possible to calculate the average duration of service requests by the formula (4.10):

$$t_i = \frac{L}{b_i} \tag{4.16}$$

Where L is the average length of one packet (bit); b_i - the average data transfer rate (bit/ms).

To illustrate performance of the proposed flying network architecture, we per-

Table 4.4 Simulation parameters.

Parameter	**Value**
Technology used for communication between aerial and terrestrial segments	IEEE 802.11n
Technology used for communication between UAVs	IEEE 802.11p
Number of UAVs in one cluster	7
Number of clusters	$\{2 \dots 37\}$
Service model	M/M/1
System load coefficient, ρ	$\{0.1, 0.2 \dots 0.9\}$
Data packet payload size, Bytes	1024

Table 4.5 Maximum number of UAV clusters capable to support required quality of service

System load coefficient, ρ	**Maximum number of UAV clusters**
0.1	65
0.2	58
0.3	50
0.4	43
0.5	35
0.6	28
0.7	20
0.8	13
0.9	5

formed a simulation campaign in WinterSIM. The main simulation parameters are summarized in the Table 4.4.

Selected simulation results, which are summarized in Table 4.5, show the dependence of the number of UAV clusters, which are providing the required quality of service ($\leq 100ms$), on the system load coefficient. Further, Fig. 4.8 indicates how the average e2e latency experienced by users increases with a larger number of UAV clusters, and $\rho = 0.5$.

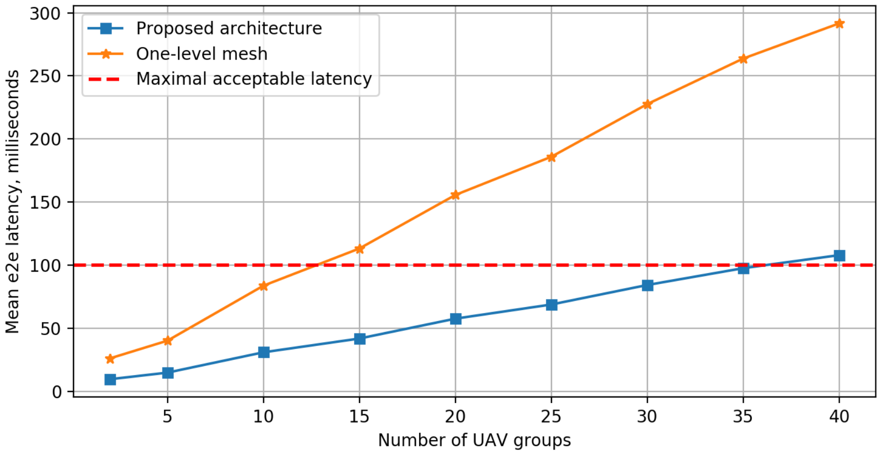


Figure 4.8 Dependence of the e2e latency from number of UAV clusters

5 UTILIZING MILLIMETER WAVE MESH TECHNOLOGIES FOR BROADBAND WIRELESS ACCESS ON THE EDGE

Emerging IoT applications heavily utilize media services that demand high throughput on the access level. In 5G systems, this challenge is attempted to address by millimeter-wave (mmW) technology.

With undeniable advantages, mmW networks also bringing new challenges, such as significant path loss, and sensitivity to blocking obstacles (due to short wavelength) [19]. As a consequence, mmW links are narrowly targeted. Additionally, mmW links are highly susceptible to atmospheric and molecular influence [33]. Together, all these factors considerably reduce the reliability of mmW communications.

The extreme throughput demands on the edge require extending conventional wireless access utilizing multi-hop D2D mesh networks. Related to mmW, such meshes are expected to improve coverage of mmW, as well as communication reliability [2].

5.1 Millimeter wave mesh networks in environments with dynamic blockages

The benefits of mmW meshes have been investigated for the outdoor use cases [51, 58, 70]. These works confirm that meshes may considerably improve connection reliability and throughput for demanding IoT devices on the edge. However, indoor mmW deployments bring new challenges caused by increased dynamic blockages. In indoor scenarios, links can be blocked by the internal constructions of a building, and by mobile objects. Considering the mobility of mesh users, indoor scenarios are feature continuously changing the connectivity between the network participants.

To address the high throughput demands of emerging IoT applications in environments with dynamic blockages, this chapter considers network connectivity and the throughput characteristics of mmW mesh networks deployed in an indoor environment. The considered scenario is characterized by the mobility of users, blocking dynamic, 3GPP propagation model, and multi-connectivity operation of the users in the mesh. For a better illustration of the dynamic blockage environment scenario, later in this chapter, we consider an illustrative use case of a fire suppression mission.

5.2 System modeling

The abovementioned illustrative scenario assumes a fire suppression mission in a single-floor indoor layout. The firefighting team extensively uses media-heavy applications enabled via a mmW mesh network. Communication with the command center is provided via a gateway (access point) installed on the border of the layout (e.g., window).

The users' mobility of the considered mesh network follows the random direction mobility (RDM, [50]) model. According to RDM, a user initially selects the random direction of movement in the range $(0, 2\pi)$ and then moving in the selected direction with a constant velocity of v_B. The user stops after an exponentially distributed time period with the parameter $\gamma = 1/E[\tau]$, where τ is the mean movement duration. After a user stops, the procedure is repeated to select a new direction of movement. While moving, if a user collides with an obstacle, it chooses a new random direction of movement.

The received signal power at the users devices can be obtained from

$$P_R(x) = P_T G_T G_R - PL, \tag{5.1}$$

where: P_T is the transmission power, PL is the path loss, G_T, G_R are the antenna gains of transmitter and the receiver, respectively. These gains depend on the antenna array used, and follow the beamforming model which is introduced later in this chapter. Following 3GPP TR 38.901, the mmW path loss in dB for the line-of-sight (LoS) links is given by

$$PL_{InH-LOS} = 32.4 + 17.3\,\mathrm{lg}(d_{3D}) + 20\,\mathrm{lg}(f_c), \tag{5.2}$$

where f_c is the carrier frequency, d_{3D} is the distance between the antennas of two communicating nodes (given in 3D space).

For the non-LoS (NLoS) links, the path loss can be expressed as

$$PL_{InH-NLOS} = \max(PL_{InH-LOS}, PL'_{InH-NLOS}), \tag{5.3}$$

where

$$PL'_{InH-NLOS} = 38.3 l g(d_{3D}) + 17.3 + 24.9\,\mathrm{lg}(f_c). \tag{5.4}$$

In the considered scenario, it is assumed that linear antenna arrays are utilized by all users. Then, half-power beamwidth (HPBW) of the array, α, is given by [10]

$$\alpha = 2|\theta_m - \theta_{3db}|, \tag{5.5}$$

where θ_{3db} is the half-power point and $\theta_m = \arccos(-\beta/\pi)$ is the array maximum, β is the array direction angle. Letting $\beta = 0$, then upper and lower 3-dB points are

$$\theta^{\pm}_{3db} = \arccos[-\beta \pm 2.782/(N\pi)], \tag{5.6}$$

where N is the number of elements in the antenna array.

Following [10] the mean antenna gain over HPBW is

$$G = \frac{1}{\theta^{+}_{3db} - \theta^{-}_{3db}} \int_{\theta^{-}_{3db}}^{\theta^{+}_{3db}} \frac{\sin(N\pi\cos(\theta)/2)}{\sin(\pi\cos(\theta)/2)} d\theta. \tag{5.7}$$

Finally, the blockages model includes (i) building constructions; (ii) user self-blocking; and (iii) blockage by dynamic objects. The blockages by inherent building layouts are taken into account in the 3GPP propagation model, which is used (it includes the mean probability of blockages for indoor). Self-blocking occurs when a users' device cannot beamform its signal towards the intended receiver. The third blocking factor is presented by a spatially-temporal blockage model. Following this model, the blockers' appearance is uniformly distributed over the considered layout, according to a homogeneous temporal Poisson process with the intensity of λ. Every blocker's appearance exists for an exponentially distributed time with the mean $1/\mu$. Notably, this process is inherent of the M/M/∞ type, and the Poisson distribution provides the number of active blockers with the parameter λ/μ. Assuming r_B is the radius of a blocker, the total area of blockers on the floor can be formulated via integral geometry [53]

$$p_C = (1 - f_{C,1})^{\lambda/\mu},\ p_{C,1} = \frac{2\pi S_B}{2\pi(S_A + S_B) + L_A L_B}, \tag{5.8}$$

where S_A is the floor area, and $S_B = \pi r_B^2$ is the blocker radius.

5.3 Evaluating system capacity

For the performance assessment, we used the following indicators: (i) the fraction of time when at least one node is disconnected from the mesh network, (ii) the mean number of disconnected nodes, and (iii) the mean throughput per-node of the mesh network.

In the considered scenario, it is assumed that devices of all users support multi-connectivity [2]. Consequently, a capable node instantly switches to another connection in cases where the current link is unavailable. In this work, the number of simultaneously supported links is referred to as the degree of multi-connectivity (M). Depending on the instant positions of users on the floor, the number of simultaneous links can vary in the range from 0 to M.

The performance evaluation of the considered system was performed in a system-level simulator. This simulator is based on the Stage source code [29, 74]. The simulation of the mesh network was performed over the 3D model of a building floorplan (Fig. 5.1). The 3D model allows for a determination of LoS conditions between two arbitrary points in the environment. In the initial stage of simulation, we executed a coordinate simulation of the users' mobility and blockage dynamics. The simulation process utilized a discrete timeline, where at every iteration, the LoS condition was checked for all the users of the mesh. The data about users coordinates in every iteration, as well as LoS conditions, were recorded to the database.

In the next step, we worked exclusively with the database for evaluating the target metrics. It starts with estimating path loss among all the users. If the signal level is lower than a preset threshold, the simulator considers that there is no direct link

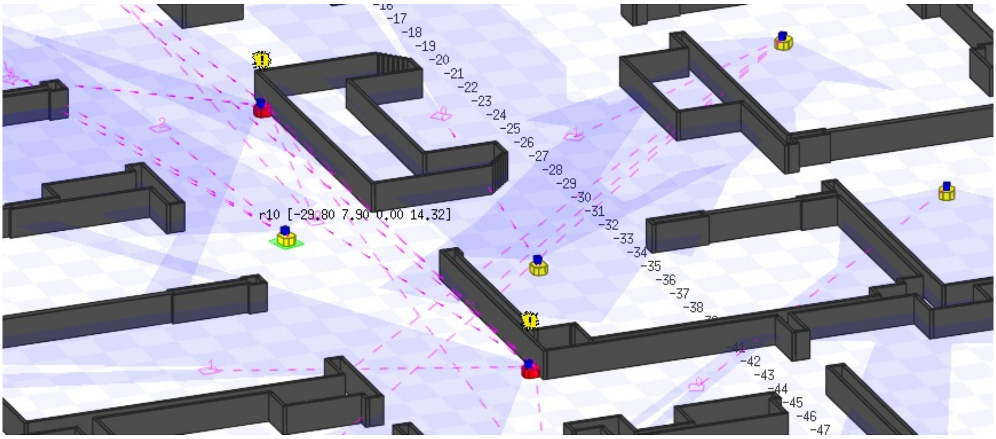


Figure 5.1 Visualization of the simulation

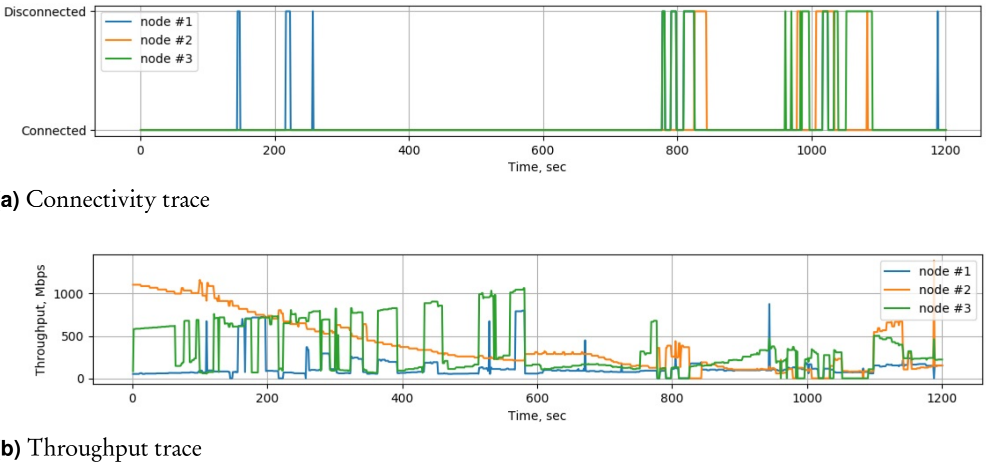
between these users. Contrary, if a direct connection between two users exists, the simulator estimates the throughput utilizing the Shannon formula. In the third step, the simulator estimates overheads introduced by medium access control and routing procedures. These results are in a network-level topology graph, where every link is associated with throughput.

During the simulation experiments, each set of input parameters (referred to as a simulation round), run for 1200 seconds of the simulation time, with a time step of 0.25 seconds. It is assumed that all of the involved processes (arrival, service, blockage) are stationary; the steady state always exists in the system. The starting point of the steady-state period has been defined from the exponentially weighted moving average (EWMA) statistics (weight parameter 0.05) and follows the procedure described in [52]. Table 5.1 summarizes simulation parameters used in our experiments.

Our analysis starts with connectivity and throughput of time-dependent behavior analysis. For this purpose, we used traces of these metrics. An example of such traces is demonstrated in Fig. 5.2. Particularly, the connectivity trace illustrated in Fig. 5.2a, allows us to conclude that disconnection intervals are comparably short but frequent. Such behavior can be explained by short-lived outage events caused by user mobility and dynamic blockage. An example of throughput variations among users is shown in Fig. 5.2b. As it can be observed, the throughput significantly deviates. For some users, the deviations are rather smooth, primarily caused by users' mobility. For other users, the throughput trace is characterized by sharp peaks; these

Table 5.1 Simulation parameters.

Parameter	Value
Carrier frequency, f_c	28 GHz
Antenna array	16 × 16 el. (planar array)
Channel model	3GPP InH
Transmission power	1 W
Receiver sensitivity	-91 dBm
Mean expected area covered by blockers, p_C	0.15
Number of users in the mesh network	$\{8, 10, 12, 14, 16\}$
Velocity of users	1 m/s
Mobility of users	RDM model
Degree of multi-connectivity	$\{2, 3, 4, \infty\}$



(a) Connectivity trace

(b) Throughput trace

Figure 5.2 Example of connectivity and throughput traces

are associated with the blockages (building geometry and dynamic blockers).

Fig. 5.3 the time fraction when at least one node is disconnected from the mesh, as a function of the number of nodes and multi-connectivity degree (M). This metric characterizes an integral measure of connectivity in the network. It demonstrated fundamentally different behavior for different degrees of multi-connectivity. Particularly, for multi-connectivity $M = 2, 3$, the metric increases as the number of users in the network grows. The explanation for this behavior is obvious: the probability that at least one user finds itself in unfavorable position becomes higher with greater number of users in the network, and the number of available simultaneous connections may be insufficient to eliminate this. However, if the degree of multi-connectivity is high enough, the effects of connection diversity enhances the overall connectivity in the mesh. Therefore, we may conclude that multi-connectivity in mmW meshes significantly improves connectivity for scenarios with dynamic blockages. However, the number of simultaneously supported connections should be rather high, which may cause notable control overheads.

Fig. 5.4 shows the mean number of disconnected users depending on the degree of multi-connectivity and the total number of users in the network. For $M = 2, 3$ the network is not able to scale appropriately as the number of disconnected users grows. Further increasing M, allows the mean number of disconnected users to dip below 1. and the mean number of disconnected users decreases with a greater number of users in the mesh network.

Fig. 5.5 indicates the mean throughput experienced by a user in the mesh, depending on the number of nodes and a number of simultaneously supported links.

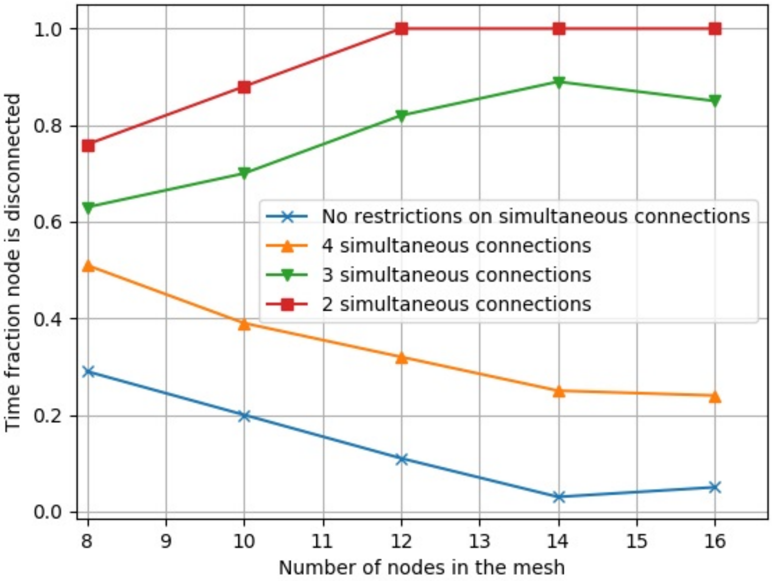


Figure 5.3 Time fraction when at least one user is disconnected

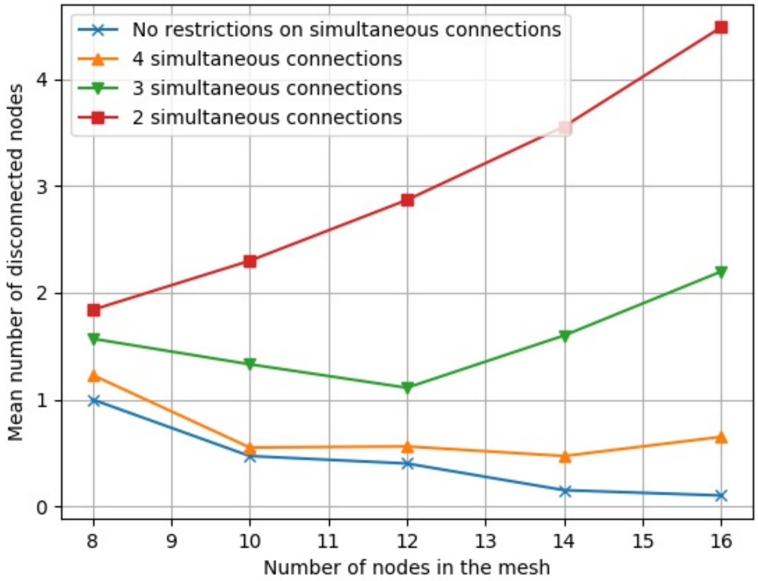


Figure 5.4 Mean number of disconnected users

As it can be noticed, the throughput behavior is qualitatively similar for all the values of M. In a case when there are no restrictions for the number of simultaneously supported links, the mean throughput experienced by the user is 3—5 times higher if compared to the $M = 2$. It is worth noting that while increasing the degree of multi-connectivity, the throughput approaches the higher bound relatively slow, e.g., the mean throughput at $M = 4$ is only a half of that for $M = \infty$. Such a behavior is differ from the one reported for outdoor scenarios in, e.g., [26, 28], where both capacity and outage probabilities grow exponentially with M.

The performance evaluation results can be summarized as follows:

- the use of multi-connectivity operation drastically improves mmW mesh connectivity in indoor deployments, but its effect on the per-node throughput is minor;
- to augment connectivity and throughput in dense indoor mmW mesh topologies, the number of simultaneously supported links needs to be greater than two, thus implying considerable control signaling overheads.

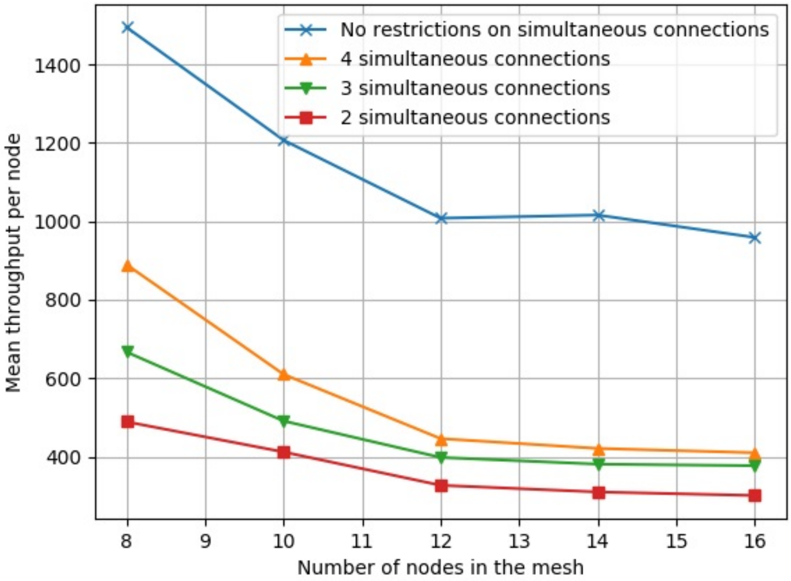


Figure 5.5 Mean throughput per-user

5.4 Improving the mesh reliability using machine learning

This section targets the improvement of the mmW mesh performance utilizing machine learning methods for avoiding blockages. Particularly, we assume that a fire spreads dynamically in the considered layout, and the firefighters not aware of actual locations of fire Fig. 5.6. In addition, the firefighting team utilizes autonomous robots, which are capable of independent movement, and assist the crew if required (e.g., acting as relays for communication if a direct connection between them is unavailable).

Firefighters and assisting robots utilize heavy media application which are streaming data to the command center for the processing which requires high throughput (120 Mbit/s per user) and connection reliability. These requirements expected to be satisfied by the mmW mesh network maintained among firefighters and robots, which using for multi-hop data transfer to the command center. Following the results discussed in the previous section, we also assume that the node of the mesh supports M simultaneous connections with the neighbor nodes.

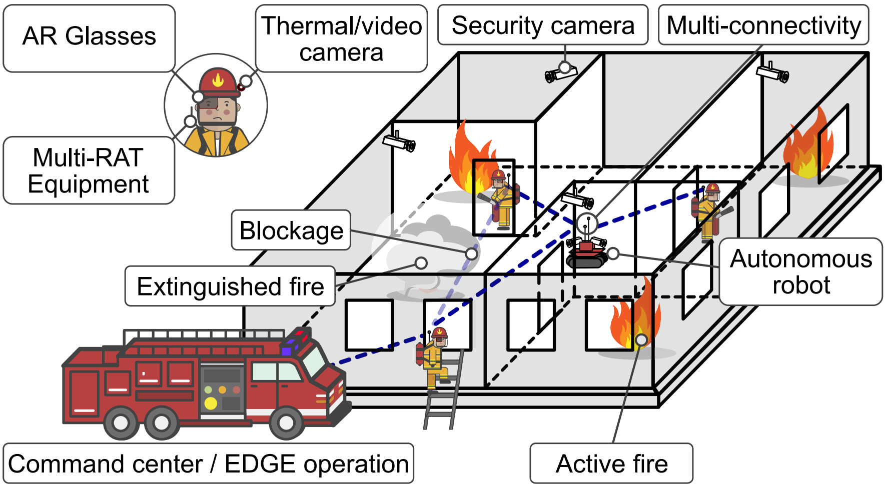


Figure 5.6 Scenario of interest

5.4.1 Artificial intelligence for millimeter wave blockage detection

Presently, computer vision technologies attract remarkable attention in the field of fire detection because of their time-efficient response. Such systems allow the monitoring of large areas with relatively low cost. The majority of early-stage vision-based methods for fire recognition rely on color and shape properties when detecting fire on an image. These methods suffered from false alarms [30]. To enhance these methods, subsequent works proposed to consider motion features. As a result, the rate of false alarms was reduced; however, the efficient operational distance sharply decreased. Recently proposed adaptive algorithms utilize CNN [25] addressing both of the drawbacks and enabling fire detection over long distances with acceptable accuracy.

In this work, we focus on the CNN-based solution to indicate the potential or AI for detecting blockages in our proactive approach. After the fire is detected, we consider two paths of action: (i) if there are alternative links existing for a node whose connection is expected to be interrupted, the traffic flow proactively switching to one of those; (ii) if the node doesn't have alternative links, the robot selects a new location which allows the establishment of a backup connection. Otherwise, the node becomes disconnected from the command center.

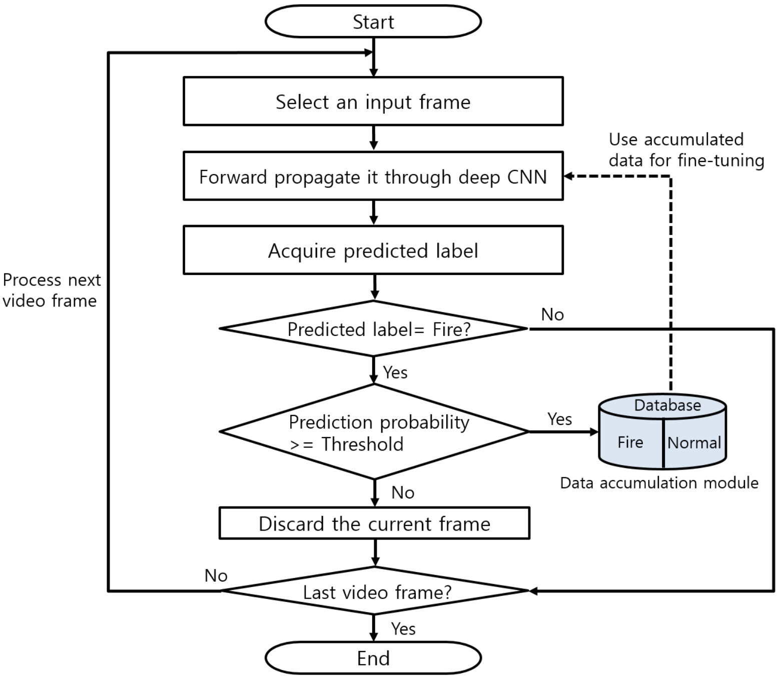


Figure 5.7 Algorithm for fire detection

5.4.2 System design details

Our proactive approach relies on CNN implementation delivered by GoogleNet. Their use of a neural network provides wide functionalities for fire detection, including localization, and semantic understanding of the scene of the fire. As an output, the CNN specifies the piece of video as “Fire” or “Non-Fire”. Our choice is motivated by the high performance of the algorithm for fire detection task [49], and the acceptable size of the model, allowing the deployment of the method on the edge. An algorithm behind the developed fire detection module shown in Fig. 5.7.

Architecturally, our proposed system consists of two network overlays that operate cooperatively. The first overlay utilizes the mmW wireless interface (28 GHz) for high rate communication demanded by the heavy traffic applications. The second overlay relies upon a long-distance wireless technology (IEEE 802.11ah), which provides robust long-range connections among all the nodes of a mmWave mesh [21], which are used for signaling and managing of the mmW mesh operation.

The system operates following the repeating cycle shown in Fig. 5.8. This cycle

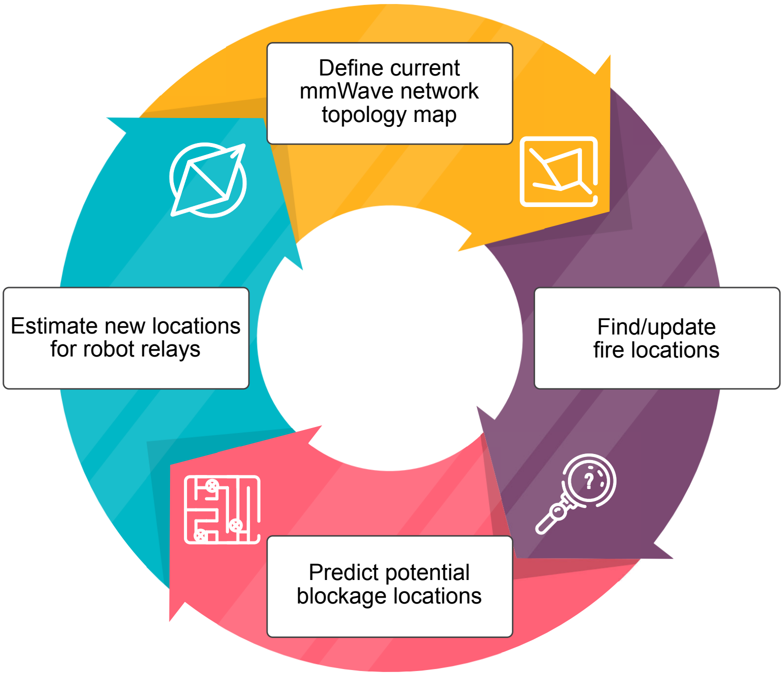

Define current mmWave network topology map
Find/update fire locations
Predict potential blockage locations
Estimate new locations for robot relays

Figure 5.8 System operation cycle

enables timely updating of the information about the mesh operation details. It is worth noting that if the operation dynamic is high, the cycle should be repeated more frequently. The remarkable factors which can affect this dynamic may include the number and density of nodes in the mmW mesh and intensity of the blockages.

The cycle starts with determining current mmW mesh topology. For this purpose, every node sends its coordinates and mmW link-state advertisements (LSAs) via the IEEE 802.11ah interface. Based on the received LSAs, the command center reconstructs the actual mmW mesh topology. Simultaneously, the command center processes video streams from the operational area to detect their fire zones and update their locations by utilizing the available media and sensor information obtained from the fire suppression team members via the mmW mesh. Further, the command center performs a mapping between estimated fire locations, the building plan, and the mmW mesh topology to define which links can be blocked. Finally, based on the information about expected blockages, the system estimates where to move the robot relays to reduce the risk of disconnecting mesh nodes Fig. 5.9.

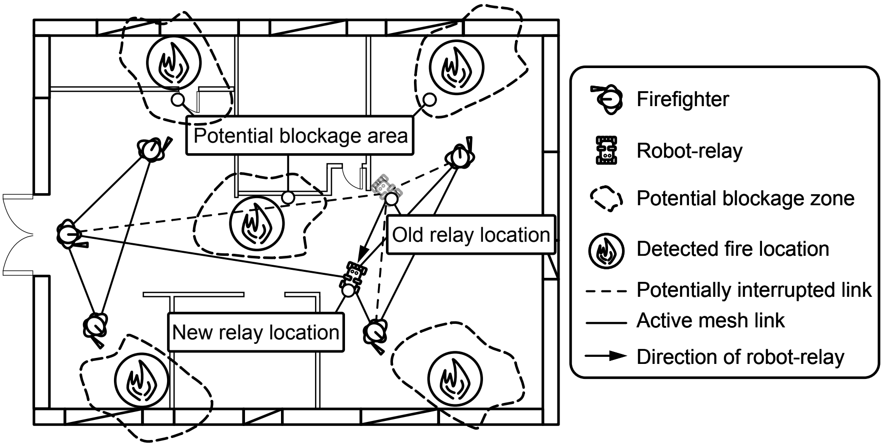


Figure 5.9 Blockage avoidance: top view

5.4.3 Selected numerical results

To explore the benefits of using a proactive blockage avoidance system, we repeated the simulation described earlier in this chapter. To represent the proposed system, we added to the deployment K robots and parameterized blockages (by the temporal intensity of fire locations, the radius of vapor cloud over the fire location, the mean duration of evaporated water after the fire suppression). As it follows from [35], the attenuation caused by water particles causes significant power loss in mmW signals. All new parameters distinguishing this simulation campaign from Table 5.1, are expanded further in Table 5.2.

In the *baseline* scenario, we assume no proactive protocol enabled, which replicates blockage avoidance techniques based on multi-connectivity [17], discussed in the first part of this chapter. While in the *AI-aided* scenario, the mmW mesh utilizes our blockage avoidance method.

To evaluate the gain from using the *AI-aided* method we selected the performance metrics related to mmW system reliability, related to those considered in the first part of this chapter: (i) fraction of time when a node in the mesh is disconnected, (ii) probability of that a number of nodes are disconnected at random time instance, (iii) intensity of node disconnections from the mesh, and (iv) data rate at the AP.

The fraction of time an arbitrarily chosen node is disconnected from the network for both scenarios of interest, as a function of the temporal intensity of fire locations,

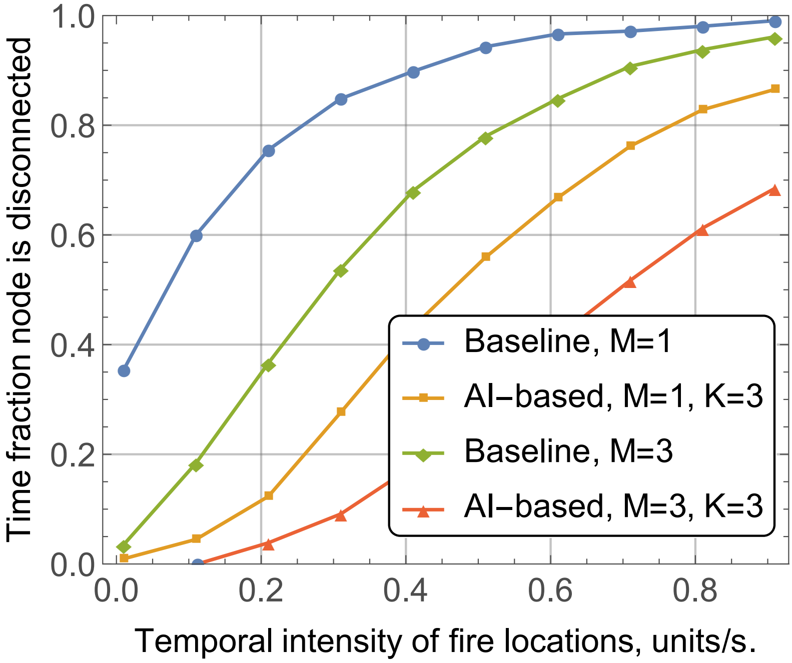


Figure 5.10 Fraction of disconnect time from a mesh

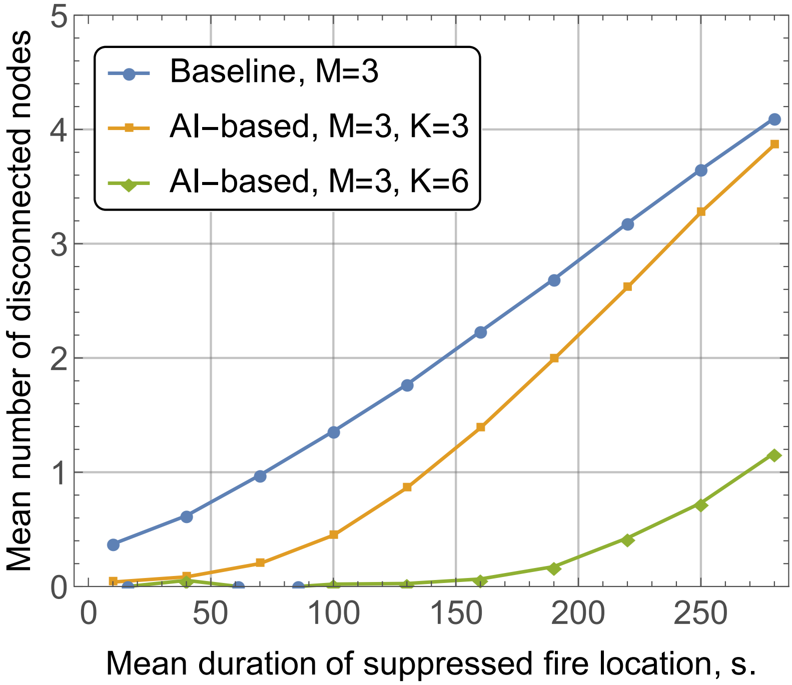


Figure 5.11 Mean number of disconnected nodes

Table 5.2 Default system parameters.

Parameter	Value
Number of static blockers	10
Radius of static blockers	5 m
Attenuation by static blockers	40 dB
Temporal intensity of fire locations	0.1 events/s
Mean duration of suppressed fire location	120 s
Radius of suppressed fire location	3 m
Attenuation by suppressed fire location	20 dB
Number of firefighting crew members	$\{10, 20, 50\}$
Number of autonomous robots	$\{1, 2 \dots 10\}$
Number of simultaneously supported links	$\{1, 2 \dots 10\}$

is presented in Fig. 5.10. As it follows from the plot, the path diversity allows a notable reduction of the negative effect of blockages. However, the system based only on multi-connectivity shows weaker performance if compared to the AI-aided case. The joint use of multi-connectivity and our proactive method strikes a strong balance in performance and shows the best results.

The mean of number disconnected nodes at a random time instance is presented in Fig. 5.11. This parameter provides insights about network ability to support application operation deployed over the mmW mesh. The more nodes disconnected from the mesh, the less efficiency of such an application. The results obtained for blockage intensity fixed on the level of 0.1 events/s, while the mean water vapor duration used as a variable. The plot indicates a significant improvement of the mesh reliability if the number of robots approaches the number of firefighters (there were ten firefighters in this simulation round). With a lower number of robots, reliability is low even in AI-aided cases.

We further characterize time-dependent performance – the intensity of node disconnections from a mesh in Fig. 5.12 as a function of the degree of multi-connectivity, M, the number of autonomous robot relays, K, and the mean water vapor duration in suppressed fire locations, $1/\theta$. As we learn, the response of the system is qualitatively similar for both the baseline and the AI-aided scenarios. An initial increase

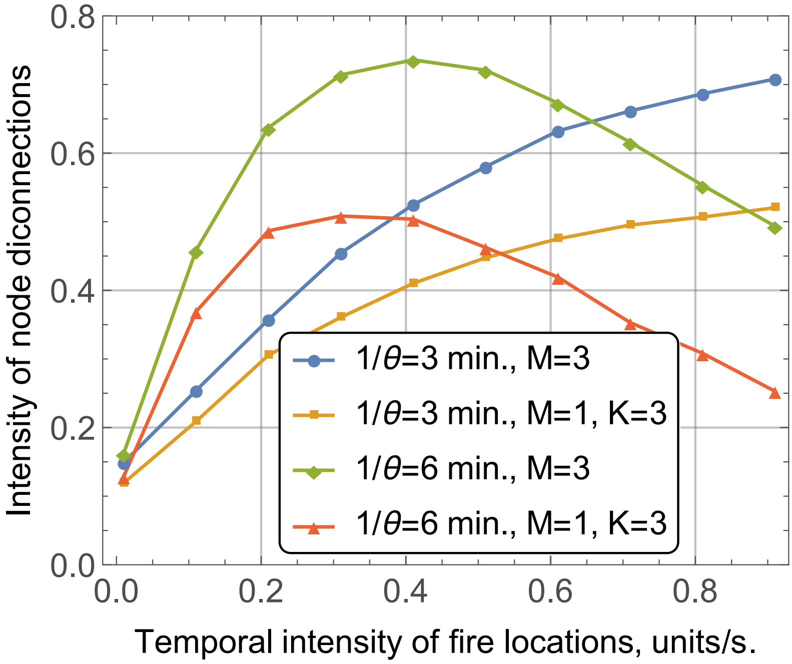


Figure 5.12 Intensity of node disconnects from a mesh

in the intensity of node disconnections is explained by the fact that the temporal intensity of node locations adds to the blockage dynamics as moving firefighters begin to experience link interruptions more frequently. However, when the intensity of dynamic blockers exceeds a certain value that generally depends on the type of the scenario and the selected system parameters, the intensity of node disconnections begins to decrease. The reason is that in this regime the number of static and dynamic blockers becomes so high that individual blockage periods merge into longer ones, thus forcing a node to spend more time in the disconnected state, see Fig. 5.10.

In Fig. 5.13 we evaluate the throughput via the data rate at the access point (for $M = 3$ and $K = 3$). The upper limit of the throughput is only achievable if dynamic blockages are disabled. Depending on the number of users generating traffic, the upper bound values results in approximately 5.4, 3.7, and 2.0 Gbps. Further, these values decrease with the growth of dynamic blockage intensity. This can be explained by higher disconnection (Fig. 5.10) leading to the packet losses. Notably, a lower fraction of the delivered data then leads to fewer fire locations detected, which, in turn, increases the fraction of disconnect time following an avalanche-like trend. Finally, this may results in full corruption of the mmW mesh since the system no

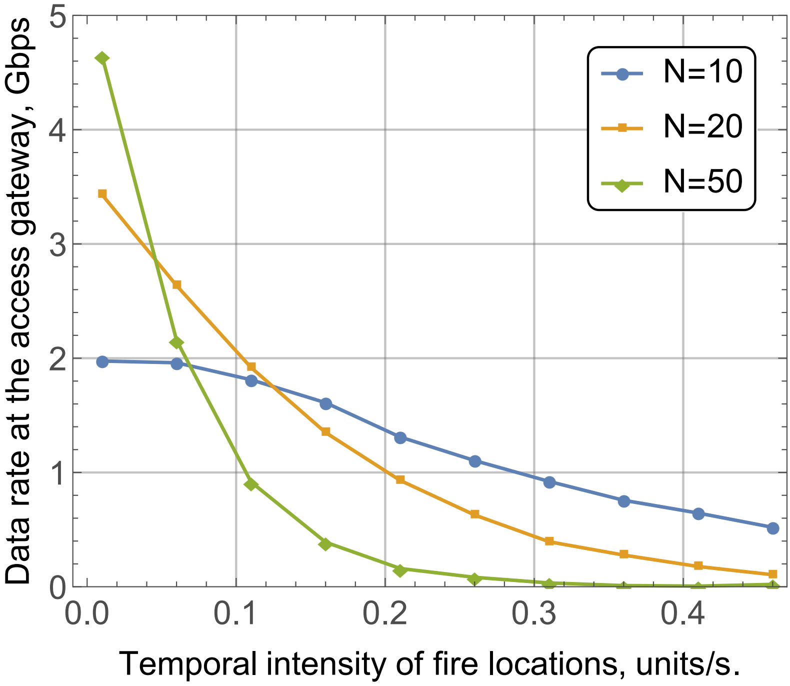


Figure 5.13 Data rate at access gateway

longer capable of maintaining its intended functionality.

6 CONCLUSIONS

6.1 Summary of the work carried out

This thesis covers a set of challenges associated with the provisioning of advanced IoT services in various scenarios, including ones with limited infrastructure support on the edge of the mobile network. The main results achieved in this thesis include the following:

- The evolution of IoT applications analyzed from the perspective of their deeper integration into everyday life and employment of advanced technologies for the design of the devices.
- Designed and evaluated a method for network identification of IoT devices.
- Developed and evaluated a set of UAV-based technologies for improving the connectivity of the wireless device. The considered scenarios include the deployment of constrained micro-sized sensor devices over large areas and the cooperation of IoT devices in locations with limited or unavailable mobile network services.
- Analyzed the potential of millimeter-wave mesh networks used for serving heavy traffic IoT applications that operate in environments with dynamic blockages.
- Developed a method for proactive blockage mitigation in millimeter-wave networks.

The comprehensive research work presented in the thesis provides a better understanding of the technical trade-offs that exist behind different solutions aimed at enhancing network experience at the wireless access level. We expect that the proposed solutions will become essential enablers for future IoT services delivering a robust communication layout for new communication services at the mobile network edge.

6.2 Future perspectives

The results of this thesis indicated the potential of multi-hop mesh networks based on a high-level evaluation. While the practical implementation of mmW meshes requires advanced beamforming and user tracking methods, this creates new research challenges. Moreover, considered scenarios imply merging of meshes and enabling connectivity on the fly, which calls for efficient methods of neighbor discovery and routing protocols.

In addition to conventional IP architecture, future research is expected to explore alternative networking approaches, such as ICN. The ICN concept relies on content identification instead of the device identification and extensively utilizes in-network caching of the data. Overall it allows for a reduction in the latency for IoT applications on the edge and eliminates the fundamental problems of meshes, referred to as a throughput bottleneck, in the middle of the network and user awareness about data and services available in the network.

Finally, there is a trend in shifting network intelligence to the edge, which requires new methods for computational resource orchestration and software container management. Existing solutions relying on Docker and Kubernetes are not able to provide acceptable scalability and latency, which calls for game-changing innovations from the network technologies side.

REFERENCES

[1] URL: `http://hitran.iao.ru/`.

[2] 3GPP. *NR; Multi-connectivity; Overall description (Release 15)*. 3GPP TS 37.340 V15.2.0. June 2018.

[3] J. H. Abawajy and M. M. Hassan. Federated internet of things and cloud computing pervasive patient health monitoring system. *IEEE Communications Magazine* 55.1 (2017), 48–53.

[4] I. F. Akyildiz, F. Brunetti and C. Blázquez. Nanonetworks: A new communication paradigm. *Computer Networks* 52.12 (2008), 2260–2279.

[5] I. F. Akyildiz and J. M. Jornet. Electromagnetic wireless nanosensor networks. *Nano Communication Networks* 1.1 (2010), 3–19.

[6] I. F. Akyildiz and J. M. Jornet. *Graphene-based plasmonic nano-antenna for terahertz band communication*. US Patent 9,643,841. May 2017.

[7] M. S. Anwer and C. Guy. A survey of VANET technologies. *Journal of Emerging Trends in Computing and Information Sciences* 5.9 (2014), 661–671.

[8] D. Aznakayeva, I. Yakovenko and E. Aznakayev. Numerical calculation of passive acoustic graphene nanosensor parameters. *Radar Methods and Systems Workshop (RMSW), IEEE*. IEEE. 2016, 95–98.

[9] D. Aznakayeva, I. Yakovenko and E. Aznakayev. Passive acoustic graphene nanosensor modeling. *Radar Methods and Systems Workshop (RMSW), IEEE*. IEEE. 2016, 91–94.

[10] C. A. Balanis. *Antenna theory: analysis and design*. John wiley & sons, 2016.

[11] R. Bez, E. Camerlenghi, A. Modelli and A. Visconti. Introduction to flash memory. *Proceedings of the IEEE* 91.4 (2003), 489–502.

[12] F. Bibi, C. Guillaume, N. Gontard and B. Sorli. A review: RFID technology having sensing aptitudes for food industry and their contribution to tracking and monitoring of food products. *Trends in Food Science & Technology* 62 (2017), 91–103.

[13] S. Blackman, C. Matlo, C. Bobrovitskiy, A. Waldoch, M. L. Fang, P. Jackson, A. Mihailidis, L. Nygård, A. Astell and A. Sixsmith. Ambient assisted living technologies for aging well: a scoping review. *Journal of Intelligent Systems* 25.1 (2016), 55–69.

[14] A. Bohák, L. Buttyán and L. Dóra. An authentication scheme for fast handover between WiFi access points. *Proc. of ACM Wireless Internet Conference (WICON).* Citeseer. 2007.

[15] P. Boronin, V. Petrov, D. Moltchanov, Y. Koucheryavy and J. M. Jornet. Capacity and throughput analysis of nanoscale machine communication through transparency windows in the terahertz band. *Nano Communication Networks* 5.3 (2014), 72–82.

[16] V. Brik, S. Banerjee, M. Gruteser and S. Oh. Wireless device identification with radiometric signatures. *Proceedings of the 14th ACM international conference on Mobile computing and networking*. ACM. 2008, 116–127.

[17] S. A. Busari, K. M. S. Huq, S. Mumtaz, L. Dai and J. Rodriguez. Millimeter-wave Massive MIMO Communication for Future Wireless Systems: A Survey. *IEEE Communications Surveys & Tutorials* 20.2 (2017), 836–869.

[18] G. Cecchini, A. Bazzi, B. M. Masini and A. Zanella. Performance comparison between IEEE 802.11 p and LTE-V2V in-coverage and out-of-coverage for cooperative awareness. *2017 IEEE Vehicular Networking Conference (VNC).* IEEE. 2017, 109–114.

[19] M. Cheffena. Industrial wireless communications over the millimeter wave spectrum: opportunities and challenges. *IEEE Communications Magazine* 54.9 (2016), 66–72.

[20] A. Cherchali, M. J. Gudelis Jr, W. G. Lester and R. J. McLaughlin. *Technique for automated MAC address cloning*. US Patent 9,124,474. Sept. 2015.

[21] P. Di Marco, R. Chirikov, P. Amin and F. Militano. Coverage analysis of Bluetooth low energy and IEEE 802.11 ah for office scenario. *Proc. of 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*. IEEE. 2015, 2283–2287.

[22] T. D. Dinh, R. Pirmagomedov, V. D. Pham, A. A. Ahmed, R. Kirichek, R. Glushakov and A. Vladyko. Unmanned aerial system–assisted wilderness search and rescue mission. *International Journal of Distributed Sensor Networks* 15.6 (2019), 1550147719850719. DOI: `10.1177/1550147719850719`.

[23] K. Doppler, E. Torkildson and J. Bouwen. On wireless networks for the era of mixed reality. *2017 European Conference on Networks and Communications (EuCNC)*. 2017, 1–5. DOI: `10.1109/EuCNC.2017.7980745`.

[24] F. Dressler and S. Fischer. Connecting in-body nano communication with body area networks: Challenges and opportunities of the Internet of Nano Things. *Nano Communication Networks* 6.2 (2015), 29–38.

[25] S. Frizzi, R. Kaabi, M. Bouchouicha, J.-M. Ginoux, E. Moreau and F. Fnaiech. Convolutional neural network for video fire and smoke detection. *Proc. of 42nd Annual Conference of the IEEE Industrial Electronics Society*. IEEE. 2016, 877–882.

[26] M. Gapeyenko, V. Petrov, D. Moltchanov, M. R. Akdeniz, S. Andreev, N. Himayat and Y. Koucheryavy. On the degree of multi-connectivity in 5G millimeter-wave cellular urban deployments. *IEEE Transactions on Vehicular Technology* 68.2 (2019), 1973–1978.

[27] I. Gepko. General requirements and security architecture for mobile phone anti-cloning measures. *IEEE EUROCON 2015 - International Conference on Computer as a Tool (EUROCON)*. Sept. 2015, 1–6. DOI: `10.1109/EUROCON.2015.7313666`.

[28] M. Gerasimenko, D. Moltchanov, M. Gapeyenko, S. Andreev and Y. Koucheryavy. Capacity of Multi-Connectivity mmWave Systems with Dynamic Blockage and Directional Antennas. *IEEE Transactions on Vehicular Technology* (2019).

[29] B. Gerkey, R. T. Vaughan and A. Howard. The player/stage project: Tools for multi-robot and distributed sensor systems. *Proceedings of the 11th international conference on advanced robotics*. Vol. 1. 2003, 317–323.

[30] O. Giwa and A. Benkrid. Fire detection in a still image using colour information. *arXiv preprint arXiv:1803.03828* (2018).

[31] A. Hegde. MAC spoofing detection and prevention. *Int. J. Adv. Res. Comput. Commun. Eng* 5.1 (2016), 229–232.

[32] S. A. Hoseini, M. Ding and M. Hassan. Massive MIMO performance comparison of beamforming and multiplexing in the Terahertz band. *2017 IEEE Globecom Workshops (GC Wkshps)*. IEEE. 2017, 1–6.

[33] R. Humpleman and P. Watson. Investigation of attenuation by rainfall at 60 GHz. *Proceedings of the Institution of Electrical Engineers*. Vol. 125. 2. IET. 1978, 85–91.

[34] I. IAMSAR. International aeronautical and maritime search and rescue manual. *Mission coordination* 2 (2007).

[35] T. Ismail, E. Leitgeb and T. Plank. Free space optic and mmWave communications: technologies, challenges and applications. *IEICE Transactions on Communications* 99.6 (2016), 1243–1254.

[36] M. Jakobsson and K. Johansson. Unspoofable Device Identity Using NAND Flash Memory. *SecurityWeek* (2010). URL: `https://www.securityweek.com/unspoofable-device-identity-using-nand-flash-memory`.

[37] S. Jeschke, C. Brecher, T. Meisen, D. Özdemir and T. Eschert. Industrial internet of things and cyber manufacturing systems. *Industrial Internet of Things*. Springer, 2017, 3–19.

[38] S. Jia, L. Xia, Z. Wang, J. Lin, G. Zhang and Y. Ji. Extracting robust keys from nand flash physical unclonable functions. *International Information Security Conference*. Springer. 2015, 437–454.

[39] J. M. Jornet and I. F. Akyildiz. Channel modeling and capacity analysis for electromagnetic wireless nanonetworks in the terahertz band. *IEEE Transactions on Wireless Communications* 10.10 (2011), 3211–3221.

[40] J. M. Jornet and I. F. Akyildiz. Information capacity of pulse-based wireless nanosensor networks. *Sensor, Mesh and Ad Hoc Communications and Networks (SECON), 2011 8th Annual IEEE Communications Society Conference on*. IEEE. 2011, 80–88.

[41] J. M. Jornet and I. F. Akyildiz. Low-weight channel coding for interference mitigation in electromagnetic nanonetworks in the terahertz band. *Communications (ICC), 2011 IEEE International Conference on*. IEEE. 2011, 1–6.

[42] J. M. Jornet and I. F. Akyildiz. Joint energy harvesting and communication analysis for perpetual wireless nanosensor networks in the terahertz band. *IEEE Transactions on Nanotechnology* 11.3 (2012), 570.

[43] R. Kirichek, R. Pirmagomedov, R. Glushakov and A. Koucheryavy. Live substance in cyberspace – Biodriver system. *Proc. of 18th International Conference on Advanced Communication Technology (ICACT)*. IEEE. 2016, 274–278.

[44] T. Kohno, A. Broido and K. C. Claffy. Remote physical device fingerprinting. *IEEE Transactions on Dependable and Secure Computing* 2.2 (2005), 93–108.

[45] J. Kokkoniemi, J. Lehtomäki, K. Umebayashi and M. Juntti. Frequency and time domain channel models for nanonetworks in terahertz band. *IEEE Transactions on Antennas and Propagation* 63.2 (2015), 678–691.

[46] M. Leo, G. Medioni, M. Trivedi, T. Kanade and G. M. Farinella. Computer vision for assistive technologies. *Computer Vision and Image Understanding* 154 (2017), 1–15.

[47] L. Militano, M. Condoluci, G. Araniti, A. Molinaro and A. Iera. When D2D communication improves group oriented services in beyond 4G networks. *Wireless Networks* 21.4 (2015), 1363–1377.

[48] H. Mora-Mora, V. Gilart-Iglesias, D. Gil and A. Sirvent-Llamas. A computational architecture based on RFID sensors for traceability in smart cities. *Sensors* 15.6 (2015), 13591–13626.

[49] K. Muhammad, J. Ahmad and S. W. Baik. Early fire detection using convolutional neural networks during surveillance for effective disaster management. *Neurocomputing* 288 (2018), 30–42.

[50] P. Nain, D. Towsley, B. Liu and Z. Liu. Properties of Random Direction Models. *Proc. of 24th Annual Joint Conference of the IEEE Computer and Communications Societies*. Vol. 3. IEEE. 2005, 1897–1907.

[51] Y. Niu, Y. Li, D. Jin, L. Su and A. V. Vasilakos. A survey of millimeter wave communications (mmWave) for 5G: opportunities and challenges. *Wireless Networks* 21.8 (2015), 2657–2676.

[52] H. Perros. Computer simulation techniques. *The definitive introduction. North Carolina State University* (2009).

[53] V. Petrov, D. Moltchanov, P. Kustarev, J. M. Jornet and Y. Koucheryavy. On the use of integral geometry for interference modeling and analysis in wireless networks. *IEEE Communications Letters* 20.12 (2016), 2530–2533.

[54] R. Pirmagomedov, D. Moltchanov, A. Ometov, K. Muhammad, S. Andreev and Y. Koucheryavy. Facilitating mmWave Mesh Reliability in PPDR Scenarios Utilizing Artificial Intelligence. *IEEE Access* (2019). ISSN: 2169-3536. DOI: 10.1109/ACCESS.2019.2958426.

[55] R. Pirmagomedov, R. Kirichek, M. Blinnikov and A. Koucheryavy. UAV-based gateways for wireless nanosensor networks deployed over large areas. *Computer Communications* 146 (2019), 55–62. DOI: 10.1016/j.comcom.2019.07.026.

[56] R. Pirmagomedov and Y. Koucheryavy. IoT Technologies for Augmented Human: a Survey. *Internet of Things* (2019), 100120. DOI: 10.1016/j.iot.2019.100120.

[57] R. Pirmagomedov, D. Moltchanov, V. Ustinov, M. N. Saqib and S. Andreev. Performance of mmWave-Based Mesh Networks in Indoor Environments with Dynamic Blockage. *International Conference on Wired/Wireless Internet Communication*. Ed. by D. F. M., N. E., B. R. and K. A. 2019, 129–140. DOI: 10.1007/978-3-030-30523-9_11.

[58] J. Qiao, X. S. Shen, J. W. Mark, Q. Shen, Y. He and L. Lei. Enabling device-to-device communications in millimeter-wave 5G cellular networks. *IEEE Communications Magazine* 53.1 (2015), 209–215.

[59] A. Ramamurthy and P. Jain. The Internet of Things in the Power Sector Opportunities in Asia and the Pacific. (2017).

[60] S. P. Rao, S. Holtmanns, I. Oliver and T. Aura. Unblocking Stolen Mobile Devices Using SS7-MAP Vulnerabilities: Exploiting the Relationship between IMEI and IMSI for EIR Access. *2015 IEEE Trustcom/BigDataSE/ISPA*. Vol. 1. 2015, 1171–1176. DOI: 10.1109/Trustcom.2015.500.

[61] N. D. Rasmussen, B. S. Morse, M. A. Goodrich and D. Eggett. Fused visible and infrared video for use in wilderness search and rescue. *2009 Workshop on Applications of Computer Vision (WACV)*. IEEE. 2009, 1–8.

[62] K. Remley, C. A. Grosvenor, R. T. Johnk, D. R. Novotny, P. D. Hale, M. McKinley, A. Karygiannis and E. Antonakakis. Electromagnetic signatures of WLAN cards and network security. *Signal Processing and Information Technology, 2005. Proceedings of the Fifth IEEE International Symposium on*. IEEE. 2005, 484–488.

[63] L. S. Rothman, C. Rinsland, A. Goldman, S. Massie, D. Edwards, J. Flaud, A. Perrin, C. Camy-Peyret, V. Dana, J.-Y. Mandin et al. The HITRAN molecular spectroscopic database and HAWKS (HITRAN Atmospheric Workstation): 1996 edition. *Journal of Quantitative Spectroscopy and Radiative Transfer* 60.5 (1998), 665–710.

[64] M. Rusci, D. Rossi, M. Lecca, M. Gottardi, E. Farella and L. Benini. An event-driven ultra-low-power smart visual sensor. *IEEE Sensors Journal* 16.13 (2016), 5344–5353.

[65] K. Sekaran, M. S. Khan, R. Patan, A. H. Gandomi, P. V. Krishna and S. Kallam. Improving the Response Time of M-Learning and Cloud Computing Environments Using a Dominant Firefly Approach. *IEEE Access* 7 (2019), 30203–30212.

[66] E. Suh and S. Devadas. Physical Unclonable Functions for Device Authentication and Secret Key Generation. *Proc 44th ACM/IEEE Design Automation Conference*. July 2007, 9–14. ISBN: 978-1-59593-627-1. DOI: 10.1109/DAC.2007.375043.

[67] A. Tal. Two flash technologies compared: NOR vs NAND. *White Paper of M-Systems* (2002).

[68] K. I. Talbot, P. R. Duley and M. H. Hyatt. Specific emitter identification and verification. *Technology Review* 113 (2003).

[69] M. Thibaud, H. Chi, W. Zhou and S. Piramuthu. Internet of Things (IoT) in high-risk Environment, Health and Safety (EHS) industries: A comprehensive review. *Decision Support Systems* 108 (2018), 79–95.

[70] A. Thornburg, T. Bai and R. W. Heath Jr. Performance Analysis of Outdoor mmWave Ad Hoc Networks. *IEEE Trans. Signal Processing* 64.15 (2016), 4065–4079.

[71] Toshiba. *NAND vs. NOR Flash Memory Technology Overview*. Tech. rep. Technical Report, 2006.

[72] K. Tsujimura, K. Umebayashi, J. Kokkoniemi and J. Lethomäki. A study on channel model for THz band. *Antennas and Propagation (ISAP), 2016 International Symposium on*. IEEE. 2016, 872–873.

[73] S. Vaidya and K. J. Christensen. A single system image server cluster using duplicated MAC and IP addresses. *Proceedings LCN 2001. 26th Annual IEEE Conference on Local Computer Networks*. 2001, 206–214. DOI: `10.1109/LCN.2001.990789`.

[74] R. Vaughan. Massively multi-robot simulation in stage. *Swarm intelligence* 2.2-4 (2008), 189–208.

[75] D. M. S. Velandia, N. Kaur, W. G. Whittow, P. P. Conway and A. A. West. Towards industrial internet of things: Crankshaft monitoring, traceability and tracking using RFID. *Robotics and Computer-Integrated Manufacturing* 41 (2016), 66–77.

[76] K. Venugopal and R. W. Heath. Millimeter wave networked wearables in dense indoor environments. *IEEE Access* 4 (2016), 1205–1221.

[77] K. Venugopal, M. C. Valenti and R. W. Heath. Analysis of millimeter wave networked wearables in crowded environments. *2015 49th Asilomar Conference on Signals, Systems and Computers*. IEEE. 2015, 872–876.

[78] R. Vijayakumar, K. Selvakumar, K. Kulothungan and A. Kannan. Prevention of multiple spoofing attacks with dynamic MAC address allocation for wireless networks. *2014 International Conference on Communication and Signal Processing*. Apr. 2014, 1635–1639. DOI: `10.1109/ICCSP.2014.6950125`.

[79] S. S. Vladimirov, R. Pirmagomedov, R. Kirichek and A. Koucheryavy. Unique Degradation of Flash Memory as an Identifier of ICT Device. *IEEE Access* 7 (2019), 107626–107634.

[80] S. S. Vladimirov, R. Pirmagomedov, R. Kirichek and A. Koucheryavy. Unique Degradation of Flash Memory as an Identifier of ICT Device. *IEEE Access* 7 (2019), 107626–107634. DOI: `10.1109/ACCESS.2019.2932804`.

[81] Y. Wang, W. Yu, S. Wu, G. Malysa, G. E. Suh and E. C. Kan. Flash Memory for Ubiquitous Hardware Security Functions: True Random Number Generation and Device Fingerprints. *2012 IEEE Symposium on Security and Privacy*. May 2012, 33–47. DOI: `10.1109/SP.2012.12`.

[82] E. W. Weisstein. Circle line picking. *From MathWorld–A Wolfram Web Resource* (2004).

[83] C. Xia and P. Maes. The design of artifacts for augmenting intellect. *Proceedings of the 4th Augmented Human International Conference*. ACM. 2013, 154–161.

[84] Z. Xu, X. Dong and J. Bornemann. Design of a reconfigurable MIMO system for THz communications based on graphene antennas. *IEEE Transactions on Terahertz Science and Technology* 4.5 (2014), 609–617.

[85] J. Zhang, G. Tian, A. Marindra, A. Sunny and A. Zhao. A review of passive RFID tag antenna-based sensors and systems for structural health monitoring applications. *Sensors* 17.2 (2017), 265.

[86] L. Zhang, I. K. Dabipi and W. L. Brown Jr. Internet of Things Applications for Agriculture. *Internet of Things A to Z: Technologies and Applications* (2018), 507–528.

PUBLICATIONS

PUBLICATION

I

IoT Technologies for Augmented Human: a Survey

R. Pirmagomedov and Y. Koucheryavy

Internet of Things (2019), 100120
DOI: 10.1016/j.iot.2019.100120

IoT Technologies for Augmented Human: a Survey

Rustam Pirmagomedov
Tampere University
Tampere, Finland
Email: rustam.pirmagomedov@tuni.fi

Yevgeni Koucheryavy
Tampere University
Tampere, Finland
Email: evgeny.kucheryavy@tuni.fi

***Abstract*—Internet of Things (IoT) technology has delivered new enablers for improving human abilities. These enablers promise an enhanced quality of life and professional efficiency; however, the synthesis of IoT and human augmentation technologies has also extended IoT-related challenges far beyond the current scope. These potential challenges associated with IoT-empowered Augmented Human (AH) have so far not been well-investigated. Thus, this article attempts to introduce readers to AH concept as well as summarize notable research challenges raised by such systems, in order to facilitate reader's further interest in this topic. The article considers emerging IoT applications for human augmentation, devices and design principles, connectivity demands, and security aspects.**

I. Introduction

Present efforts in human augmentation (sometimes referred to as "Human 2.0") focus on the creation of cognitive and physical improvements as an integral part of the human body [1]. These improvements are enabled by specially designed devices, such as leg or hand prosthesis, implants, artificial vision connected to the neural system of an organism, augmented reality glasses, hearing aids, and insulin pumps. Artificially recreated or extended abilities may improve quality of life and even give some competitive advantages for users.

Currently, the progress in human augmentation is driven by the interconnected Internet of Things (IoT) devices. The performance of these devices relies heavily on communication technologies. Commonly, such devices are located in close proximity to the human body. More specific applications may utilize bio-integrated devices, for example, neurally-controlled artificial limbs. All the devices used by an individual form an integrated ecosystem and should work coherently, which enabled by appropriate communication technologies. Depending on the type of the device the utilized communication technology may vary from traditional wireless protocols such as Bluetooth or Wi-Fi to highly specific technologies, such as electromagnetic or molecular nanonetworks. Therefore, a network of assisting devices can be considered as a highly heterogeneous Body Area Network (BAN). In addition to local communication, applications of Augmented Human (AH) require an internet connection (e.g., to be aware of context, offload of difficult computational tasks, upgrade software). As a whole, the concept of the Augmented Human creates a new segment of communication challenges, since the reliable performance of communication technologies in such systems is the essential enabler for the users' well-being.

Presently, the research on AH is spread across many different communities. From the perspective of communication technologies, devices for human augmentation have a lot in common with wearable electronics. However, being a branch of the IoT concept, AH devices perhaps provide the most critical class of services, because humans do not exist independently but rather as a part of human-centric AH systems in which one is trained [1]. A failure in an AH application would cause chaos in this system and make the human vulnerable. Inherently, a communication failure will reduce a user's physical or cognitive abilities. Thus, the AH applications require comprehensive analysis from the technological perspective to define their place in emerging network services.

This paper provides an overview of the IoT technologies for AH and defines relevant research challenges in this innovative area.

The article is organized in the following way: Section II provides an overview of AH applications. Section III considers the aspects and trends in device design. In Section IV we consider connectivity demands of the AH applications. Section V we discuss the security concerns for AH applications, and conclude the article in Section VI.

II. Applications of AH Technologies

Attempts to recover or improve human abilities began in ancient times. The majority of these attempts aimed at replacing a lost body part with an artificial one, for example, a leg or hand prosthesis. Some enthusiastic inventors aimed to go beyond the natural capabilities of the human organism by developing "upgrades", such as wings for flying. These two vectors of augmentation development are still relevant and form an augmentation continuum as shown in the Fig.1. Initially, the majority of efforts towards human augmentation were focused on the improvement of physical abilities, while in the 20th century, due to progress in microelectronics, augmentation has been extended by advanced sensing and cognitive improvements. Small-sized electronic devices are capable of assisting in performing specific tasks, e.g., a hearing aid assists people with auditor disorders or the use of AR glasses capable of providing both navigation support and object recognition.

[1] https://www.gartner.com/it-glossary/human-augmentation/

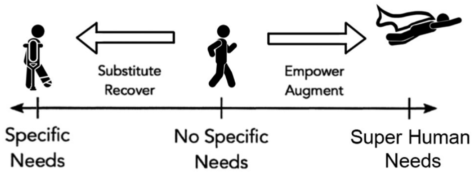


Fig. 1. Assistive augmentation continuum according to [2]

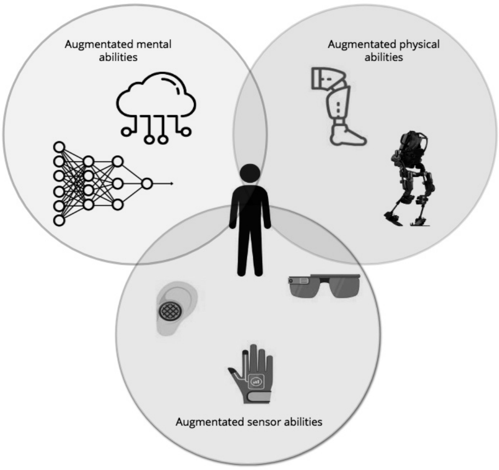

Augmentated mental abilities
Augmentated physical abilities
Augmentated sensor abilities

Fig. 2. Areas of a human augmentation

A. Objectives of AH applications

In a general case, AH systems assisting in daily routines [3] using electronic devices, which are connected to a single BAN via different communication technologies. The network of interconnected wearables serves as a technological layout for high-level applications of AH, which enable physical augmentation, advanced sensing, and mental assistance Fig. 2.

Physical augmentation aims at enhancement of an individual's ability to move and manipulate objects. Examples of tools used for physical augmentation include exoskeleton, artificial arms and legs, or even a jet pack. Failures in physical augmentation can be hazardous for human safety and health; therefore physical augmentation tools should be capable of providing basic functionality even when network service or other resources are unavailable.

Sensory abilities allow a person to be aware of the environment and the context surrounding them. Sensory abilities may include vision, touch, hearing, smell, and taste. Augmentation may facilitate these senses by amplifying them or, in the case of having lost a sense, augmentation allows a transformation of the characteristics of one sensory modality into stimuli of another sensory modality [4], e.g., visualizing speech or smells.

Mental (or cognitive) augmentation provides data processing assistance and facilitates decision making. An illustrative example of a cognitive augmentation is a personal planning application, where users can save time and resources when planning daily routines. The application may plan optimal logistics during the day, select and book lunch at the highest quality and yet affordably priced restaurant within the defined location, find parking spaces and car charging plugs, integrate recommended physical activity into the day's timeline, and automatically revise plans in accordance with changing conditions (automatic negotiations with involved parties and reconfiguration of schedule). All these functions can be performed in a background mode, increasing the efficiency of a working day and saving time for creative activities or leisure. Currently, cognitive augmentation is the most familiar branch of human augmentation because of its widespread use in mobile applications. As one may observe, technologically such applications rely on machine learning [5] and entirely hinge on information about the environment, while also being significantly dependent on an internet connection.

B. Classification of AH applications

Taxonomy of AH applications include three major classes: (i) supporting independent living (e.g., for the aging population or people with impairments); (ii) facilitating the professional performance; (iii) self-efficiency and entertainment.

The applications which support independent living allow users to satisfy their basic daily needs without the assistance of other people. In addition, such applications monitor users' health conditions in real-time and increase their safety (e.g., by protecting aging people from occasional fall). As a result, nursing costs can be considerably reduced while also improving the quality of life for both aging and disabled people.

The AH applications for improving professional performance focus on augmenting the abilities relevant to the professional areas of an individual. For instance, an exoskeleton for a worker allows for moving heavy weights without harmful consequences for the spine. Another illustrative example comes from emergency response, where AH may enhance the performance of rescue team members by providing augmented sensing (e.g., sensing of hazard gases, utilizing thermal vision), empowered physical abilities (e.g., exoskeleton), and efficient decision making (e.g., AI-assisted operation).

The entertainment class aims to provide unusual user experiences (e.g., flying with a jet pack) or an immersive experience of extreme situations (e.g., virtual reality gaming) without physical risks to the user.

It is worth noting that AH systems may encompass the entire context where a person exists [6], which include interaction with proximate entities such as buildings, city infrastructure,

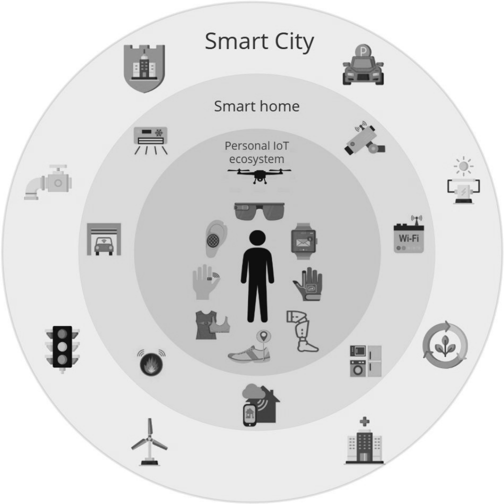


Fig. 3. Integrated smart environment

and other individuals. Communicating with each other, these form an integrated smart environment (Fig.3).

C. Challenges

Ethical and social aspects. Innovative technologies for human augmentation will undoubtedly bring new challenges in ethical and social fields. Augmented features may become trendy especially among a younger generation [7], or the border between artificial and natural abilities may become blurred. It is not evident which specific issues will become a part of the agenda, nevertheless, ethical and social aspects require primary consideration.

Human-machine interaction. User experience issues are of the utmost importance when designing IoT systems for human augmentation. The IoT-empowered AH systems should provide functional but straightforward user interfaces to avoid certain user groups (e.g., older people) from feeling uncomfortable when using such systems (due to the system complexity).

III. Devices and Design Principles

Technological challenges related to AH devices are primarily shaped by design principles, which rely on users' demands and expectations. Regarding the wearable IoT devices (which AH systems are), the users' expectations are primary: small size and weight of devices, enhanced reliability, and long battery lifetime.

A. Flexible Hybrid Electronics

Flexible Hybrid Electronics (FHE) integrates devices from thinned flexible materials with electric circuits in formats that can be thin, light-weight, flexible, bendable, conformal, potentially stretchable and disposable [8]. FHE offer notable advantages over the conventional electronic systems that are made of bulky and rigid materials [9]. Recent advancements in the advanced materials and soft mechanics have enabled a successful integration of rigid, miniaturized chips with flexible/stretchable circuit interconnects. Such FHE in AH applications enhances signal processing, memory, and wireless power transfer in wearable systems [10], [11]. For example, in real-time monitoring of health parameters, FHE enables bio-friendly devices on biological tissues, such as artificial human skin, or internal organs with time-dynamic motions [12]. In general, implementation of FHE in AH enabling improved wearability and performance for the devices, and as a result, facilitating their use among individuals.

B. Reduced size of the devices

Due to the progress in nanotechnologies, IoT wearables can be deployed at the nano level (named as Nanonetworks) [13]. Such nanodevices employ unique properties of graphene, which allows a significant size decrease for electronic elements, including antennas, processors, receivers and transmitters [14]–[16], as well as sensors and actuators [17], [18]. The graphene-based nanoantennas enable communication in the THz frequency band [15]. However, the distance of communication in the THz frequency band is substantially limited by the high signal power losses during propagation [19], [20]. The distance of communication will not exceed 2 meters, even in an air environment with minimal humidity; if the communication is performed in an environment with a high concentration of liquids, such as a human body, the distance of transmission will decrease to several millimeters [21] establishing new challenges related to enabling communication within such networks. These graphene-based devices (antennas and transceivers for THz communication) are small enough to be integrated into biological systems (on the border between the organism and the environment) and can be easily integrated into modern communication devices (e.g., smartphones) as they are based on existing electronic technologies.

C. Improving energy effeciency

A power unit used in wearables typically the most significant contributor to both the size and weight of the devices [22]. As a consequence, developers must balance between size and the capability for autonomous operation when designing wearables. A majority of devices are designed with the priority given to size and weight, and thus have minimal operation time between recharging [23]. However, users' expectations continue to move toward fully autonomous devices without recharges or other maintenance operations. To address these demands, recent research efforts have targeted enhanced battery lifetime through improving the energy efficiency of the devices. Significant energy costs in wearables come from network functions, data acquiring, and processing [24]–[26].

The networking overheads of wearable devices was investigated in [27], [28]. More specifically, these works considered digital traffic generated by the wearable network in real-time mode. Results of the study demonstrated that network resource utilization in wearable systems is extremely low due to signaling overheads. However, the efficiency can be improved if an advanced data management algorithm is utilized on the BAN gateway. One such algorithm was proposed and evaluated in [29]. The reported results demonstrated improved networking efficiency by approximately 80 percent via the reduction in network signaling overheads, while the performance of applications decreased negligibly. Despite the notable improvements in networking, the energy efficiency of the considered systems is far from optimal and has massive potential for further improvement.

From the perspective of data processing, a drastic improvement is the promise of Approximate Computing (AC) [30]. Approximate Computing is inspired by the Pareto Principle according to which, roughly 80 percent of the effects come from 20 percent of the causes. Regarding the wearable networks, this principle can be formulated in the following way: capturing just 20 percent of the data may enable 80 percent of the application's performance. It should be noted that the actual percentage can be different; however, the general principle remains the same – a minority of efforts provides the majority of results.

Wearable applications work with noisy data; thus they are natively resilient to error [31], moreover, most of the applications do not require extremely precise results, thus the paradigm of an acceptable margin of error as introduced by AC promises significant energy-efficiency gains for AH systems.

D. Reliability

Reliability issues need to be addressed long before a device could be considered for any mission-critical application. However, the reliability and validity of existing wearable devices is concerning. The majority of available devices are not verified in terms of accuracy and reliability [32]. Recent tests among wearables showed significant variations of accuracy with error margins of up to 25 percent [33].

In addition to device reliability, but by no means less important, is the enabling quality of server platforms. Possible adverse effects from a cloud server failure are widely discussed in the literature [22], [34], [35], and can be considerably mitigated via placement optimization [36].

E. Challenges

Developing networks of nanodevices. Recent developments in nanotechnologies have enabled tiny-sized devices with both sensor and actuator functionality. However, due to multiple limitations, these devices are not capable of supporting standard communication protocols, including medium access control, routing, and security. Although networking among nanodevices is widely discussed in the literature, commercially available solutions have yet to be delivered, which keeps the door open for transferring theoretical findings to the real world.

Power supply. Emerging AH systems should fully utilize the benefits of efficient wireless power harvesting and energy transmission [37], as well as low energy technologies, for reducing a user's routine in its relation to the charging of devices.

Requirements for the devices and testing specifications. Despite, the notable progress in provisioning reliable AH operation, there is a lack of systematic perspective on the reliability of mission-critical IoT systems. This gap is expected to be fulfilled by the efforts of international standardization bodies (e.g., SG11 ITU-T) which perform extensive work towards the standardization of unified testing procedures for such systems.

Balancing the trade-offs between energy efficiency and accuracy. Implementation of AC promises a reduction of energy consumption by computing and sensing blocks of AH systems. However, the balance between energy efficiency and application performance should be clearly defined.

IV. Connectivity Demands of AH

The connectivity demands of AH include intra-BAN and inter-BAN considerations and cover physical interfaces, networking architecture, and AH integration in emerging network infrastructure (5G/5G+).

A. Multi-tier networking architecture

To enable the sustainable operation of AH devices, and context-awareness, AH systems must support multi-connectivity when operating in a multi-tier network environment (Fig.4).

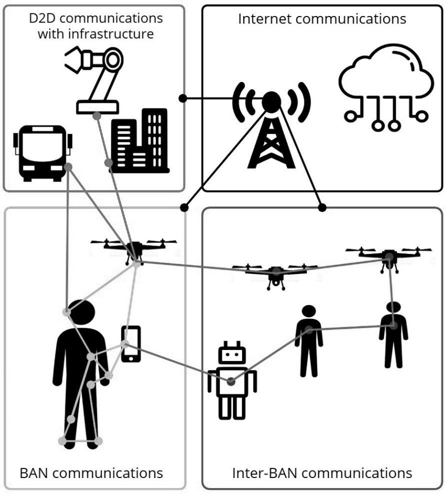


Fig. 4. Integrated smart environments

Intra-BAN communications integrate all personal devices of an individual into one network. Such a network can operate in a distributed way or can be orchestrated by the head node (e.g., smartphone or body gateway). The orchestrated BAN is less reliable, because the fault of orchestrating devices causes a disruption of the whole BAN, while in a case of a distributed approach the network is resilient to the faults of individual devices [38]. On the contrary, an orchestrated BAN demonstrates better quality of service (QoS) and energy efficiency [39]–[41].

Inter-BAN communication covers the interaction between devices of two or more individuals. Such interaction often relies on device-to-device communication (D2D) and is required for enabling synchronization among AH systems when users collaborate. This type of communication is commonly characterized by a higher temporal and spatial dynamic (e.g., link blockages and outages). To improve the stability of sessions, the communication links can be established via assisting robot relays, such as drones. Drones can be considered as a part of a personal IoT ecosystem where they contribute to sensor augmentation, providing additional information about the environment. Simultaneously they may serve as relays for reliable D2D communication (e.g., connection with a user around the corner). In total, communication with infrastructure in direct mode allows an AH to be context-aware without a load on the mobile infrastructure. An internet connection via the mobile network infrastructure can be used for accessing cloud servers and other devices, which are not reachable via D2D communication (e.g. offloading of computation to the edge) [42].

B. *Wired and wireless*

Wireless interfaces allow the creation of flexible connectivity within BAN. The tree of IEEE 802.15 standards specified wireless technologies adjusted for BAN use cases. Commonly used wireless technologies for intra-BAN communication include Bluetooth (IEEE 802.15.1), ZigBee (IEEE 802.15.4) and more recent WiMedia (IEEE 802.15.3) for ultra-wideband links [22], [37]. Inter-BAN connections are enabled by Wi-Fi (IEEE 802.11) and mobile networks (e.g., LTE, 5G NR).

Wired interfaces enable improved reliability and stable quality of connections, which can be fundamentally important for critical elements of AH. Moreover, wired devices (almost unsusceptible to radio interference) reduce the problem of the radio-noisy environment when plenty of wireless devices are working in close proximity (ultra-dense scenario). In addition, a wired connection can be used for an energy supply which is a notable advantage of such systems. However, the low flexibility of wired networks significantly limits their utilization in AH systems. Wired connections are currently selected exclusively for intra-BAN communication. The most suitable niche for wired communications is related to cases where connected devices are not expected to be moved considerably in relation to each other. For example, elements of an exoskeleton, elements of smart textile, and sensors embedded in the skin and connected via smart tattoo.

Recent advances in inductive links and intrabody links may establish a new branch of communication technologies for AH systems, based on using human body tissues as a transmission medium (e.g., molecular communication) [13], [43].

C. *Increasing throughput*

Initially, wireless technologies for machine type communications (interconnected devices) were developing with a focus to low rate traffic (e.g., telemetry), and a limited density of devices in a network. Presently, due to the reduced size of wearables, the density of connected devices can be significant. In addition, their services have spread far beyond simple telemetry, and they now use media extensively (e.g., AR/VR video services). As a result, IoT devices are generating a notable portion of data in the network, which can be expected to continue to increase in the future.

Supporting a high data rate among wireless wearable devices, especially in dense deployment (e.g., crowded streets of the city, stadiums) is a challenging task. The primary concern is interference when many devices operate simultaneously. As an alternative to an extensively employed microwave spectrum, it is proposed to use millimeter-wave (mmWave) links [44]. Due to the higher spectrum and less interference (because of greater signal loss at these frequencies), mmWave links are considered as a solution for the mitigation of interference and throughput concerns in emerging wearable networks [45].

D. *Augmented Human in 5G/5G+ landscape*

The connectivity challenges of AH in 5G/5G+ networks are driven by the spontaneous forming, maintaining and termination of heterogeneous networks of AH devices, and traffic flow balancing. Mission-critical communication has already been deeply investigated and discussed in the literature. However, network demands in the considered scenarios all occur in predefined locations (e.g., manufacturing, transportation hubs, medical facilities) [46]–[48], while network demands of AH applications are characterized by a high degree of temporal and spatial variations [49]. Therefore, conventional “static” network planning methods are inefficient for AH and require development of adaptive methods.

In comparison with legacy mobile networks, 5G systems bring a considerable shift in the quality of services offering Ultra-Reliable Low-Latency Communication (URLLC) for delay-sensitive applications, which open new horizons for AH applications. More specifically, the dynamic demands of AH are expected to be addressed in 5G by utilizing mobile access points (e.g., cell on wheels, aerial access point) and traffic offloading on D2D mesh networks. Additionally, connectivity of AH can be considerably enhanced by utilizing multiband access (e.g., using sub-6 GHz and millimeter-wave bands of 5G NR simultaneously).

Nevertheless, the mission-critical services natively supported by 5G systems require standardization efforts to meet AH demands. These efforts should result in a prioritized network service of AH applications and support interoperability between AH systems in 5G and beyond.

E. *Challenges*

D2D mesh networking. Secure inter-BAN multi-hop D2D communications are required for supporting merging and consequent splitting of AH systems employed by different users during their collaborative activities. Merging of AH systems means the incorporation of corresponding BANs. Such connectivity on the fly, requires robust devices identification method, neighbor's discovery, routing, automatic choice and assignment of devices acting as a heterogeneous gateway to connect devices with different radio access technologies. For the last, but not least, it is essential to incentivize users to share their resources and participate in mesh networks; otherwise, the performance of the meshes will be very limited.

Health Concerns. Wide use of wireless wearables raises concerns related to the effects caused by the high-frequency electromagnetic waves on people's health. The sensitivity of human tissues and skin to electromagnetic radiation, as well as long term effects caused by wireless devices, needs to be analyzed carefully.

Enabling directional wireless communications in BAN Directional wireless communication is extensively discussed in the literature as a feature of emerging air interfaces operating using high frequencies (e.g., millimeter-wave or THz communication). Utilization of directional antennas in wearable networks significantly increases the complexity of wireless interfaces, but promising lower interference among devices and gigabits-per-second rates (if mmWave links used) [45]. To enable directional links in BAN, research challenges related to beamforming techniques must be addressed [50].

Adaptive network management mechanisms. A novel signaling architecture is required for capturing and predicting AH demands, in order to enable real-time network adaptation to varying demands of AH application and the varying available network resource. The promising solutions for addressing this challenge may come from the synthesis of machine learning approach and SDN/NFV technologies.

V. Security Considerations

Applications of AH bring security concerns to the top, as security breaches in such enablers can have dramatic results to both the infrastructure and the individuals who rely on them. International standardization bodies are considering security challenges architecturally [51]–[53]. Following this, Fig. 5 summarized in a layered manner, common security threats relevant for AH applications.

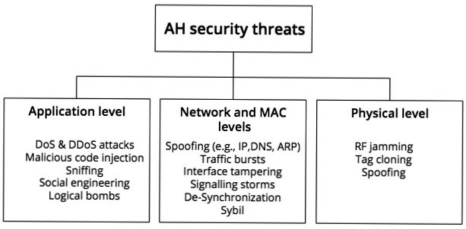


Fig. 5. Common security threats in layered approach

A. Physical level

Attacks on the physical level may disrupt the normal operation of connected devices even if the high level (MAC, network and application) is well designed. For example radio frequency (RF) jamming may imply interruption of wireless communication using high power radio signals of the same frequency as used by AH devices. RF jamming may entirely block communication or interfere with it. The later may exhaust batteries of wearables due to additional energy costs required for numerous retransmissions, using higher transmit power, idle listening, etc.

A wireless medium is essentially a broadcast one, which makes such systems vulnerable to eavesdropping (e.g., attackers may eavesdrop on ongoing transmissions and hijack the contents or spoof the other user) [54].

B. Network and MAC levels

A significant issue of network and MAC security is caused by a lack of robust device identification methods [54]–[56]. There are many solutions proposed for the identification [57]. These technologies can be classified into two groups: virtual identifiers and physical identifiers. Currently, the most popular identifiers (IMEI, MAC-address) are recorded into the memory of the device [58], which makes them vulnerable to cloning and tampering [59]–[62]. An alternative recently was proposed a concept of hybrid identifier [58], which is significantly more resilient to tampering and potentially may address the issue of reliable device identification in the network. A reliable identification method is required to enable the blocking of untrusted devices on MAC and network levels which reduces the risk of attacks based on accessing the network, including Sybil, tampering server or client interface, spoofing (e.g., DNS, ARP), signaling storms (redundant signaling messages), traffic bursts (e.g., extensive request or data forwarding), and de-synchronization.

Concerning the network level, IEEE 802.15.6 defines three levels of security with the focus on critical applications of BAN. According to the standard, each security level has different properties, security levels, and data frame formats. The lowest level of security is provided on level 0, which employs an unsecured data frame for communication. This level has no mechanism for data integrity, confidentiality and privacy protection, and replay defense. The next level provides authentication for enhancing security; however, data is not encrypted. Thus confidentiality and privacy issues are not addressed. Finally, the third level enables authentication and encryption, providing maximal security. The required security level can be selected when a new device associates BAN. The security mechanisms proposed in IEEE 802.15.6 support both unicast and multicast [63].

C. Application level

Software (including firmware) quality and immutability are primary concerns at the application level. Most common attacks exploit the vulnerabilities of software to inject malicious code or logical bombs, performing DoS attacks, and sniffing. It is worth noting that untrusted software producers may incorporate malicious code or a logical bomb into their application by default, which can make the user vulnerable. Beyond this, the most successful attacks at the application level are based on social engineering. This type of attacks exploits users' weaknesses, which is often much easier than hacking a well-designed application.

D. Challenges

Software secured from social engineering attacks. Software utilized in AH applications should be design in such a way to a users' actions by enabling protection from social engineering attacks, by limiting their rights in a system. Recent machine learning algorithms are expected to enable monitoring and dynamic protection from social engineering attacks [64].

Standardization of requirements and testing procedures. Security and reliability requirements for AH applications have to be standardized to provide a validated design framework for developers. New applications can then be considered as ready for AH services if an appropriate testing campaign has certified their conformity to the standard.

Device identification and validation Counterfeit devices still have a notable share in the market. Such devices may operate incorrectly and reduce the performance of the system as a whole. Thus, it is especially important to provide a robust device identification system for AH applications. Such a system can be used for blocking counterfeit and untrusted devices in the network which facilitates the security of the applications.

VI. Conclusion

Communication technologies of the past decades notably shaped social and lifestyle changes. Internet-related technologies accelerated lifestyle via efficient and prompt information exchange. Currently, in the era of IoT one may observe how connected devices have become fully autonomous, delivering advanced services to their users. Emerging IoT applications are facilitating human augmentation via enhanced sensing, increased physical power, or cognitive performance. These applications form a new area for research and development, promising to become one of the most impactful technologies in the foreseeable future.

This paper covered the main aspects of IoT technologies for human augmentation and identified possible future research directions. The topic of human augmentation is highly interdisciplinary; thus, the defined challenges are not limited to communication technologies only, and their mitigation requires efforts in ethics, security, and natural sciences. Only collaborative work on this topic enables real opportunities for human wellbeing via IoT augmentation.

References

[1] C. Xia and P. Maes, "The design of artifacts for augmenting intellect," in *Proceedings of the 4th Augmented Human International Conference*. ACM, 2013, pp. 154–161.

[2] S. Nanayakkara, J. Huber, and P. Sridhar, "Introduction," in *Assistive Augmentation*. Springer, 2018, pp. 4–7.

[3] A. H. Maslow, "A theory of human motivation." *Psychological review*, vol. 50, no. 4, p. 370, 1943.

[4] M. Leo, G. Medioni, M. Trivedi, T. Kanade, and G. M. Farinella, "Computer vision for assistive technologies," *Computer Vision and Image Understanding*, vol. 154, pp. 1–15, 2017.

[5] K. Sekaran, M. S. Khan, R. Patan, A. H. Gandomi, P. V. Krishna, and S. Kallam, "Improving the response time of m-learning and cloud computing environments using a dominant firefly approach," *IEEE Access*, vol. 7, pp. 30 203–30 212, 2019.

[6] S. Blackman, C. Matlo, C. Bobrovitskiy, A. Waldoch, M. L. Fang, P. Jackson, A. Mihailidis, L. Nygård, A. Astell, and A. Sixsmith, "Ambient assisted living technologies for aging well: a scoping review," *Journal of Intelligent Systems*, vol. 25, no. 1, pp. 55–69, 2016.

[7] Y. Koucheryavy, R. Kirichek, R. Glushakov, and R. Pirmagomedov, "Quo vadis, humanity? ethics on the last mile toward cybernetic organism," *Russian Journal of Communication*, vol. 9, no. 3, pp. 287–293, 2017.

[8] J. Lombardi, R. Malay, J. Schaffner, H. J. Song, M.-H. Huang, S. Pollard, M. Poliks, and T. Talty, "Copper transparent antennas on flexible glass by subtractive and semi-additive fabrication for automotive applications," in *2018 IEEE 68th Electronic Components and Technology Conference (ECTC)*. IEEE, 2018, pp. 2107–2115.

[9] R. Herbert, J.-H. Kim, Y. Kim, H. Lee, and W.-H. Yeo, "Soft material-enabled, flexible hybrid electronics for medicine, healthcare, and human-machine interfaces," *Materials*, vol. 11, no. 2, p. 187, 2018.

[10] A. D. Valentine, T. A. Busbee, J. W. Boley, J. R. Raney, A. Chortos, A. Kotikian, J. D. Berrigan, M. F. Durstock, and J. A. Lewis, "Hybrid 3d printing of soft electronics," *advanced Materials*, vol. 29, no. 40, p. 1703817, 2017.

[11] Y. M. Roshan and E. J. Park, "Design approach for a wireless power transfer system for wristband wearable devices," *IET Power Electronics*, vol. 10, no. 8, pp. 931–937, 2017.

[12] S. R. Gutbrod, M. S. Sulkin, J. A. Rogers, and I. R. Efimov, "Patient-specific flexible and stretchable devices for cardiac diagnostics and therapy," *Progress in biophysics and molecular biology*, vol. 115, no. 2-3, pp. 244–251, 2014.

[13] I. F. Akyildiz, F. Brunetti, and C. Blázquez, "Nanonetworks: A new communication paradigm," *Computer Networks*, vol. 52, no. 12, pp. 2260–2279, 2008.

[14] J. M. Jornet and I. F. Akyildiz, "Graphene-based plasmonic nano-antenna for terahertz band communication in nanonetworks," *IEEE Journal on selected areas in communications*, vol. 31, no. 12, pp. 685–694, 2013.

[15] J. M. Jornet and I. F. Akyildiz, "Graphene-based plasmonic nano-transceiver for terahertz band communication," in *Proc. of 8th European Conference on Antennas and Propagation (EuCAP)*. IEEE, 2014, pp. 492–496.

[16] J.-S. Moon, H.-C. Seo, K.-A. Son, B. Yang, D. Le, H. Fung, and A. Schmitz, "Zero-bias THz detection using graphene transistors," in *Proc. of MTT-S International Microwave Symposium (IMS)*. IEEE, 2015, pp. 1–4.

[17] T. Le, T. Thai, V. Lakafosis, M. Tentzeris, Z. Lin, Y. Fang, K. Sandhage, and C. Wong, "Graphene enhanced wireless sensors," in *IEEE Sensors*. IEEE, 2012, pp. 1–4.

[18] G. Chen, T. M. Paronyan, E. M. Pigos, and A. R. Harutyunyan, "Enhanced gas sensing in pristine carbon nanotubes under continuous ultraviolet light illumination," *Scientific reports*, vol. 2, p. 343, 2012.

[19] I. F. Akyildiz and J. M. Jornet, "Electromagnetic wireless nanosensor networks," *Nano Communication Networks*, vol. 1, no. 1, pp. 3–19, 2010.

[20] J. Kokkoniemi, J. Lehtomäki, K. Umebayashi, and M. Juntti, "Frequency and time domain channel models for nanonetworks in terahertz band," *IEEE Transactions on Antennas and Propagation*, vol. 63, no. 2, pp. 678–691, 2015.

[21] K. Yang, A. Pellegrini, M. O. Munoz, A. Brizzi, A. Alomainy, and Y. Hao, "Numerical analysis and characterization of THz propagation channel for body-centric nano-communications," *IEEE Transactions on Terahertz Science and technology*, vol. 5, no. 3, pp. 419–426, 2015.

[22] A. Mosenia, S. Sur-Kolay, A. Raghunathan, and N. K. Jha, "Wearable medical sensor-based system design: A survey," *IEEE Transactions on Multi-Scale Computing Systems*, vol. 3, no. 2, pp. 124–138, 2017.

[23] T. Rault, A. Bouabdallah, and Y. Challal, "Energy efficiency in wireless sensor networks: A top-down survey," *Computer Networks*, vol. 67, pp. 104–122, 2014.

[24] M. R. Nakhkash, T. N. Gia, I. Azimi, A. Anzanpour, A. M. Rahmani, and P. Liljeberg, "Analysis of performance and energy consumption of wearable devices and mobile gateways in iot applications," in *Proceedings of the International Conference on Omni-Layer Intelligent Systems*. ACM, 2019, pp. 68–73.

[25] J. Williamson, Q. Liu, F. Lu, W. Mohrman, K. Li, R. Dick, and L. Shang, "Data sensing and analysis: Challenges for wearables," in *The 20th Asia and South Pacific Design Automation Conference*. IEEE, 2015, pp. 136–141.

[26] M. Javed, G. Ahmed, D. Mahmood, M. Raza, K. Ali, and M. Ur-Rehman, "Taeo-a thermal aware & energy optimized routing protocol for wireless body area networks," *Sensors*, vol. 19, no. 15, p. 3275, 2019.

[27] R. Pirmagomedov, I. Hudoev, and D. Shangina, "Simulation of medical sensor nanonetwork applications traffic," in *International Conference on Distributed Computer and Communication Networks*. Springer, 2016, pp. 430–441.

[28] T. M. Behera, S. K. Mohapatra, U. C. Samal, M. S. Khan, M. Daneshmand, and A. H. Gandomi, "Residual energy based cluster-head selection in wsns for iot application," *IEEE Internet of Things Journal*, 2019.

[29] R. Pirmagomedov, M. Blinnikov, R. Glushakov, A. Muthanna, R. Kirichek, and A. Koucheryavy, "Dynamic data packaging protocol for real-time medical applications of nanonetworks," in *Internet of Things, Smart Spaces, and Next Generation Networks and Systems*. Springer, 2017, pp. 196–205.

[30] Q. Xu, T. Mytkowicz, and N. S. Kim, "Approximate computing: A survey," *IEEE Design & Test*, vol. 33, no. 1, pp. 8–22, 2015.

[31] R. Pirmagomedov, M. Blinnikov, A. Amelyanovich, R. Glushakov, S. Loskutov, A. Koucheryavy, R. Kirichek, and E. Bobrikova, "Iot based earthquake prediction technology," in *Internet of Things, Smart Spaces, and Next Generation Networks and Systems*. Springer, 2018, pp. 535–546.

[32] L. Piwek, D. A. Ellis, S. Andrews, and A. Joinson, "The rise of consumer health wearables: promises and barriers," *PLoS Medicine*, vol. 13, no. 2, p. e1001953, 2016.

[33] J.-M. Lee, Y.-W. Kim, and G. J. Welk, "Track it: Validity and utility of consumer-based physical activity monitors," *ACSM's Health & Fitness Journal*, vol. 18, no. 4, pp. 16–21, 2014.

[34] C. Cérin, C. Coti, P. Delort, F. Diaz, M. Gagnaire, Q. Gaumer, N. Guillaume, J. Lous, S. Lubiarz, J. Raffaelli *et al.*, "Downtime statistics of current cloud solutions," *International Working Group on Cloud Computing Resiliency, Tech. Rep*, 2013.

[35] K. V. Vishwanath and N. Nagappan, "Characterizing cloud computing hardware reliability," in *Proceedings of the 1st ACM symposium on Cloud computing*. ACM, 2010, pp. 193–204.

[36] A. Zhou, S. Wang, B. Cheng, Z. Zheng, F. Yang, R. N. Chang, M. R. Lyu, and R. Buyya, "Cloud service reliability enhancement via virtual machine placement optimization," *IEEE Transactions on Services Computing*, vol. 10, no. 6, pp. 902–913, 2016.

[37] V. Singanamalla, R. Patan, M. S. Khan, and S. Kallam, "Reliable and energy-efficient emergency transmission in wireless sensor networks," *Internet Technology Letters*, vol. 2, no. 2, p. e91, 2019.

[38] M. Khan, A. Kumar, and B. Xie, "Stitching algorithm: A network performance analysis tool for dynamic mobile networks," *Procedia Technology*, vol. 3, pp. 41–51, 2012.

[39] P. Rashidi and A. Mihailidis, "A survey on ambient-assisted living tools for older adults," *IEEE journal of biomedical and health informatics*, vol. 17, no. 3, pp. 579–590, 2012.

[40] A. Ylisaukko-oja, E. Vildjiounaite, and J. Mantyjarvi, "Five-point acceleration sensing wireless body area network-design and practical experiences," in *Eighth International Symposium on Wearable Computers*, vol. 1. IEEE, 2004, pp. 184–185.

[41] C. C. Poon, Y.-T. Zhang, and S.-D. Bao, "A novel biometrics method to secure wireless body area sensor networks for telemedicine and m-health," *IEEE Communications Magazine*, vol. 44, no. 4, pp. 73–81, 2006.

[42] A. Kovtunenko, M. Timirov, and A. Bilyalov, "Multi-agent approach to computational resource allocation in edge computing," in *Internet of Things, Smart Spaces, and Next Generation Networks and Systems*. Springer, 2019, pp. 135–146.

[43] B. Atakan, O. B. Akan, and S. Balasubramaniam, "Body area nanonetworks with molecular communications in nanomedicine," *IEEE Communications Magazine*, vol. 50, no. 1, pp. 28–34, 2012.

[44] K. Venugopal, M. C. Valenti, and R. W. Heath, "Analysis of millimeter wave networked wearables in crowded environments," in *2015 49th Asilomar Conference on Signals, Systems and Computers*. IEEE, 2015, pp. 872–876.

[45] K. Venugopal and R. W. Heath, "Millimeter wave networked wearables in dense indoor environments," *IEEE Access*, vol. 4, pp. 1205–1221, 2016.

[46] K. O. Olasupo, I. Kostanic, C. E. Otero, and T. O. Olasupo, "Link performance modeling of wireless sensor network deployment for mission-critical applications (underground deployment)," in *2017 16th Annual Mediterranean Ad Hoc Networking Workshop (Med-Hoc-Net)*. IEEE, 2017, pp. 1–7.

[47] N. Brahmi, O. N. Yilmaz, K. W. Helmersson, S. A. Ashraf, and J. Torsner, "Deployment strategies for ultra-reliable and low-latency communication in factory automation," in *2015 IEEE Globecom Workshops (GC Wkshps)*. IEEE, 2015, pp. 1–6.

[48] A. Orsino, A. Ometov, G. Fodor, D. Moltchanov, L. Militano, S. Andreev, O. N. Yilmaz, T. Tirronen, J. Torsner, G. Araniti *et al.*, "Effects of heterogeneous mobility on d2d-and drone-assisted mission-critical mtc in 5g." *IEEE Communications Magazine*, vol. 55, no. 2, pp. 79–87, 2017.

[49] G. Wang, G. Feng, S. Qin, and R. Wen, "Efficient traffic engineering for 5g core and backhaul networks," *Journal of Communications and Networks*, vol. 19, no. 1, pp. 80–92, 2017.

[50] O. Galinina, A. Pyattaev, K. Johnsson, A. Turlikov, S. Andreev, and Y. Koucheryavy, "Assessing system-level energy efficiency of mmwave-based wearable networks," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 923–937, 2016.

[51] I. Recommendation, "Security architecture for systems providing end-to-end communications," 2002.

[52] X. Zhang, A. Kunz, and S. Schröder, "Overview of 5g security in 3gpp," in *2017 IEEE Conference on Standards for Communications and Networking (CSCN)*. IEEE, 2017, pp. 181–186.

[53] R. P. Jover and V. Marojevic, "Security and protocol exploit analysis of the 5g specifications," *IEEE Access*, vol. 7, pp. 24 956–24 963, 2019.

[54] F. Hussain, L. Ferdouse, A. Anpalagan, L. Karim, and I. Woungang, "Security threats in m2m networks: A survey with case study," *COMPUTER SYSTEMS SCIENCE AND ENGINEERING*, vol. 32, no. 2, pp. 117–135, 2017.

[55] A. Paranjothi, M. S. Khan, and M. Nijim, "Survey on three components of mobile cloud computing: offloading, distribution and privacy," *Journal of Computer and Communications*, vol. 5, no. 06, p. 1, 2017.

[56] S. A. Salehi, M. A. Razzaque, I. Tomeo-Reyes, and N. Hussain, "Ieee 802.15.6 standard in wireless body area networks from a healthcare point of view," in *2016 22nd Asia-Pacific Conference on Communications (APCC)*, Aug 2016, pp. 523–528.

[57] M. G. Samaila, J. B. Sequeiros, M. M. Freire, and P. R. Inácio, "Security threats and possible countermeasures in iot applications covering different industry domains," in *Proceedings of the 13th International Conference on Availability, Reliability and Security*. ACM, 2018, p. 16.

[58] S. S. Vladimirov, R. Pirmagomedov, R. Kirichek, and A. Koucheryavy, "Unique degradation of flash memory as an identifier of ict device," *IEEE Access*, vol. 7, pp. 107 626–107 634, 2019.

[59] S. P. Rao, S. Holtmanns, I. Oliver, and T. Aura, "Unblocking Stolen Mobile Devices Using SS7-MAP Vulnerabilities: Exploiting the Relationship between IMEI and IMSI for EIR Access," in *2015 IEEE Trustcom/BigDataSE/ISPA*, vol. 1, Aug 2015, pp. 1171–1176.

[60] I. Gepko, "General requirements and security architecture for mobile phone anti-cloning measures," in *IEEE EUROCON 2015 - International Conference on Computer as a Tool (EUROCON)*, Sept 2015, pp. 1–6.

[61] S. Vaidya and K. J. Christensen, "A single system image server cluster using duplicated MAC and IP addresses," in *Proceedings LCN 2001. 26th Annual IEEE Conference on Local Computer Networks*, Nov 2001, pp. 206–214.

[62] A. Cherchali, M. J. Gudelis Jr, W. G. Lester, and R. J. McLaughlin, "Technique for automated MAC address cloning," Sep. 1 2015, US Patent 9,124,474.

[63] K. S. Kwak, S. Ullah, and N. Ullah, "An overview of ieee 802.15. 6 standard," in *2010 3rd International Symposium on Applied Sciences in Biomedical and Communication Technologies (ISABEL 2010)*. IEEE, 2010, pp. 1–6.

[64] R. Heartfield and G. Loukas, "A taxonomy of attacks and a survey of defence mechanisms for semantic social engineering attacks," *ACM Computing Surveys (CSUR)*, vol. 48, no. 3, p. 37, 2016.

PUBLICATION

II

Unique Degradation of Flash Memory as an Identifier of ICT Device

S. S. Vladimirov, R. Pirmagomedov, R. Kirichek and A. Koucheryavy

IEEE Access 7.(2019), 107626–107634
DOI: 10.1109/ACCESS.2019.2932804

IEEE Access
Multidisciplinary ⁝ Rapid Review ⁝ Open Access Journal

Received May 22, 2019, accepted July 29, 2019, date of publication August 2, 2019, date of current version August 19, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2932804

Unique Degradation of Flash Memory as an Identifier of ICT Device

SERGEY S. VLADIMIROV[1], RUSTAM PIRMAGOMEDOV[2,3], RUSLAN KIRICHEK[1], AND ANDREY KOUCHERYAVY[1]

[1]Department of Communication Networks, St. Petersburg State University of Telecommunications, 193232 St. Petersburg, Russia
[2]Applied Mathematics & Communications Technology Institute, Peoples Friendship University of Russia (RUDN University), 117198 Moscow, Russia
[3]Electrical Engineering Unit, Tampere University, FI-33720 Tampere, Finland

Corresponding author: Rustam Pirmagomedov (rustam.pirmagomedov@tuni.fi)

The publication was prepared with the support of the ‘‘RUDN University Program 5-100’’.

ABSTRACT The counterfeit and stolen information and communication technologies (ICT) devices are an essential and growing problem. Reliable technology for the identification of ICT devices is required to enable blocking of these devices in the network worldwide. Motivated by this challenge, we elaborate on the idea of using the unique degradation image of flash memory chip (DFMC) as the identifier of the device. This idea is based on the assumption that the distribution of degraded segments in the memory chip is unique enough to provide reliable identification of the device. In this paper, we provide a proof of concept through a hardware experiment. For this experiment, we developed a custom test bed and special software enabling the forced degradation of NOR-flash memory chips. We, then, consider the uniqueness of such identifiers using combination theory and consider practical issues of DFMC implementation, including the initial identification procedure, light dynamic identifiers, and identification using a cross-correlation function and options of dynamic identification. We conclude that using DFMC addresses relevant challenges of ICT devices identification.

INDEX TERMS Internet of things, counterfeiting, system identification, flash memory cells, physical identification, communication system security, network security.

I. INTRODUCTION

The recent development of electronics has led to active penetration of information and communication technologies (ICT) to all spheres of life (e.g., automatization of industrial processes, e-health, smart homes, smart wearables) [1]–[5]. In the beginning, ICT devices mostly had an entertainment purpose. Now there are many mission-critical applications, such as e-health [6] or remote surgery [7]. These applications need more rigid requirements for reliability, security, and quality of services [6], [8]. To satisfy these higher requirements, manufacturers made significant efforts to improve their devices [9]. Consequently, these devices tend to be more expensive and often become a target of counterfeiting.

The counterfeiting of ICT devices is a significant and growing problem according to recent report of the International Telecommunication Union (ITU) [10]. The report shows that investments in the economy, employment, right holders and users are suffering economically from counterfeiting. Moreover, counterfeit devices cause severe health and safety risks due to the use of low-quality materials and incorrect manufacturing processes.

International standardization bodies consider combating counterfeits as a crucial technological challenge for modern society. For example, combating counterfeit and stolen ICT equipment was the aim of Q15/11 ITU-T. Current efforts are focused on developing solutions to block (or deny ICT services) any stolen or counterfeit ICT devices worldwide. Reliable technology for the identification of ICT devices is required to enable these solutions.

Existing identification solutions can be classified into two subordinate groups: virtual identifiers and physical identifiers. The implementation of physical methods is limited by the local network since these require direct contact. The virtual methods are easily scalable, thus can be employed for any size networks. However, during the ITU workshop ‘‘Global approaches on combating counterfeiting and stolen ICT devices’’ held on 23 July 2018 in Geneva[1] it was noted that virtual identifiers (e.g. IMEI and other identifiers

The associate editor coordinating the review of this manuscript and approving it for publication was Mamoun Alazab.

[1] https://www.itu.int/dms_pub/itu-t/opb/tut/T-TUT-CCICT-2015-PDF-E.pdf

recorded in erasable memory) are not robust enough to prevent tampering.

In this paper, we elaborate on the recently presented approach of using a degraded segment in flash memory chips as an identifier of the device [11]. The idea is based on the assumption that the distribution of degraded segments in the memory chip is unique enough to provide reliable identification of the device [11]–[13]. The main objectives of this paper are: (i) to provide a proof of concept through hardware experiments, (ii) to define approaches of using the degraded flash memory as an identifier of ICT devices and (iii) to define essential challenges for future work.

The paper is structured as follows. In Section II we provide an overview of the identification systems. In Section III we describe the proposed method. In Section IV, we provide results of the hardware testing of the proposed method and study the repeatability of the identifiers. We consider use case scenario of the method and some implementation issues in Section V. In Section VI we conclude the paper.

II. RELATED WORKS

Identification of the ICT devices is an essential enabler for the management of network and services. There are myriad of technologies developed for the identification. These technologies can be classified into two groups: virtual identifiers and physical identifiers. In order to set the niche for the proposed approach, we first provide an overview of notable identification systems.

A. VIRTUAL IDENTIFIERS

The virtual methods can be roughly classified to software-based fingerprinting and virtual identifiers which are recorded in devices. The software-based fingerprinting methods rely on differences in the software configuration of an ICT device. Internet companies widely utilize such a method for tracking users and targeted advertising. The software-based fingerprint consists of a number of parameters such as IP address, operating system version, device location and time settings, battery status, screen resolution, touch support, language setup, etc. Utilizing all of these parameters together creates a unique identifier for an ICT device. The more parameters used in the software-based fingerprinting, the more unique the identifier. Thus, in the case of a simple device, such as IoT sensor device, the software-based fingerprint will consist of only a few parameters and the identifier will be weak. Therefore, implementation of the software-based fingerprinting is limited to relatively powerful devices such as smartphones, laptops or tablet PCs.

The virtual identifiers which are recorded in devices are applicable for network identification of major types of the devices, including simple IoT devices. Currently, the virtual identifiers are recorded into the memory of the ICT device during the manufacturing process [14], [15]. For example, Ethernet MAC-address or International Mobile Equipment Identity (IMEI). The latter is the most widespread identifier type for mobile ICT devices. However, these identifier are vulnerable to cloning and tampering [16]–[19].

In [20], and [21] approaches are discussed for the detection of devices with the same MAC address, utilizing analysis of frame sequence. However, these approaches can be implemented only in a local area and are not applicable for the identification of the devices beyond a local network. Moreover, since all of the approaches are software-based, they can be hacked by using a different system configuration or communication behavior.

B. PHYSICAL IDENTIFIERS

Alternatively to virtual methods, the physical methods utilize the individuality of ICT device hardware. Notable examples of the physical methods are: integrated clock skew estimation, radio-frequency fingerprinting, and utilizing degradation uniqueness of the flash memory.

An identification method based on the integrated clock skew was proposed by Kohno *et al.* [22]. The method employs TCP and ICMP time stamps, thus, the identification can be tampered by means of software (e.g., TCP and ICMP time stamps used by the method could be altered or disabled) [23].

The radio-frequency fingerprinting came from the military sphere, where transmitter individuality has been used to distinguish between friendly and enemy radars since the Vietnam War era [23]. Later, similar systems were implemented in mobile networks to prevent access from unauthorized phones [24]. Further, K. Remley et al. experimentally showed differences between signals of different 802.11 transmitters [25]. The experiments proved that RF fingerprinting methods could be used to identified features of various interfaces in an anechoic chamber. However, in a practical case, the fingerprints can be altered by environmental effects, such as interference with other nodes. Moreover, a considerable part of ICT devices connect to the operator network utilizing intermediate nodes (e.g., gateways, relays in mesh networks), while the RF fingerprinting methods require a direct connection to a trustable authority (e.g. service provider). Thus, the RF fingerprinting is suitable for only highly specific purposes while the IoT era requires mass methods.

C. THE IDENTIFICATION OF ICT DEVICE USING UNIQUE PROPERTIES OF FLASH MEMORY CHIPS

The concept of physical unclonable functions (PUFs) and their implementation using integrated circuits introduced in [26] provides a theoretical foundation for the authentication device using individual properties of flash memory chips. The theory was developed further in [12] and [13], where authors proposed to utilize variations in threshold voltages as a unique property of a NAND flash memory chip. The proposed approach was considered as an enabler for several application scenarios: the random number generator, generator of cryptographic key and device fingerprinting. However, the voltage variations may significantly change in time due to the degradation of the chip. Thus the proposed approach can be efficient only for short-term use.

Alternative PUF-based method relies on degraded blocks of memory which appeared randomly due to microscopic manufacturing defects in a NAND memory chip [27]. A pattern of the degraded memory cells was suggested for consideration as a physically unique identifier. However, the uniqueness of such an identifier was not numerically evaluated. Moreover, using a NAND flash drive for such a method of identification is complicated because modern NAND memory chips contain a special microcontroller to prevent the degradation of memory blocks and prevent using degraded memory blocks [28]. Finally, it should be noted that while NAND flash memory is used in storage devices such as USB-flash drives or SSD drives, which are widely employed in complex computer systems, it is rarely used in simple devices (e.g., most of IoT devices) due to technological reasons [28].

Further development of PUF-based methods is related to NOR flash memory. In contrast with NAND, the life span of NOR flash memory is about 10 times shorter; thus it is degrading faster. Moreover, NOR flash memory widely is used in network devices including resource-constrained IoT devices for storing the device microprogram. In our previous work [11], it was proposed to use a forcibly degraded segment of a NOR flash memory chip as a unique identifier of the whole network device. In this paper, we elaborate on the idea and provide hardware testing.

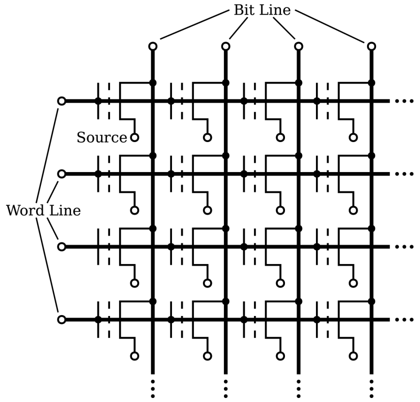


FIGURE 1. Cell architecture of a NOR-flash memory [32].

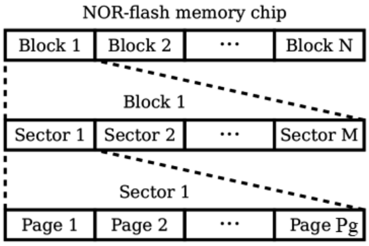


FIGURE 2. Hierarchic structure of a NOR-flash memory chip.

III. METHOD DESCRIPTION

A. THE DEGRADATION OF FLASH MEMORY

The NOR-flash memory is a two-dimensional array of low-level memory cells, located on the matrix of the conductor [29], [30]. Each of these low-level cells stores from one to four bits of information, depending on the type of the cell. To record or erase the information in the memory cell its charge must be changed [31]. The architecture of a NOR flash memory cell presented in Fig. 1. During the recording process, specific cells change the initial state of a bit (usually it is "1"), to the opposite ("0"). Each change in the state of charge of the cell causes the accumulation of irreversible changes in its structure, and at a particular moment, a cell may cease to change its state. Further, we name such cells as bad-cells and the process of the appearance of bad-cells as the degradation of a memory chip.

The bad-cells in the NOR-flash memory always retain the same state of charge regardless of the erasing/recording procedures that are carried-out [28], [33]. Occasionally, separate bad-cells may change their charge during the recording process. However, during the subsequent erasing/recording, they again drop to their constant initial state.

B. MEMORY CHIP DESIGN

The NOR-flash memory chip has a hierarchic structure, as shown in Fig. 2. The total volume of the chip can be determined by a formula $N \times M \times Pg \times V_{Pg}$, where N is the amount of the memory blocks, M is the amount of memory sectors in one block, and P is the amount of memory pages with the volume V_{Pg} (Bytes) in one sector [30].

A sector is the minimal erasable part of a NOR-flash memory. Consequently, to change at least one bit of information in the chip, the following algorithm must be utilized [12], [28], [33]:

- Read the whole memory sector;
- Change the necessary information;
- Erase the whole read sector;
- Record the changed information in the erased memory sector (page-by-page).

C. IDENTIFICATION PROCEDURE

An image (S) of a degraded memory sector (the map of bad-cells in the sector) can be used to establish a procedure of chip identification. Since flash memory chips are used in most modern ICT devices, one memory sector in the chip can be allocated for the identification purpose (for example, the first or last sector) and then a chip, and consequently, a device can be identified [11].

It should contain a certain number of bad-cells before being installed into the device to enable using the chip as an identifier. These bad-cells can be created by overwriting the memory sector until the appearance of stable bad-cells (forcible degradation). Since the device with the degraded chip inside is manufactured, the image (S) must be recorded by the manufacturer into the database of the released devices.

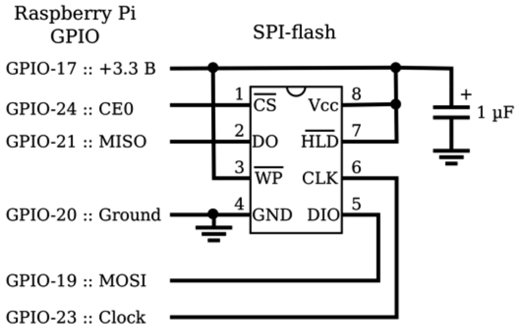


FIGURE 3. Connection scheme of NOR-flash memory chips to the SPI interface of Raspberry Pi 3.

TABLE 1. The matrix of bad-cells matches.

C	Identifiers $S(C)$											
	$S(1)$	$S(2)$	S(3)	S(4)	S(5)	S(6)	S(7)	S(8)	S(9)	S(10)	S(11)	S(12)
1	43	0	0	1	0	0	0	0	0	0	0	0
2	0	95	0	1	1	1	1	0	2	1	2	1
3	0	0	36	2	0	0	1	0	0	0	0	0
4	1	1	2	32	0	0	2	0	0	0	0	0
5	0	1	0	0	48	0	4	1	0	0	0	0
6	0	1	0	0	0	62	0	0	0	2	0	2
7	0	1	1	2	4	0	348	3	0	0	0	0
8	0	0	0	0	1	0	3	62	0	0	0	0
9	0	2	0	0	0	0	0	0	117	0	0	0
10	0	1	0	0	0	2	0	0	0	49	0	0
11	0	2	0	0	0	0	0	0	0	0	81	0
12	0	1	0	0	0	2	0	0	0	0	0	39

Throughout the lifetime of the devices, the sector allocated for identification should be accessible for remote reading by third-party (e.g., communication services provider, authorities). Afterward, the pattern of bad-cells can be checked in the database (for instance, it can be a corporate, local, national, or international register of the devices). If the bad-cell pattern is not in the database, the device will be denied of service.

IV. PROOF OF CONCEPT TESTING AND DUPLICATION EVALUATION OF IDENTIFIERS

A. HARDWARE TESTING OF THE PROPOSED METHOD

For testing of the proposed method, a hardware experiment was performed. During the experiment, a forced degradation of NOR-flash memory chips was executed. All chips in the experiment were connected to a custom test bed. A Raspberry Pi 3 (Raspbian Linux OS) was selected for this trial. The chips were connected to the SPI interface of the Raspberry according to the standard scheme shown in Fig. 3. To enable the testing, a custom library of functions (pi-spiflash) was developed and is currently available for public. This library is currently uploaded on the Github[2] for free use.

The testing was performed using 2 Mbit chips W25X16VSIG, W25X16AVSIG, and W25Q16VSIG with a standard command system. There were a total of 120 chips in the experiment. The chips had the similar memory structure: N = 32 blocks; M = 16 sectors in a block (512 sectors per chip); K = 16 pages in a sector, (8192 pages per a chip). The volume of each page is 256 bytes. The first memory sector (4096 bytes with the address 000000h) of each chip was forcibly degraded through multiple rewriting.

During the testing, each chip was subjected to a forced degradation (350,000 write cycles for each chip). The degradation affected only the first memory sector. As a result of forced degradation, a set of images (bad-cell maps) of the first memory sector, was obtained. The number of bad-cells for most chips tested (111 pieces) ranged from 30 to 100.

The next phase was the identification of every chip using the image of the first sector. After the images of the first sector were read, positions of the bad-cells were compared among all the chips being tested. To illustrate it, Table 1 shows the matrix of bad-cell matches for 12 randomly picked chips.

Rows of Table 1 correspond with numbers C of chips. Columns correspond with the identifiers $S(C)$ of these chips. The intersection of rows and columns shows how many bad-cells were matched. To find the matches, we erased the first sector of each of 12 selected chips and then compared them with each identifier. For example, identifier $S(3)$ refers to chip no.3 (36 bad-cells matched), but also two bad-cells of $S(3)$ are matched with chip no.4 and one bad-cell matched with chip no.7. The obtained results show that we can correctly identify every chip using majority decision.

Also we get mutual correlation (including autocorrelation) of defined identifiers $S(C)$. The graphs of the normalized value of mutual correlation of identifiers $S(1)$–$S(12)$ are plotted (Fig. 4). On the ordinate axis, the correlation value is indicated, where one corresponds to the complete coincidence of the bad-cell positions. The abscissa axis indicates the numbers $C = \{1, 2, \ldots, 12\}$ of identifiers (chips) to be compared.

From Fig. 4 it can be seen that the rate of duplication in the positions of bad-cells between different identifiers within the set of chips is negligible. Consequently, the experiment proved that the proposed method could be used for identification of the devices.

B. THE REPEATABILITY OF THE IDENTIFIERS

The repeatability of the identifiers depends on the several parameters: (i) size of the memory area allocated for identification (the larger it is, the less repeatability), (ii) number of devices to be identified, (iii) the share of bad-cells. If the number of bad-cells m in sector used for identification is in a range from m_1 to m_2 and degradation (bad-cells) occurs equiprobable amongst cells, then the total number of unique combinations of the bad-cells can be estimated using

$$K = \sum_{m=m_1}^{m_2} C_T^m, \tag{1}$$

[2] https://github.com/vlad-ss/pi-spiflash

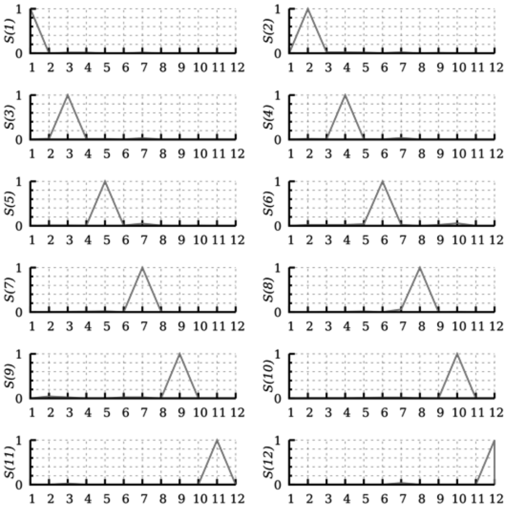


FIGURE 4. The correlation of identifiers.

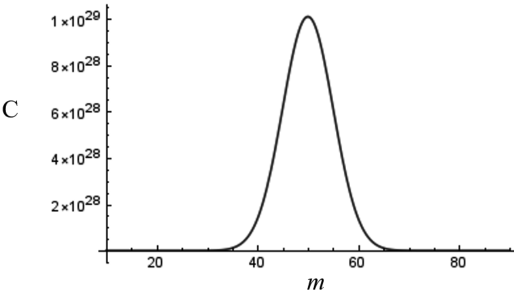


FIGURE 5. Unique combinations of bad-cells for $T = 100$.

where m is amount bad-cells, T is the total amount of cells in a sector of chip allocated for identification, and C is the number of unique combinations of the bad-cells. To determine the probability of the appearance of two equal identifiers (δ) among devices (d), the solution of "birthday paradox" [34] can be utilized

$$\delta = 1 - \frac{K!}{K^d (K-d)!}, \tag{2}$$

where d is a total number of devices to be identified.

The target bad-cells shared in a sector, which should be reached during degradation process, can be determined by a simplified example: let us assume a total number of cells in a sector (T) is 100. By changing the number of bad-cells in the sector from 1 to 99, the number of unique combinations of bad-cells (C_{100}^{m}) was estimated and presented on a graph (Fig. 5). The graph clearly shows that the highest value of unique combination can be reached if the share of bad-cells is 50% of the total cells in the sector.

Further, the probability of the appearance of two equal identifiers (δ) has to be estimated. The calculation of (δ) using equation 2 required high computational performance, even for the simplified example. As an alternative, equation 2 can be approximated by the expansion of an exponential function in a Taylor series

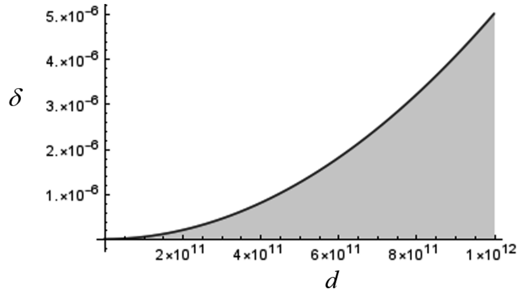


FIGURE 6. The probability of appearance at least two equal identifiers for $T = 100$.

$$\delta \approx 1 - e^{-\frac{d^2}{2\times K}}. \tag{3}$$

For the calculation of (3), $K = 10^{29}$ was taken, it corresponded to $m = 50$. The calculation made for variable d is presented on a graph (Fig. 6). Here, the probability of the appearance two equal identifiers is extremely low even if the total number of devices to be identified is one trillion ($d = 10^{12}$).

It is must be considered that the continuous procedure of erase-and-read of the allocated sector during the flash memory chip lifetime can lead to the appearance of new bad-cells. As it follows from the experiment described in Section IV.A the mean number of the erase cycles required for appearance a new bad-cell is about 3940. This number of cycles corresponds to more than ten years of device use with daily identification. However, if the bad-cell appeared during the lifetime of a device and there is another device with an ID (bad-cell distribution) differing by one bad-cell, the identification process can be failed for these devices. The possibility of such an event is following from (3), and can be estimated using formula (4)

$$\delta \approx 1 - (e^{-\frac{d^2}{2\times C_T^m}} \times e^{-\frac{d^2}{2\times C_T^{m-1}}}) \tag{4}$$

For general case (n of new bad-cells appeared), the (4) can be expressed in the following form:

$$\delta \approx 1 - \prod_{i=1}^{n} e^{-\frac{d^2}{2\times C_T^{m-i}}} \tag{5}$$

In figure Fig. 7 we provide numerical results for the (5) within the considered simplified scenario ($T = 100$).

As it can be seen from the Fig. 7, the probability of appearance at least two equal identifiers slight even if five new bad-cells appeared during the lifetime of the device. Moreover, this calculation was made for the simplified scenario ($T = 100$), whereas the real number of cells in a sector is $T = 32768$. Consequently, in a real case, the probability of the appearance of at least two equal identifiers will be statistically infinitesimal.

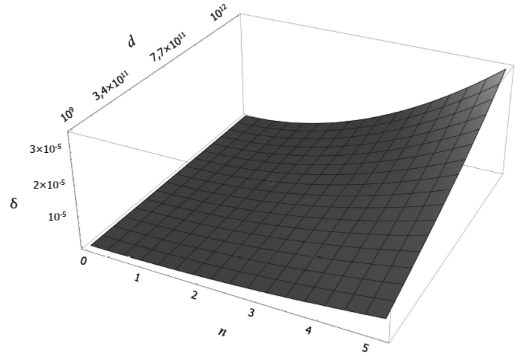


FIGURE 7. The probability of appearance at least two equal identifiers because of the natural degradation of memory chips.

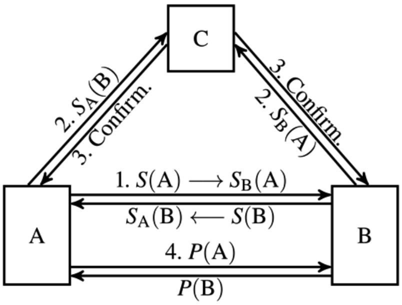


FIGURE 8. The scheme of the identification.

V. CONSIDERATION OF IMPLEMENTATION ISSUES

A. MUTUAL AUTHENTICATION OF TWO NETWORK DEVICES

In a case of M2M communication, devices have to identify each other. The scheme of mutual identification for two ICT devices with the use of external trusted identifiers' database is shown in Fig. 8.

The procedure of identification for this use case is executed in four steps:

1) Network devices A and B get each other's full identifiers $S(A)$ and $S(B)$ correspondingly;
2) Each network device checks the received full identifier in the trusted database (C);
3) After the confirmation from the C, network devices start the data exchange process;
4) Regular confirmation utilizes the short identifiers $P(A)$ and $P(B)$.

Therefore, a differentiated approach can be applied: for (i) initial identification devices use full identifiers S, for (ii) dynamic or confirmatory identification with the use of short identifiers P.

B. THE DYNAMIC IDENTIFICATION USING A CROSS-CORRELATION FUNCTION

The use of short identifiers for dynamic identification has its drawbacks, among which there is the necessity to transfer a large amount of data (identifier P, which takes up one page, already requires the transfer at least 256 bytes). The alternative solution for dynamic identification is the use of a specific mathematical function using S as a parameter. This function should be known to the devices beforehand. That function may be the cross-correlation (CCF) function for two different sectors of the identifier S of one side or the cross-correlation function for having one address of identifier's sectors of each side of the transfer.

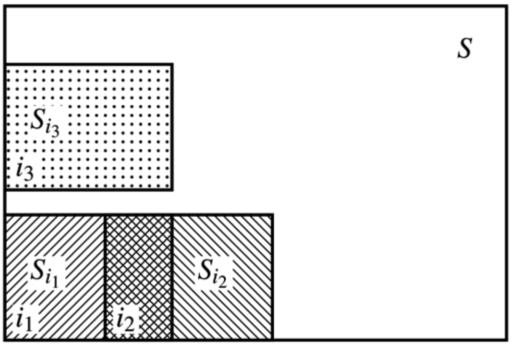


FIGURE 9. Example of rectangular sectors S_i.

Let us consider the example of using such dynamic identifiers. Assume that $S(A)$ and $S(B)$ are identifiers of the network devices A and B respectively. $S_B(A)$ is the image of $S(A)$ in the memory of the device B, received and confirmed during the initial identification. Likewise, $S_A(B)$ is the image $S(B)$ in the memory of the device A. During the dynamic identification of the device A by the device B, a particular sector of the identifier with the address i is used. The device B sends the identification request mentioning address i to device A. Then, the device A calculates the dynamic identifier $D_i(A)$ as a cross-correlation function of the corresponding sector of its own identifier $S_i(A)$ and the image $S_{Ai}(B)$

$$D_i(A) = \mathrm{CCF}[S_i(A), S_{Ai}(B)]. \qquad (6)$$

The calculated identifier $D_i(A)$ is transferred to the device B, which calculates the corresponding dynamic identifier $D_{Bi}(A)$, using its identifier $S_i(B)$ and the image $S_{Bi}(A)$

$$D_{Bi}(A) = \mathrm{CCF}[S_{Bi}(A), S_i(B)]. \qquad (7)$$

If the identifier $D_i(A)$, received from A and the identifier $D_{Bi}(A)$ calculated by B are equal, then identification of the device A by the device B is completed successfully.

The sectors S_i, used for calculating the dynamic identifier D_i, can be chosen in different ways. The first option is the use of the rectangular sector as it is shown in Fig. 9.

The second way involves the use of cross sectors. In this case, the address i points to a memory cell in the S. The cell i is a crossing point of a line and a column, this line and column form a sector S_i. The example of such sectors is shown in Fig. 10.

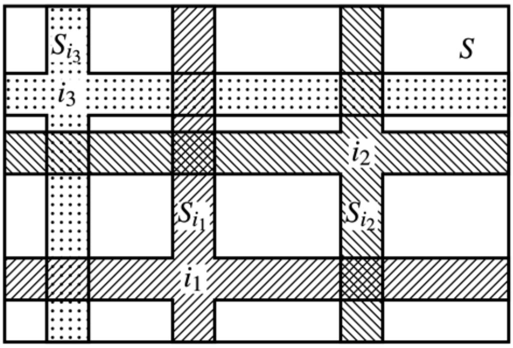


FIGURE 10. Example of the "cross" sectors S_i.

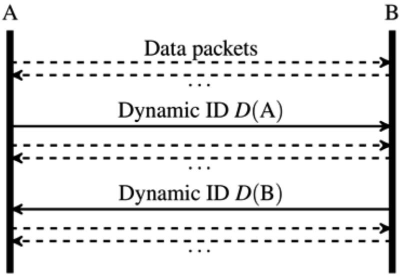


FIGURE 11. Diagram of the "confirmatory" dynamic identification.

If the devices A and B are synchronized with each other, it is not required to transfer the address i with every message. The i can be established at the initial identification and each data package, which was successfully received, will change the address i by one step. This sequential address changing can be performed by shifting registers.

It is worth noting that the dynamic identification is not limited to using CCF. The cross-correlation function was chosen to demonstrate a simple and universal way to compare identifiers with their casts. The method considered in this subsection is convenient to use for debugging the identification protocol, and for its initial implementation. In future works, other methods can be considered and compared, in order to define the most efficient one.

C. OPTIONS OF THE DYNAMIC IDENTIFICATION

Dynamic identification may have two options. The first option is a "confirmatory" dynamic identification (Fig. 11). In this case, every network device sends its dynamic identifier, informing the device-interlocutor that important data was received correctly.

The second option is the identification by request. In this case, the network device requests the identifier from its interlocutor (Fig. 12). The rate of requests depends on the importance of data. It may be sent in a case if there is uncertainty in the validity of the interlocutor, for example, when incorrect data packets appear.

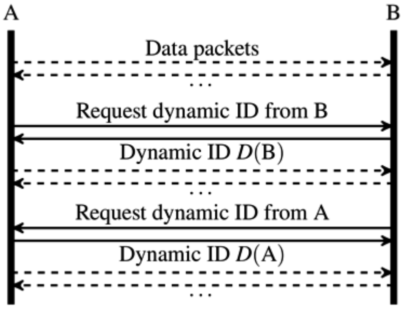


FIGURE 12. Diagram of the dynamic identification by request.

TABLE 2. Average time of an ID search in a database of 10,000 unique records.

	S	$P+S$	IMEI	IMEI $+S$
Identification time t_s, sec	0.371	0.272	0.243	0.265

D. MITIGATING COMPUTATIONAL OVERHEADS

In this subsection, we consider the computational overhead introduced by the proposed identification method. To compare the performance of different identification methods, we measure the mean time required for the identification of an ICT device using each of the methods.

Performance of the proposed method depends on the number of records in a database (number of registered devices) and the size of the identifiers. If the number of records in the database is large and the size of an identifier (the volume of the memory chip allocated for identification) is significant, the computational complexity of a device verification procedure may become considerably high. To reduce the computational complexity, the initial search in a database can be performed using a conventional identifier (e.g., MAC address, IMEI) as a key. After the record is found, the device identity can be verified by a comparison of the bad cell pattern.

For a numerical evaluation of the time required to search for an identifier in a database, we performed an express test using a system of computer algebra GNU/Octave. The test allows us to define the relative difference between conventional (IMEI) and proposed identification methods. During the test we considered four options: (i) search of full identifier S; (ii) search of S using short pointer P (first 256 bytes of S) with subsequent verification of S; (iii) search of IMEI; (iv) search of S using IMEI as a pointer with subsequent verification of S. For the experiment, we used a simplified database containing 10,000 records. For each option, we executed 1,000 search attempts. The meantime t_s required for searching for the identifiers during the test is shown in Table 2.

Table 2 demonstrates that time required for the identification of an ICT device using unique degradation of NOR flash is approximately 53 percent higher than using IMEI. However, in cases of using a pointer, the time difference with the conventional method is only 10-15 percent.

It should be noted that the performed test indicates the relative difference in identification time for considered methods. Absolute values of identification time can be reduced considerably if specialized computational systems and acceleration methods are utilized, while the relative difference between the times required by each method remains the same.

VI. CONCLUSION

As it was demonstrated in this paper, the degraded flash memory sector can be used as a unique identifier of a device. One memory sector has an extremely high number of possible bad-cell combinations. Moreover, the combination of bad-cells is difficult to tamper. Due to these two facts, the reliable identification for quadrillions of devices can be provided by only one memory sector. This number of devices is far beyond the existing demand (existing number of ICT devices).

Considering security issues of the method it should be noted each identification system that is based on transmitting identifiers through a public network is vulnerable to a substitution attack. The traditional solution is to use cryptography protected channels.

For future development of the method, many implementation challenges should be addressed. Firstly, the proposed approach can be tremendously improved by specifically designing chips optimized for use as identifiers. These chips should contain low volume sectors and include an option for fast degradation. The optimized chips will allow a reduction in the time and cost for the degradation process and simplify the identifiers (identifiers are currently too redundant). Secondly, the network protocol has to be developed for remote identification of the ICT devices using degraded flash memory. This protocol should be integrated into the OSI model and serve as a tool for blocking counterfeit or stolen devices. Thirdly, the database architecture for both storing and managing the identifiers must be developed. This architecture should be highly reliable and protected from malicious changes.

Finally, further development of this theory is required. Particularly, theoretical foundations of the method should consider the natural appearance of bad-cells in a memory during the lifespan of a device. Due to this process, these new bad-cells may change the identifier, causing a device to be blocked even if it is not counterfeit or stolen.

REFERENCES

[1] S. Jeschke, C. Brecher, T. Meisen, D. Özdemir, and T. Eschert, ‘‘Industrial Internet of Things and cyber manufacturing systems,’’ in *Industrial Internet of Things*. Cham, Switzerland: Springer, 2017, pp. 3–19.

[2] A. Ramamurthy and P. Jain, ‘‘The Internet of Things in the power sector opportunities in Asia and the Pacific,’’ Asian Develop. Bank, Mandaluyong, Philippines, Tech. Rep. 48, Aug. 2017.

[3] L. Zhang, I. K. Dabipi, and W. L. Brown, Jr., ‘‘Internet of Things applications for agriculture,’’ in *Internet of Things A to Z: Technologies and Applications*. Hoboken, NJ, USA: Wiley, May 2018, pp. 507–528.

[4] J. H. Abawajy and M. M. Hassan, ‘‘Federated Internet of Things and cloud computing pervasive patient health monitoring system,’’ *IEEE Commun. Mag.*, vol. 55, no. 1, pp. 48–53, Jan. 2017.

[5] M. Thibaud, H. Chi, W. Zhou, and S. Piramuthu, ‘‘Internet of Things (IoT) in high-risk environment, health and safety (EHS) industries: A comprehensive review,’’ *Decis. Support Syst.*, vol. 108, pp. 79–95, Apr. 2018.

[6] R. Pirmagomedov, I. Hudoev, R. Kirichek, A. Koucheryavy, and R. Glushakov, ‘‘Analysis of delays in medical applications of nanonetworks,’’ in *Proc. IEEE 8th Int. Congr. Ultra Mod. Telecommun. Control Syst. Workshops (ICUMT)*, Oct. 2016, pp. 49–55.

[7] A. Kumcu, L. Vermeulen, S. A. Elprama, P. Duysburgh, L. Platiša, Y. Van Nieuwenhove, N. Van De Winkel, A. Jacobs, J. Van Looy, and W. Philips, ‘‘Effect of video lag on laparoscopic surgery: Correlation between performance and usability at low latencies,’’ *Int. J. Med. Robot. Comput. Assist. Surg.*, vol. 13, no. 2, p. e1758, 2017.

[8] R. Pirmagomedov, M. Blinnikov, R. Glushakov, A. Muthanna, R. Kirichek, and A. Koucheryavy, ‘‘Dynamic data packaging protocol for real-time medical applications of nanonetworks,’’ in *Internet of Things, Smart Spaces, and Next Generation Networks and Systems*. Cham, Switzerland: Springer, 2017, pp. 196–205.

[9] N. A. Johansson, Y.-P. E. Wang, E. Eriksson, and M. Hessler, ‘‘Radio access for ultra-reliable and low-latency 5G communications,’’ in *Proc. IEEE Int. Conf. Commun. Workshop (ICCW)*, Jun. 2015, pp. 1184–1189.

[10] ‘‘Counterfeit ICT equipment,’’ ITU-T, Geneva, Switzerland, ITU-T Tech. Rep., Dec. 2015.

[11] S. Vladimirov and R. Kirichek, ‘‘The IoT identification procedure based on the degraded flash memory sector,’’ in *Internet of Things, Smart Spaces, and Next Generation Networks and Systems*. Cham, Switzerland: Springer, 2017, pp. 66–74.

[12] Y. Wang, W.-K. Yu, S. Wu, G. Malysa, G. E. Suh, and E. C. Kan, ‘‘Flash memory for ubiquitous hardware security functions: True random number generation and device fingerprints,’’ in *Proc. IEEE Symp. Secur. Privacy*, May 2012, pp. 33–47.

[13] S. Jia, L. Xia, Z. Wang, J. Lin, G. Zhang, and Y. Ji, ‘‘Extracting robust keys from NAND flash physical unclonable functions,’’ in *Proc. Int. Conf. Inf. Secur.* Cham, Switzerland: Springer, 2015, pp. 437–454.

[14] R. Vijayakumar, K. Selvakumar, K. Kulothungan, and A. Kannan, ‘‘Prevention of multiple spoofing attacks with dynamic MAC address allocation for wireless networks,’’ in *Proc. Int. Conf. Commun. Signal Process. (ICCSP)*, Apr. 2014, pp. 1635–1639.

[15] A. Hegde, ‘‘MAC spoofing detection and prevention,’’ *Int. J. Adv. Res. Comput. Commun. Eng.*, vol. 5, no. 1, pp. 229–232, 2016.

[16] S. P. Rao, S. Holtmanns, I. Oliver, and T. Aura, ‘‘Unblocking stolen mobile devices using SS7-MAP vulnerabilities: Exploiting the relationship between IMEI and IMSI for EIR access,’’ in *Proc. IEEE Trustcom/BigDataSE/ISPA*, vol. 1, Aug. 2015, pp. 1171–1176.

[17] I. Gepko, ‘‘General requirements and security architecture for mobile phone anti-cloning measures,’’ in *Proc. IEEE Int. Conf. Comput. Tool (EUROCON)*, Sep. 2015, pp. 1–6.

[18] S. Vaidya and K. J. Christensen, ‘‘A single system image server cluster using duplicated MAC and IP addresses,’’ in *Proc. IEEE 26th Annu. Conf. Local Comput. Netw. (LCN)*, Nov. 2001, pp. 206–214.

[19] A. Cherchali, M. J. Gudelis, Jr., W. G. Lester, and R. J. McLaughlin, ‘‘Technique for automated MAC address cloning,’’ U.S. Patent 9 124 474 B2, Sep. 1, 2015.

[20] J. Wright, ‘‘Detecting wireless LAN MAC address spoofing,’’ Help Net Secur., Kastav, Croatia, White Paper, Jan. 2003.

[21] F. Guo and T.-C. Chiueh, ‘‘Sequence number-based MAC address spoof detection,’’ in *Proc. Int. Workshop Recent Adv. Intrusion Detection*. Berlin, Germany: Springer, 2005, pp. 309–329.

[22] T. Kohno, A. Broido, and K. Claffy, ‘‘Remote physical device fingerprinting,’’ *IEEE Trans. Dependable Secure Comput.*, vol. 2, no. 2, pp. 93–108, Apr./Jun. 2005.

[23] V. Brik, S. Banerjee, M. Gruteser, and S. Oh, ‘‘Wireless device identification with radiometric signatures,’’ in *Proc. 14th ACM Int. Conf. Mobile Comput. Netw.*, 2008, pp. 116–127.

[24] K. I. Talbot, P. R. Duley, and M. H. Hyatt, ‘‘Specific emitter identification and verification,’’ *Technol. Rev.*, vol. 113, pp. 113–133, Jan. 2003.

[25] K. Remley, C. A. Grosvenor, R. T. Johnk, D. R. Novotny, P. D. Hale, M. McKinley, A. Karygiannis, and E. Antonakakis, ‘‘Electromagnetic signatures of WLAN cards and network security,’’ in *Proc. IEEE 5th Int. Symp. Signal Process. Inf. Technol.*, Dec. 2005, pp. 484–488.

[26] G. E. Suh and S. Devadas, ‘‘Physical unclonable functions for device authentication and secret key generation,’’ in *Proc. ACM/IEEE 44th Design Automat. Conf.*, Jun. 2007, pp. 9–14.

[27] M. Jakobsson and K.-A. Johansson, ‘‘Unspoofable device identity using NAND flash memory,’’ *SecurityWeek*, Oct. 2010.

[28] A. Tal, "Two flash technologies compared: NOR vs NAND," M-Syst. Flash Disk Pioneers, Kfar Saba, Israel, White Paper 91-SR-012-04-8L, Oct. 2002.

[29] L. Forbes, "NOR flash memory cell with high storage density," U.S. Patent 6 996 009 B2, Feb. 7, 2006.

[30] K. Itoh, *VLSI Memory Chip Design*, vol. 5. Berlin, Germany: Springer, 2013.

[31] Toshiba, "NAND vs. NOR flash memory technology overview," Toshiba Amer. Electron. Compon., Irvine, CA, USA, Tech. Rep., 2006, pp. 1–4.

[32] D. N. Nguyen, S. M. Guertin, G. M. Swift, and A. H. Johnston, "Radiation effects on advanced flash memories," *IEEE Trans. Nucl. Sci.*, vol. 46, no. 6, pp. 1744–1750, Dec. 1999.

[33] R. Bez, E. Camerlenghi, A. Modelli, and A. Visconti, "Introduction to flash memory," *Proc. IEEE*, vol. 91, no. 4, pp. 489–502, Apr. 2003.

[34] A. DasGupta, "The matching, birthday and the strong birthday problem: A contemporary review," *J. Statist. Planning Inference*, vol. 130, nos. 1–2, pp. 377–389, 2005.

SERGEY S. VLADIMIROV received the Cand.Sc. degree from the St. Petersburg State University of Telecommunications (SPbSUT), in 2013, where he has been an Associate Professor, since 2015. His research interests include error-correcting coding, network identification, network protocols, and the tactile internet.

RUSTAM PIRMAGOMEDOV received the Dipl.Eng. and Cand.Sc. degrees from SPbSUT, in 2010 and 2014, respectively. Since 2016, he has been focused on IoT enablers. He was with an industry, from 2010 to 2018. He is currently with RUDN University, Moscow, Russia, and Tampere University, Tampere, Finland. His research interests include the IoT, body area networks, and mesh networks.

RUSLAN KIRICHEK graduated from SPbSUT, in 2007. He received the Cand.Sc. degree from SPbSUT, in 2012. He was an Associate Professor with SPbSUT, where he has been the Head of the Internet of Things Laboratory, since 2012. His research interests include the IoT, wireless sensor networks, and identification of network devices.

ANDREY KOUCHERYAVY graduated from the Leningrad University of Telecommunication, in 1974. He received the Cand.Sc. and Dr.Sc. degrees, in 1982 and 1994, respectively. He joined the Telecommunication Research Institute (LONIIS), where he was the First Deputy Director, from 1986 to 2003. He has been a Professor with the St. Petersburg State University of Telecommunication (SUT), since 1998. He has been the Head of the Department of Telecommunication Networks and Data Transmission, since 2011. His research interests include the network planning, teletraffic theory, and the IoT and its enablers.

• • •

PUBLICATION

III

UAV-based gateways for wireless nanosensor networks deployed over large areas

R. Pirmagomedov, R. Kirichek, M. Blinnikov and A. Koucheryavy

Computer Communications 146.(2019), 55–62
DOI: 10.1016/j.comcom.2019.07.026

Computer Communications 146 (2019) 55–62

Contents lists available at ScienceDirect

Computer Communications

journal homepage: www.elsevier.com/locate/comcom

UAV-based gateways for wireless nanosensor networks deployed over large areas☆

Rustam Pirmagomedov [a,b,*], Ruslan Kirichek [c], Mikhail Blinnikov [b], Andrey Koucheryavy [c]

[a] *Tampere University, Tampere, 33720, Finland*
[b] *Peoples' Friendship University of Russia, Moscow, 117198, Russian Federation*
[c] *St. Petersburg State University of Telecommunication, St. Petersburg, 193232, Russian Federation*

ARTICLE INFO

Keywords:
IoT
UAV
Wireless sensor networks
Nanonetworks

ABSTRACT

This article combines passive wireless nanosensor networks deployed over a large area and unmanned aerial vehicles (UAVs). The use of UAVs in nano communication network applications can significantly expand their capabilities. Particularly, the highly mobile UAV-based gateways considered in the paper, enable the collection of data from thousands of nanosensors without the utilization of complicated multi-hop routing between nanodevices. The article considers the unique properties of the THz frequency range for the wireless energy transfer to nanodevices as well as for communication with them. More specifically, the energy harvested from electromagnetic waves which are radiated by the UAV-based gateway provide sufficient power for the functioning of the passive nanosensor and signal transmission to the gateway (reader). Such passive nanosensors do not require any maintenance, have a long service life and low cost. Thus, the considered case can serve as the basis for numerous monitoring scenarios, including control of the soil state in agriculture, environmental pollution monitoring, and the control of linear objects (pipelines, dams, dikes). In the considered scenario, the paper discusses technical aspects of the system design, including installation of nanosensors, data frame structure, medium access control, the energy consumption of nanosensors, and aspects of electromagnetic wave propagation. Finally, we evaluate the performance of the proposed system using a system-level simulator.

1. Introduction

Wireless sensor networks are an essential technology enabler for emerging industrial applications (Industry 4.0). Typically, nodes in such networks are equipped with a transceiver, microprocessor, power unit, and a sensor. In cases where these nodes may not have an onboard battery unit, such devices referred to as passive wireless sensors (PWS). The passive wireless sensors typically utilize energy harvesting (e.g., from electromagnetic waves) to accumulate sufficient power for the unit's function and communication (data transmission to a gateway). The notable advantages of PWS, such as low maintenance expenses, long service life, and low initial cost, have made them widely used in many areas including healthcare [1], supply chains [2], manufacturing [3] and smart cities [4].

The recent development of nanotechnologies enabled innovative sensor solutions, such as wireless nanosensor networks. A nanosensor is not necessarily a device whose size is limited to nanometers, but rather a device that utilizes unique properties of nanomaterials for the detection and measurement in the nanoscale [5]. For the time being, common wireless nanosensors have microscopic dimensions (e.g., passive acoustic nanosensor [6,7]) and utilize the unique properties of graphene to transmit data in the THz frequency range [8]. Since a battery is a primary contributor to the size of a device, wireless nanosensors are commonly designed as passive devices with energy harvesting capabilities [5].

Due to energy constraints, passive wireless nanosensors (PWNS) are unable to maintain a constant communication channel with other network elements. In order to communicate, these devices should be provided with energy sufficient to perform the measurement, and for transmission of the measured values to the network gateway. Recent research on wireless nanosensor networks has been mostly restricted to the consideration of these problems, with a focus on body area networks. Specifically, it has been proposed to power passive nanosensors using on-body gateways [9,10]. Conversely, there is little published data regarding other types of passive wireless nanosensor network deployments.

This paper considers data acquisition from wireless nanosensor networks deployed over large areas, e.g., agricultural fields, oil pipelines, or construction sites. To automate the process of gathering data, the wireless gateway (reader) can be integrated into an autonomous mobile platform, such an Unmanned Ground Vehicle (UGV) or Unmanned

☆ The publication has been prepared with the support of the "RUDN University Program 5-100".
* Corresponding author at: Tampere University, Tampere, 33720, Finland.
E-mail address: rustam.pirmagomedov@tuni.fi (R. Pirmagomedov).

https://doi.org/10.1016/j.comcom.2019.07.026

Available online 30 July 2019

Aerial Vehicle (UAV). In comparison with the terrestrial machines, UAVs allow for rapid deployment to the areas of interest and more flexible or dynamic routes (unobstructed paths are not required). Moreover, the direct impact of UAVs on the environment is negligible, while terrestrial UGVs may cause disruptions or unintended harm. For these reasons, in this work, we are focused solely on a UAV-aided solution.

This study aims to contribute to the growing branch of nanonetworks by conceptualizing both the wireless energy transfer and data transmission when utilizing a UAV-based THz-frequency wireless gateway. Particularly we developed an energy model which include the specifics of THz wave propagation and energy costs of sensor nodes. Our metric is the fraction of nodes which failed to send data to the gateway (losses). The reported numerical results demonstrate the feasibility of our proposal and illustrate how losses and redundancy depend on the velocity of the UAV and the frequencies used (we considered the first transparency window of THz range).

The remainder of this paper is organized as follows: Section 2 provides background information about passive wireless sensors and electromagnetic nanosensor networks. Section 3 describes an illustrative scenario of using UAVs with nanosensor networks. Section 4 considers the analytical model for the scenario. Section 5 presents the results of the simulation using a developed model. In the final section, the results of the work are analyzed, and ideas for future work are considered.

2. Related works

2.1. Passive RFID with sensor

Recently, passive Radio-Frequency Identification tags (RFIDs) equipped with a sensor (referred to as a passive wireless sensor) have become a popular solution for several types of industrial settings and functions (e.g., manufacturing, agricultural, healthcare, construction, and retail environments). The acquisition of data from such sensors is performed using a reader. The reader radiates electromagnetic waves used by sensors for energy harvesting. Passive wireless sensors then convert these radio waves into DC power and use it for both the environmental sensing function and communication with the reader. After the sensor has harvested the necessary energy, it starts transmitting data to the reader. In the next step, and depending on the specific requirements of the application, the reader may store the data locally or forward it to a remote server/user (Fig. 1).

Passive sensor devices all include a microchip with integrated energy harvesting function, an antenna, and a sensor. In some cases, the sensing functionality can be performed by changing antenna impedance correlated to the measured parameter (e.g., pressure, temperature, or moisture level).

Most of the existing passive RFIDs operate in a narrow frequency range, which requires a revision of the medium access control (MAC) mechanism [11], especially if the reader interacts with multiple sensor devices simultaneously. A considerable part of the developed RFID MAC protocols relies on time division multiple access, which limits the number of simultaneous communication sessions [12]. A protocol specially adapted to the demands of passive wireless sensors is proposed in [13]. This protocol requires strictly centralized control and utilization of an out-of-band RF power transfer, which considerably limits its implementation. Alternatively, the Token-MAC protocol enables fair access to the medium for RFID systems without synchronization [14]. However, this approach is both energy and computation demanding, since it relies on the local processing operations performed by nodes.

Reviewing medium access technologies for passive RFID devices, one can see that there is no universal solution. Choice of an appropriate MAC protocol is a vital part of application design, one which depends on the parameters of the application (e.g., the number of simultaneous connections, computational capabilities of the microchips, energy models).

2.2. Electromagnetic nanodevices

A new network paradigm, named Internet of Nano-Things (IoNT) was suggested by I. F. Akyildiz et al. [15]. The IoNT is defined as the interoperability of nanodevices (or nanomachines) with conventional networks, including the Internet. The reference architecture of IoNT [15], includes the following elements:

1. Nanonodes — simple nanosized devices with highly limited computational, communicational and energy capabilities. On the one hand, these nanosized devices can be easily embedded in the environment. On the other hand, due to the size, the devices cannot be equipped with powerful batteries or processors, which leads to significant limitations. Depending on the application, the nanonodes can serve as a sensor or actuator, or in some cases, perform both functions simultaneously [15]. To expand the capabilities and applications of a single nanonode, they can be interconnected and execute collaborative tasks in a distributive way. The network of nanonodes (nanonetwork) enables communication among nanonodes utilizing broadcasting and multihop communication methods [16].
2. Nanorouters — nanodevices, providing aggregation of information from nanonodes and interoperability of different nanonetworks. Optionally, nanorouters can be used to orchestrate the nanonodes (e.g., switching sleep mode on and off). It is assumed that nanorouters have more powerful capabilities in comparison with nanonodes. As a consequence, nanorouters have larger physical sizes.
3. Nano/micro gateways — devices responsible for acquiring the data from nanorouters or directly from nanonodes and creating interoperability with traditional networks (e.g., Wi-Fi, LTE, 5G).

IoNT technologies facilitated improvements in various scientific and practical spheres, including biomedical, industrial, and environmental applications. The biomedical applications of IoNT promise the kind of continuous online monitoring of an organism which opens may open a new era of personalized healthcare, one where diagnosis and treatment are performed considering unique properties and circumstances of each patient. The industrial applications of IoNT aim towards total control of quality during every step of manufacturing [16]. In environmental deployments, nanodevices create cost-effective solutions for monitoring large physical areas. This extensive area monitoring can provide early detection of aggressive chemical and biological agents (e.g., technogenic disaster scenario) and coordinate the appropriate response [17].

Wireless communication with electromagnetic nanodevices is enabled by the novel properties of graphene antennas, which are capable of both transmitting and receiving modulated THz-band radiation [15]. The graphene antennas rely on the phenomena of surface plasmon polaritons (SPP). Due to the SPP effect, 1μm-sized antenna is capable of communicating in the THz band [8].

Electromagnetic nanonetworks can be easily combined with conventional telecommunication, as they utilize electromagnetic energy to transfer information. However, due to the limited capabilities of electromagnetic nanodevices, there are plenty of challenges which still need to be addressed. One of these challenges is energy supply. To provide the power for nanodevices, energy from a nano–micro gateway can be used with the same principle that works with passive RFID technology [18]: the modulated THz radiation emitted by the nano/micro gateway creates correlated SPP waves in the graphene antenna; then, injection of SPP into waves in High Electron Mobility Transistor (HEMT) results into a coupled plasma wave, the plasma wave creates an electric current which can be used for both powering of the nanosensor and for communication [19].

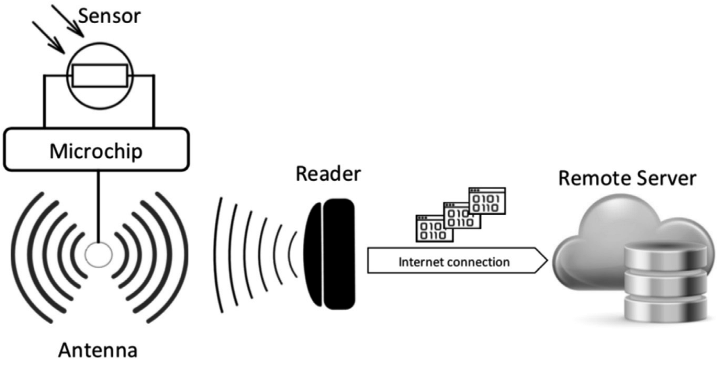


Fig. 1. Acquiring data from a passive RFID with sensor.

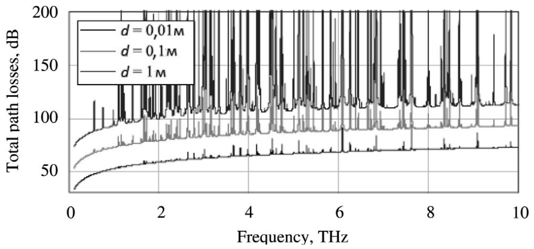


Fig. 2. The total path loss [22].

2.3. Properties of THz-band channel

The THz frequencies have specific properties, which can affect the efficiency of communication. The transmitted signal dissipates as it propagates through space. Therefore, the signal strength received by the antenna will decrease according to the distance between the receiver and the transmitter. This type of attenuation is called free-space propagation loss (FSPL). In addition to the FSPL, there is both considerable molecular absorption and the scattering of particles in the suspended matter (e.g., fog, smoke) for THz frequencies [20].

The molecular absorption in the THz band is mostly caused by vibrations and rotations of molecules of oxygen and water. Molecular absorption occurs at frequencies close to the resonant frequencies of the molecules. These molecules absorb a certain amount of signal energy and create noise (molecular noise) at the same frequencies, due to the internal kinetic energy. The ability of the molecules to absorb the energy of the signal is determined by the absorption coefficient. Thus, the range of stable communication in the THz band is highly dependent on the relative humidity of the environment [21,22].

The scattering may have a significant effect if the concentration of particles in the air is high [23], which is possible in extreme environmental scenarios such as firefighting. Under normal conditions, the impact of scattering on communication channels is negligible.

As it follows from Fig. 2, the distance of the communication in THz-band is strongly dependent on the frequency and limited by the high losses, while also requiring a line of sight (LoS) between transmitter and receiver [22].

3. Illustrative scenario

To illustrate the utilization of a UAV for collecting data from a wireless nanosensor network, we consider field A with sizes 500×500 m. For the monitoring the field, passive nanosensors were deployed, which can measure certain environmental parameters (e.g., temperature, humidity, pH). The monitoring process has a two-step execution: (i) initial installation of nanosensors and (ii) acquiring data from nanosensors.

3.1. Installation of nanonodes

To install the nanosensors on the field, the UAV is used. The UAV installs the nanosensors on the field during the initial flyover (Fig. 3). There are two options for nanonodes installation. The first option is a predetermined installation. This option takes more time but limits the overall number nanosensors (fewer number of sensors are required to cover the field). The second option is to scatter nanosensors from the UAV (randomized installation). The second option requires less time for installation, but uses more sensors to cover the field and may provide a non-uniform data collection pattern. In this subsection, we consider both options.

We assume that monitoring is a space–time discretized process and that each sensor has an effective area of work. To mathematically describe the considered scenario, we used a binary model discussed in [24]. In the binary model, the effective area of work for each sensor is equal and constant. To evaluate the portion of the area covered by each nanosensor, the binary model considers the field is covered by grid $G(x, y)$. In our scenario, we assume that each cell of the grid is 1 m^2. We then assume that the fraction of the field covered by the effective work area of the nanosensors is approximated by the number of covered grid cells.

To evaluate the fraction of the covered grid cells, we introduce a binary function $I(x, y, s_i)$ whose value is 1 if the cell falls into the sensor effective work area $s_i(x_i, y_i)$ and 0, otherwise.

$$I(x, y, s_i) = \begin{cases} 1, & \text{if } \sqrt{(x - x_i) + (y - y_i)} < R_s \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

where R_s is the radius of effective work area of a nanosensor.

In the case of random dispersal of nanosensors, the covered area can be calculated using formula (2).

$$A_{cover}(s) = U_{i=1}^{n} A_{s_i} \cong \frac{A}{500 \cdot 500} \sum_{x=1}^{500} \sum_{y=1}^{500} I(x, y, s) \tag{2}$$

here n is the total number of the nanosensors on the field.

The probability of field coverage by sensor work areas is defined by formula (3).

$$P_{cover} = \frac{A_{cover}(s)}{A} \tag{3}$$

In a special case, when the coordinates of the sensor nodes are random independent numbers distributed by a uniform law, the distribution of sensor nodes can be described by the Poisson model. Thus,

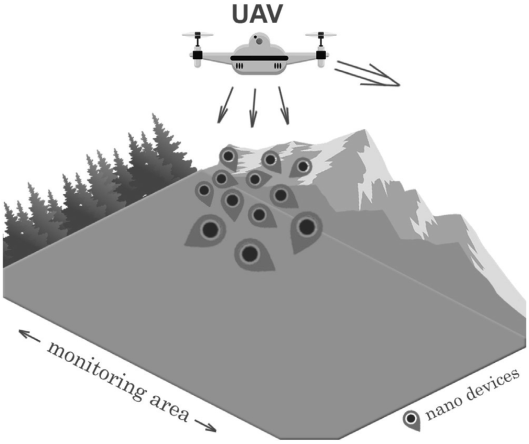


Fig. 3. Installation sensors.

Table 1
Structure of a data frame utilized by the nanosensor.

Pre bits	Start byte	Sensor model	Data	Ending byte
2 bytes	1 byte	2 byte	8 bytes	1 byte

the distance between a randomly chosen point and the nearest sensors is a stochastic value and follows the Rayleigh law. Thus, the probability that nanosensors cover the whole field can be calculated using the formula (4).

$$P(n) = 1 - e^{-\pi R_s^2 \frac{n}{A}} \tag{4}$$

The dependence of the probability $P(n)$ from different R_s in the case of a random installation is shown in Fig. 4.

In the case of a predetermined installation, we consider sensor nodes located as shown on a Fig. 5. Thus, the total number of nanosensors required for full coverage can be calculated using the formula (5). For $R_s = 1$ m, the total number of nanosensors required for complete coverage of the field is 125 000.

$$N_d = \frac{A}{2(R_s)^2} \tag{5}$$

As it follows from the provided analysis, the predetermined installation requires almost three times fewer nanosensors to cover the field.

3.2. Acquiring the data from nanosensors

To obtain data from nanosensors, we suggested using a micro/nano gateway integrated into the UAV. The UAV-based gateway flies at a low altitude over the field with existing sensors by a given route (Fig. 6).

Flying over the field where passive nanosensors are present, the UAV radiates electromagnetic (EM) waves of the THz range. The energy of these waves is harvested by the sensors (by converting the SPP waves into electricity). After accumulating enough energy, the nanosensor measures a specific parameter and sends a data frame to the UAV-based gateway, using the THz frequency range. The structure of the data frame is presented in Table 1. After the UAV receives the data frame from the sensor, the data is further supplemented by the current position of the UAV (geographic coordinates) and timestamp before sending data to the remote server. An example of the structure and data is displayed in Table 2.

In order for the passive nanosensor to harvest enough energy, it should be located in the UAV service area for a certain period of time.

Table 2
Example of data packet sent by the UAV to the remote server.

Data received from a sensor			Data added by a gateway (UAV)		
Type of sensor	Sensor model	Data	Location	Time	Date
Temperature	XFD3112	34.211	59.903176, 30.491099	12:32:03	22.09.2017

The service area determines the propagation length of EM waves with respect to the ground plane. For this article, when developing a model, the boundaries of the UAV service area are taken nominally.

3.3. Medium access control

The described scenario requires the support of multi-connectivity when collecting data from sensor nodes. Thus, to avoid collisions on the MAC level, we consider the utilization of different carrier frequencies. Due to the limitations of nanodevices mentioned in Section 2, they do not support an adaptive configuration of the carrier frequency. Therefore, we accept that the carrier frequencies of the nanosensors are fixed and randomly distributed in the considered frequency range, while the reader operates in a wide range. It should be noted that if two or more sensors with the same carrier frequencies are located in close proximity, it may cause interference and disrupt the reception of data from these nodes. The probability of such an event can be numerically evaluated using the solution of "birthday paradox" (6).

$$\delta = 1 - \frac{K!}{K^d(K-d)!}, \tag{6}$$

where K is the number of different carrier frequencies utilized by the sensor nodes, d is the maximal number of the sensor devices which are simultaneously communicating with the reader (located in the working area of the reader).

Let us assume that $d = 10$ and $K = 1000$, then δ will not exceed 0.045, which means that the considered MAC protocol demonstrates an acceptable level of reliability.

4. System model

4.1. Energy required for the sending a sensor report

In order to determine the total delay time t_{tot} between the entry of the passive nanosensor unit into the flying gateway service area and the receipt of a sensor report, it is necessary to take into account the time t_{ch}, required for energy harvesting by a nanosensor, as well as the time required for creating and sending a packet with a sensor report t_{sg}:

$$t_{tot} = t_{ch} + t_{sg} \tag{7}$$

The time t_{sg} is an integral index that directly depends on aspects such as the size of the packet being transferred, data transmission technology, time for processing and encapsulation; it was taken $t_{sg} =$ 10 ms. The time spent for harvesting the energy by a nanosensor is characterized by the formula (8):

$$t_{ch} = \frac{E_{tot}}{E_{rx} r_c} \tag{8}$$

where E_{tot} — total energy demands required for transmission of the one sensor report; E_{rx} — energy, being received by passive nanosensor per time unit; r_c — conversion coefficient of electromagnetic energy into electric energy. In this work we assumed $r_c = 0.5$.

In order to calculate electric energy expended by the passive nanosensor unit for the transmission of one sensor report to the flying gateway, it is necessary to take into account energy expended on maintaining the nanosensor in working order (until it measures the

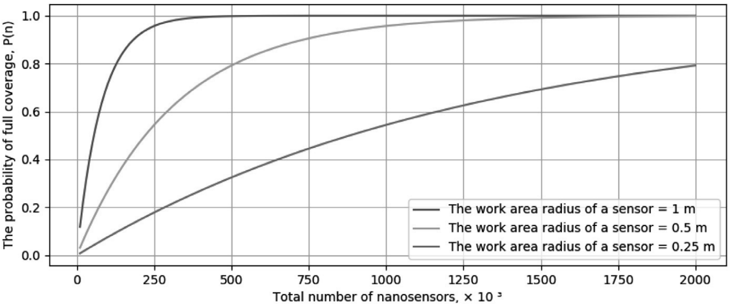


Fig. 4. The probability that a field fully covered by the nanosensors.

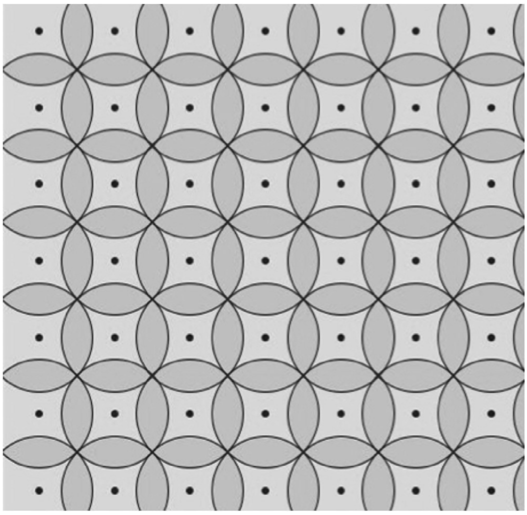


Fig. 5. A scheme for the determined installation of sensors.

required parameter), for processing the information received, and for encapsulating and sending the packet to the UAV [25]:

$$E_{tot} = E_S + E_p + E_{packet-tx} \tag{9}$$

where E_S — energy demands of nanosensor device for measuring the indicator value; E_p — energy demands for data processing; $E_{packet-tx}$ — energy demands for sending the data packet to the gateway.

To specify values of energy consumption of nanosensors, in this paper we used data on energy consumption of Ultra-Low-Power Smart Visual Sensor [26] as a reference. Thus, $E_S = 1.06$ μJ and $E_p = 0.73$ μJ.

To calculate the energy expended for sending one data packet, in [27], it is proposed to use an equation (10).

$$E_{packet-tx} = N_{bits} W E_{pulse-tx} \tag{10}$$

where N_{bits} — the number of bits contained in the transmitted packet; W — the code weight, i.e. the probability of transmitting the pulse ("1") instead of keeping the mute mode ("0"), $E_{pulse-tx}$ — the energy expended for transmission of one pulse. In the considered scenario we use the coding scheme proposed in [28], where the number of "1" and "0" bits in the packet is approximately the same, and accordingly, $W = 0.5$. The mean energy of one pulse required for transmission to a distance of 10 mm is $E_{pulse-tx} = 1$ pJ [29].

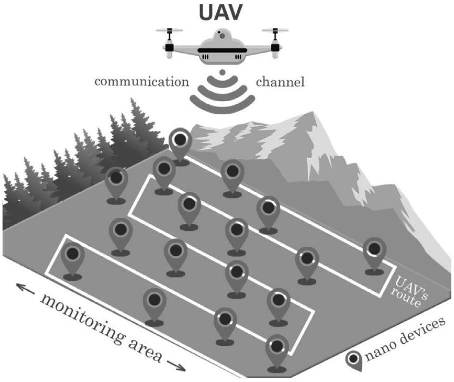


Fig. 6. Collecting data from sensors.

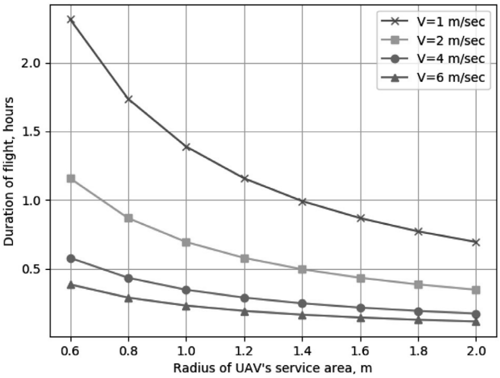


Fig. 7. Time of UAV operation required for serving one hectare.

4.2. Propagation model

The signal power on the receiver side can be expressed as (11).

$$P_{rx} = \frac{P_{tx}G(f)}{A(f)} + N_T(f) + N_{mol}(f) \tag{11}$$

where P_{tx} — power of the transmitted signal; $G(f)$ — antenna gain parameter; $A(f)$ — total attenuation ratio; $N_T(f)$ — thermal noise; $N_{mol}(f)$ — molecular-based absorption noise.

Thermal noise of graphene antennas is suggested to be negligible due to the properties of this material [21].

For the calculation of the attenuation ratio, it is necessary to take into account not only the attenuation of the signal during propagation in the space A_{fspl}, but also the molecular absorption A_{mol} [22,30,31].

Free-space path loss can be defined as follows:

$$A_{fspl}(f) = (\frac{4\pi f d}{c})^2 \tag{12}$$

where d — distance from transmitter to receiver; f — transmission frequency; c — light velocity.

A feature of using the THz range for wireless communication is the presence of molecular absorption caused by vibrations and rotation of molecules. The molecular absorption of the EM energy is an effect that occurs when the signal is transmitted at frequencies close to the resonance frequencies of molecules, which absorb part of the signal energy and produce noise $N_{mol}(f)$ at the same frequencies due to the internal kinetic energy of the molecules [16].

$$A_{mol} f = e^{k(f)d} \tag{13}$$

where k — the average absorption coefficient.

The molecular absorption coefficient determines the ability of a molecule to absorb energy [21]. According to [22], the losses of the molecular absorption are calculated depending on the transmission frequency, distance between the receiving and transmitting antenna, but also on environment conditions and composition where the signal propagates. In this article the molecular absorption coefficient was determined by means of the HITRAN database [32,33] for ambient conditions corresponding to the "USA model, mean latitude, summer", $H = 0$ (mixture correlations: $H_2O = 1.860000\%$, $CO_2 = 0.033000\%$, $O_3 = 0.000003\%$, $N_2O = 0.000032\%$, $CO = 0.000015\%$, $CH_4 = 0.000170\%$, $O_2 = 20.900001\%$, $N_2 = 77.206000\%$) at a temperature of 296 K and a pressure of 1 atm.

As it was demonstrated in [21,22] at some frequency ranges, the absorption is notably larger, which can limit the communication range. To avoid negative effects of absorption, we utilize "transparency windows" – parts of the spectrum with the low absorption. It is worth noting that k does not depend on the transmission distance, but only on the ambient conditions and the frequency of the transmitted signal.

In accordance with [21] when the transmission coefficient values are lower than 94.5% (equivalent to the absorption coefficient values of the environment higher than 5.5%), the molecular absorption noise N_{mol} becomes equal to the maximum value of −203.89 dB/Hz ($\approx 10^{-20}$ W/Hz), which corresponds to the Johnson–Nyquist thermal noise level.

Since there is no need to use high throughput for the transmission of sensor reports, we consider a bandwidth of 100 kHz per one nanodevice which is close to existing UHF RFID-based sensors. This article considers the frequency range of 0.1–0.15 THz. Taking into account that the environment transmission coefficient in the transparency windows is always above 95.5% at small and medium distances, and the molecular noise per 1 Hz of the used frequency band is about $\approx 10^{-20}$ W, we obtain $N_{mol} = 1$ fW, which allows for considering the molecular noise value as negligibly small. Thus, (11) will result in (14).

$$P_{rx} = \frac{T_{tx} G(f) c^2}{(4\pi f d)^2 e^{k(f)d}} \tag{14}$$

The main parameters used in the model, are summarized in Table 3. More details regarding the 3D channel characteristics used in this model can be found in [34].

4.3. Coverage of the UAV

When collecting the data from nanosensors the UAV flying linearly with a constant speed v over the subject area. In our model, we considered the flight of the UAV an altitude of 2 m to reduce the impact of the stability issues in our results. It worth noting that the stability of flight will directly depend on the type of the UAV (e.g., multi-rotor, fixed-wing, aerostat-based) and environmental conditions (e.g., strong wind, precipitation). For some types of UAVs, altitudes lower than 2 m may also be considered if advanced flight control methods are utilized [35].

Table 3
Parameters used in the model.

Parameter	Identification	Value
THz-reader capacity	$P_{tx} \cdot G$	1 W
Antenna of the THz-reader		Directional [37,38]
Antenna of the nanosensor		Isotropic
Frequency range	$f_1 - f_2$	0.1 – 0.15 THz
Bandwidth per sensor	$\triangle f$	100 kHz
Altitude of UAV	h	2 m
Service area radius	R	[0.8, 1.2, 1.6, 2] m
Velocity	V	[1, 2, 4, 6, 8, 10, 12] m/s
Absorption factor (0.1 THz)	k_1	2.58×10^{-5} m^{-1}
Absorption factor (0.15 THz)	k_2	1.01×10^{-4} m^{-1}
Mean energy required for transmission of one data packet from sensor to UAV	$E_{tot,min}$	2,27 μJ
Time required for creating and sending a packet with sensor report	t_{sg}	0.01 s

We assume that UAV is equipped with a linear antenna array which beamwidth defines the ground service area. In our model, the coverage area is taken as a circle with radius R. Increase of R allows serving the area faster, as is shown in Fig. 7. However, it may cause additional losses since antenna gain is reduced. To demonstrate the effect of beamwidth to system performance, in the next section, we consider different values of R, with a constant level of transmitter power.

The number of packets received by the gateway directly depends on the time of the passive nanosensor's presence in the coverage of the UAV, t_{ex}(15):

$$t_{ex} = \frac{D}{v} \tag{15}$$

where D ($D \leq 2R$) — communication range.

If the condition $t_{ex} \geq (t_{ch} + t_{sg})$ is not fulfilled, the packet loss will occur. However, if $t_{ex} \geq n(t_{ch} + t_{sg})$, then one nanodevice transmits $'n'$ copies of packets, as the energy harvesting and report transmission cycle succeeds several times.

The distance D can be expressed utilizing the circle line picking method [36]:

$$D = 2R(\frac{\pi}{2} F(S)) \tag{16}$$

$$F(S) = \frac{2}{\pi} \arcsin(\frac{S}{2}) + C \tag{17}$$

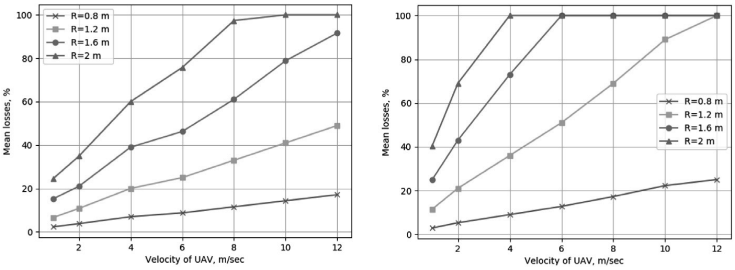
where $F(S)$ — unit circle probability density function; S — the distance between two points on the circle of the UAV service area; C — constant.

5. Simulation results

A system-level performance assessment has been conducted by utilizing WinterSIM simulator. The simulation results indicate the dependency of losses (failures of acquiring the measurement from sensors) on the UAV velocity for two frequencies (Fig. 8). During the simulation, we assumed that the UAV flight path to not overlap with the already serviced territory.

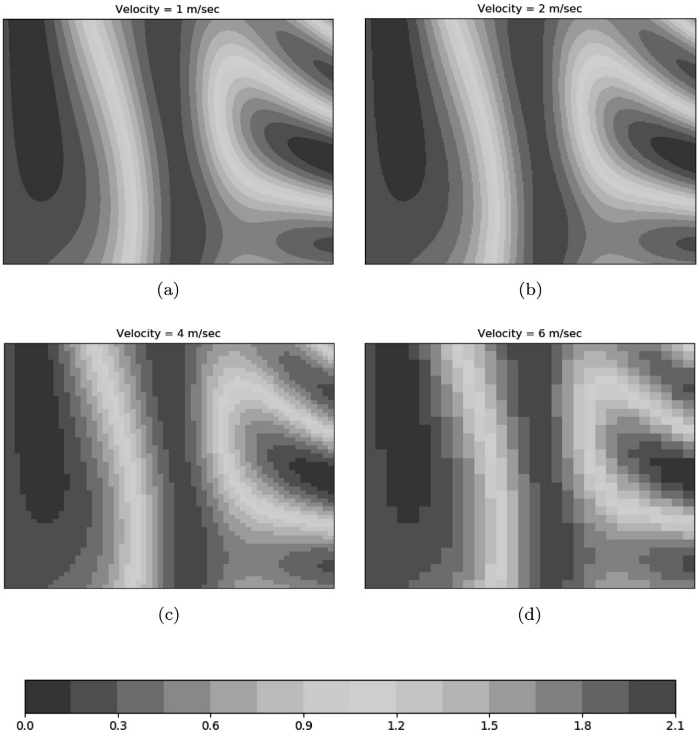
f

The results indicate that losses can be reduced utilizing low flight speeds and narrow beamwidth (short R). The increase in velocity increases losses because sensors located on the edges of the service area (edges which are orthogonal to the UAV motion vector) do not have enough time to accumulate sufficient energy needed for operation. The increase of R leads to reduced antenna gain (transmitter power is constant) and consequently a reduction of energy transmitted to sensors, which also produces increased losses. In addition, the diagrams



GHz.png GHz.png (a) (b)

Fig. 8. Dependence of packet losses on the UAV velocity (a) $f = 0.1$ THz, (b) $f = 0.15$ THz.



(a) (b) (c) (d)

Fig. 9. Maps of measurements obtained with different UAV velocity ($f = 0.1$ THz, $R = 2$ m).

clearly show that the utilization of higher frequency has dramatically higher losses.

To illustrate how the losses affect the performance of the field monitoring, we presented an example of measurement maps compiled using our simulation results (Fig. 9). The maps contain a visualization of the measurements of a virtual parameter performed by a nanosensor network facilitated with the UAV-based gateway. These show how the density of measurement decreases due to higher losses.

Sensors located closer to of UAV route are capable of harvesting energy sufficient for several measurements, which causes redundancy of packets from these sensors. The dependence of the redundancy on UAV velocity for two frequencies is shown in Fig. 10. To avoid creating high traffic in the "gateway–user" channel, the UAV should perform initial processing of raw data in order to detect and average the measured values before forwarding those through mobile network [39].

6. Conclusion

Integration of wireless sensor networks and UAVs promise a cost-effective and simplified data acquisition process. Implementation of nanotechnologies will further expand the scope of applying such networks to industries and situations where the size of the sensor device is the foremost consideration.

This paper demonstrated the feasibility of a UAV-based gateway for acquiring data from passive wireless nanosensors in the THz range. We considered the main aspects of the scenario, including the installation of nanosensors, MAC protocol, THz wave propagation model, and energy consumption of the nanosensors. Within the considered scenario,

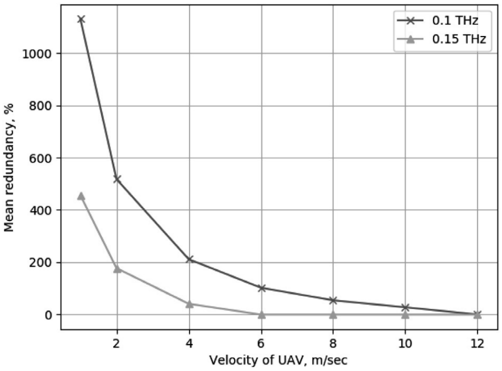


Fig. 10. Dependence of redundancy on UAV velocity.

the THz frequency range was used for both energy harvesting and data transmission. Our numerical results indicate how the losses and redundancy depend on the velocity of the UAV and the frequencies which are utilized.

The notable limitation of this study is that it does not take into account the possible influence of the weather and the presence of additional obstacles between UAV and sensors; these factors should be considered in future works. Moreover, for future works, we expect the developed model could be enhanced further with the concrete characteristics of UAV (mobility model, battery capacity, and path planning), and graphene-based antennas for nanosensors.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] J. Zhang, G. Tian, A. Marindra, A. Sunny, A. Zhao, A review of passive RFID tag antenna-based sensors and systems for structural health monitoring applications, Sensors 17 (2) (2017) 265.

[2] F. Bibi, C. Guillaume, N. Gontard, B. Sorli, A review: RFID technology having sensing aptitudes for food industry and their contribution to tracking and monitoring of food products, Trends Food Sci. Technol. 62 (2017) 91–103.

[3] D.M.S. Velandia, N. Kaur, W.G. Whittow, P.P. Conway, A.A. West, Towards industrial internet of things: Crankshaft monitoring, traceability and tracking using RFID, Robot. Comput.-Integr. Manuf. 41 (2016) 66–77.

[4] H. Mora-Mora, V. Gilart-Iglesias, D. Gil, A. Sirvent-Llamas, A computational architecture based on RFID sensors for traceability in smart cities, Sensors 15 (6) (2015) 13591–13626.

[5] I.F. Akyildiz, J.M. Jornet, Electromagnetic wireless nanosensor networks, Nano Commun. Netw. 1 (1) (2010) 3–19.

[6] D. Aznakayeva, I. Yakovenko, E. Aznakayev, Passive acoustic graphene nanosensor modeling, in: Radar Methods and Systems Workshop (RMSW), IEEE, IEEE, 2016, pp. 91–94.

[7] D. Aznakayeva, I. Yakovenko, E. Aznakayev, Numerical calculation of passive acoustic graphene nanosensor parameters, in: Radar Methods and Systems Workshop (RMSW), IEEE, IEEE, 2016, pp. 95–98.

[8] I.F. Akyildiz, J.M. Jornet, Graphene-Based Plasmonic Nano-Antenna for Terahertz Band Communication, Google Patents, 2017, US Patent 9, 643, 841.

[9] F. Dressler, S. Fischer, Connecting in-body nano communication with body area networks: Challenges and opportunities of the Internet of Nano Things, Nano Commun. Netw. 6 (2) (2015) 29–38.

[10] R. Kirichek, R. Pirmagomedov, R. Glushakov, A. Koucheryavy, Live substance in cyberspace—Biodriver system, in: 2016 18th International Conference on Advanced Communication Technology (ICACT), IEEE, 2016, pp. 274–278.

[11] M.Y. Naderi, P. Nintanavongsa, K.R. Chowdhury, RF-MAC: A medium access control protocol for re-chargeable sensor networks powered by wireless energy harvesting, IEEE Trans. Wireless Commun. 13 (7) (2014) 3926–3937.

[12] F. Iannello, O. Simeone, U. Spagnolini, Medium access control protocols for wireless sensor networks with energy harvesting, IEEE Trans. Commun. 60 (5) (2012) 1381–1389.

[13] J. Kim, J.-W. Lee, Energy adaptive MAC protocol for wireless sensor networks with RF energy transfer, in: 2011 Third International Conference on Ubiquitous and Future Networks (ICUFN), IEEE, 2011, pp. 89–94.

[14] L. Chen, I. Demirkol, W. Heinzelman, Token-MAC: A fair MAC protocol for passive RFID systems, IEEE Trans. Mob. Comput. 13 (6) (2014) 1352–1365.

[15] I.F. Akyildiz, J.M. Jornet, The internet of nano-things, IEEE Wirel. Commun. 17 (6) (2010).

[16] I.F. Akyildiz, F. Brunetti, C. Blázquez, Nanonetworks: A new communication paradigm, Comput. Netw. 52 (12) (2008) 2260–2279.

[17] H. Dai, Carbon nanotubes: Synthesis, structure, properties, and applications, in: Topics in Applied Physics, Vol. 80, Springer, 2001.

[18] E. Perret, M. Hamdi, A. Vena, F. Garet, M. Bernier, L. Duvillaret, S. Tedjini, RF And THz identification using a new generation of chipless RFID tags., Radioengineering 20 (2) (2011).

[19] J.M. Jornet, I.F. Akyildiz, Graphene-based plasmonic nano-transceiver for terahertz band communication, in: Antennas and Propagation (EuCAP), 2014 8th European Conference on, IEEE, 2014, pp. 492–496.

[20] V. Petrov, A. Pyattaev, D. Moltchanov, Y. Koucheryavy, Terahertz band communications: Applications, research challenges, and standardization activities, in: 2016 8th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), IEEE, 2016, pp. 183–190.

[21] J.M. Jornet, I.F. Akyildiz, Channel modeling and capacity analysis for electromagnetic wireless nanonetworks in the terahertz band, IEEE Trans. Wireless Commun. 10 (10) (2011) 3211–3221.

[22] P. Boronin, V. Petrov, D. Moltchanov, Y. Koucheryavy, J.M. Jornet, Capacity and throughput analysis of nanoscale machine communication through transparency windows in the terahertz band, Nano Commun. Netw. 5 (3) (2014) 72–82.

[23] M.J. Fitch, R. Osiander, Terahertz waves for communications and sensing, Johns Hopkins APL Tech. Dig. 25 (4) (2004) 348–355.

[24] M. Abo-Zahhad, S.M. Ahmed, N. Sabor, S. Sasaki, Coverage maximization in mobile wireless sensor networks utilizing immune node deployment algorithm, in: Electrical and Computer Engineering (CCECE), 2014 IEEE 27th Canadian Conference on, IEEE, 2014, pp. 1–6.

[25] M. Blinnikov, R. Pirmagomedov, Wireless identifying system based on nanotags, in: Advanced Communication Technology (ICACT), 2018 20th International Conference on, IEEE, 2018, pp. 753–757.

[26] M. Rusci, D. Rossi, M. Lecca, M. Gottardi, E. Farella, L. Benini, An event-driven ultra-low-power smart visual sensor, IEEE Sens. J. 16 (13) (2016) 5344–5353.

[27] J.M. Jornet, I.F. Akyildiz, Joint energy harvesting and communication analysis for perpetual wireless nanosensor networks in the terahertz band, IEEE Trans. Nanotechnology 11 (3) (2012) 570.

[28] J.M. Jornet, I.F. Akyildiz, Low-weight channel coding for interference mitigation in electromagnetic nanonetworks in the terahertz band, in: Communications (ICC), 2011 IEEE International Conference on, IEEE, 2011, pp. 1–6.

[29] J.M. Jornet, I.F. Akyildiz, Information capacity of pulse-based wireless nanosensor networks, in: Sensor, Mesh and Ad Hoc Communications and Networks (SECON), 2011 8th Annual IEEE Communications Society Conference on, IEEE, 2011, pp. 80–88.

[30] J. Kokkoniemi, J. Lehtomäki, K. Umebayashi, M. Juntti, Frequency and time domain channel models for nanonetworks in terahertz band, IEEE Trans. Antennas and Propagation 63 (2) (2015) 678–691.

[31] K. Tsujimura, K. Umebayashi, J. Kokkoniemi, J. Lethomäki, A study on channel model for THz band, in: Antennas and Propagation (ISAP), 2016 International Symposium on, IEEE, 2016, pp. 872–873.

[32] L.S. Rothman, C. Rinsland, A. Goldman, S. Massie, D. Edwards, J. Flaud, A. Perrin, C. Camy-Peyret, V. Dana, J.-Y. Mandin, et al., The HITRAN molecular spectroscopic database and HAWKS (HITRAN Atmospheric Workstation): 1996 edition, J. Quant. Spectrosc. Radiat. Transfer 60 (5) (1998) 665–710.

[33] [link] URL http://hitran.iao.ru/.

[34] C. Zhang, C. Han, I.F. Akyildiz, Three dimensional end-to-end modeling and directivity analysis for graphene-based antennas in the terahertz band, in: 2015 IEEE Global Communications Conference (GLOBECOM), IEEE, 2015, pp. 1–6.

[35] Y. Chen, G. Zhang, Y. Zhuang, H. Hu, Autonomous flight control for multi-rotor UAVs flying at low altitude, IEEE Access 7 (2019) 42614–42625.

[36] E.W. Weisstein, Circle line picking, From MathWorld–A Wolfram Web Resource.

[37] S.A. Hoseini, M. Ding, M. Hassan, Massive MIMO performance comparison of beamforming and multiplexing in the Terahertz band, in: 2017 IEEE Globecom Workshops (GC Wkshps), IEEE, 2017, pp. 1–6.

[38] Z. Xu, X. Dong, J. Bornemann, Design of a reconfigurable MIMO system for THz communications based on graphene antennas, IEEE Trans. Terahertz Sci. Technol. 4 (5) (2014) 609–617.

[39] R. Pirmagomedov, M. Blinnikov, R. Glushakov, A. Muthanna, R. Kirichek, A. Koucheryavy, Dynamic data packaging protocol for real-time medical applications of nanonetworks, in: Internet of Things, Smart Spaces, and Next Generation Networks and Systems, Springer, 2017, pp. 196–205.

PUBLICATION

IV

Unmanned aerial system–assisted wilderness search and rescue mission
T. D. Dinh, R. Pirmagomedov, V. D. Pham, A. A. Ahmed, R. Kirichek,
R. Glushakov and A. Vladyko

International Journal of Distributed Sensor Networks 15.6 (2019), 1550147719850719
DOI: 10.1177/1550147719850719

Research Article

International Journal of
Distributed
Sensor Networks

International Journal of Distributed Sensor Networks
2019, Vol. 15(5)
© The Author(s) 2019
DOI: 10.1177/1550147719850719
journals.sagepub.com/home/dsn
SAGE

Unmanned aerial system–assisted wilderness search and rescue mission

Truong Duy Dinh[1], Rustam Pirmagomedov[2,3], Van Dai Pham[1], Aram A Ahmed[1], Ruslan Kirichek[1], Ruslan Glushakov[4] and Andrei Vladyko[1]

Abstract

The success of the wilderness search and rescue missions is highly dependent on the time required to search for the lost person. The use of unmanned aerial systems may enhance search and rescue missions by supplying aerial support of the search process. There are unmanned aerial system–based solutions, which are capable of detecting the lost person using computer vision, infrared sensors, and detection of a mobile phone signal. The most pressing issue is reducing the cost of a search and rescue mission. Thus, to improve the efficiency of the resource utilization in wilderness search scenario, we consider the use of unmanned aerial system for both mobile phone detection and enabling Wi-Fi communication for the ground portion of the search and rescue team. Such an approach does not require specific additional tools (e.g. access point, specific user equipment) for communication, which reduces the cost and improves the scalability and coordination of the search and rescue mission. As a result, the article provides methods of searching the wilderness for a person using beacon signals from a mobile phone for two situations: when the distance to the source of emergency signals is unknown and when the distance is known. In addition, the voice transmission delay and the number of unmanned aircrafts are found to guaranty the quality of a call.

Keywords

Unmanned aerial systems, UAS, flying network, unmanned aircrafts, search and rescue, SOS signals

Date received: 1 December 2018; accepted: 27 March 2019

Handling Editor: Daniel Gutierrez-Reina

Introduction

Every year, a considerable number of people around the world get lost in the wilderness. Some of the main reasons for this lie in the inability of modern people to navigate the terrain and the overvaluation of their capabilities, including physical.

The success of a wilderness search and rescue mission is highly dependent on the time required for the search of the lost person. For algorithmization of the search and rescue operations, there were proposed POA (Probability of Area)—POD (Probability of Detection) model, which is an integral part of IAMSAR (International Aeronautical and Maritime Search And Rescue manual).[1] This model allows us to

[1]The Bonch-Bruevich Saint-Petersburg State University of Telecommunications, Saint Petersburg, Russia
[2]Tampere University, Tampere, Finland
[3]Peoples' Friendship University of Russia (RUDN University), Moscow, Russia
[4]Saint-Petersburg State Pediatric Medical University, Saint Petersburg, Russia

Corresponding author:
Ruslan Kirichek, The Bonch-Bruevich Saint-Petersburg State University of Telecommunications, 22/1 Prospekt Bolshevikov, Saint Petersburg 193232, Russia.
Email: kirichek@sut.ru

build an optimal search process based on the available data about the lost person, the area of search, physical conditions, distinguishing features.

The efficiency of the search and rescue mission can be enhanced by the use of the automated detection technologies enabled by unmanned aerial system (UAS). A flying network, which uses UAS-based technologies, allows a reduction in the time required for detection as well as reducing the number of people required for the search mission.[2,3]

Recent works on the use of unmanned aircrafts (UAs) in search and rescue missions utilize mobile phone detection technologies to locate the individual. Among the wireless technologies used for detection of mobile phones, there are 2/3/4G cellular networks, Wi-Fi and Bluetooth. Mobile phone detection technologies can be used jointly with a computer vision system, which improved their efficiency when there is a line of sight between the UA and the lost person. In this network, each UA can be considered as a heterogeneous mobile gateway.

The resource allocation for the search using UAS is a relevant research topic. The cost of a rescue mission is still high, even when automated UAS-based technologies are used. Thus, to improve the efficiency of resource utilization in wilderness search scenario, we consider the use of the UASs for both mobile phone detection and enabling Wi-Fi communication for the ground segment of the search and rescue team. Such an approach does not require specific additional means (access point, specific user equipment) for communication, which reduces the cost and improves the scalability and coordination of the search and rescue mission.

The main contribution of this article is the following:

1. We introduce the methods of wilderness search of a person using beacon signals from a mobile phone for two situations: when the distance to the source of emergency signals is unknown and when the distance is known.
2. We consider the simultaneous use of UASs involved in the search mission to support voice communication among the search and rescue team on the ground using Wi-Fi technology.

The rest of this article is organized by following: In the "Related works," we take an overview the most related work and describe that there are many ways to use a UA during a wilderness search and rescue mission, which helps the task to be completed effectively. In the "Methods of searching the wildness for a person using beacon signals from a mobile" section, we consider two scenarios methods of searching the wildness for an individual using beacon signals from a mobile phone: Methods of searching for a person based on known distance to the source of emergency signals and the method of detecting the coordinates of a mobile phone when the distance to the source of emergency signals is unknown. Voice transmission by a flying network will be shown in the following section. We conclude the article in the "Conclusion.".

Related works

Recent developments in UAS technologies have low-cost drones with considerable capabilities. Currently, the drone market is proliferating, because the UAS's capabilities are relevant in a wide range of applications. It is difficult to cover all the applications; thus, in the remainder of this section, we will only consider rescue scenarios in order to provide additional rationale for our work.

When a disaster happens in a region, the primary aim is to coordinate disaster management operations. Such coordination may become a challenging task if the communication infrastructure is unavailable (e.g. damaged by an earthquake). The UAs may create a temporary infrastructure, acting as access points to provide wireless communication.[4]

The use of UAs in Public Protection and Disaster Relief (PPDR) scenarios is not limited by launching a temporary communication infrastructure. The UAs equipped with sensors can be used to monitor the environmental pollution (e.g. composition of gases, radiations) in specific areas of damage, such as a chemical storehouse or a nuclear power plant. The information about the pollution may assist rescue teams to avoid life-treating areas.[5]

Deploying a UA network autonomously and providing communication services in a disaster scenario also were considered in Sánchez-García et al.[6] In this article, an algorithm called Distributed and dynamic Particle Swarm Optimization for UA networks (dPSO-U) was presented with two main goals: exploring a disaster scenario area, and making the UAs converge to several victim groups discovered during the exploration phase. To evaluate the algorithm, the authors compared it with other algorithms through simulations and received positive results.

The UAs equipped with video cameras may help to evaluate the consequences of the disaster and provide online surveillance for better coordination of rescue teams. Real-time onboard video processing may help to identify the most impacted areas and to assess whether any individuals need help.[5,7] The UAs can also assist the rescue teams to reach a particular location if standard routes are blocked.

Along with sensing or communications, the UAs can be used for the rapid delivery of necessary items to rescue teams (e.g. food, medicine, medical equipment) in the areas that are not accessible by roads.[8,9]

Another scenario to consider regarding UAS utilization is the UAS-assisted search and rescue of a lost person (or group of people) in the wildness. The distinguishing feature of this scenario is the enormous size of the search area and the relatively few targets (lost people). As it was mentioned in Hanna et al.,[10] survivability of the lost person is higher when it is located within 24 h. Thus, the goal of the rescue team is to check the most extensive area as soon as possible. The UAs can enhance the efficiency of that search by utilizing computer vision to detect the target from the air. To enhance the efficiency of the victim detection, in Rasmussen et al.[11] it was proposed to augment visible-spectrum searching with infrared sensing.

The UAs equipped with video cameras have been used to support rescue operations utilizing image recognition for detecting the lost person. However, there are plenty of use cases when victims are buried (e.g. a snow avalanche) or covered by the tree canopy, making the computer vision methods almost useless.

To detect lost people without video methods, wireless signal detection can be utilized. Many specialized wireless beacons allow rescuers to find lost or trouble in people. These beacons are widely used by people whose activities are closely related to risks, such as hikers and climbers. Such categories of people are prepared for the circumstances when they need to be rescued. However, ordinary people can also experience emergency situations. These people are usually not prepared for such a situation. Thus, the techniques that utilizing popular personal wireless devices (e.g. smartphones, smartwatches) as beacons are highly relevant.[12]

The idea of using a cell phone to locate the user is not new. In 1994, the US Federal Communications Commission compelled mobile operators to provide the location of mobile phones that dialed 911.[13] In Mendelson,[14] a method was proposed that employed features of a user's cellular phone to emit an emergency beacon to aid first respond search and rescue units or emergency personnel. This method may noticeably extend the detection of the mobile phone. As it was shown in Wolfe et al.,[15] the mobile phone signal is powerful enough to be detected even under a snow depth of 7 feet (the snow avalanche scenario). The existing solutions consider switching the mobile phone into a beacon mode.[16] In this mode, the mobile phone transmits a beacon signal, which can be more powerful than a regular one. The beacon signal may be used to locate the user.

The UA equipped with a mobile base station or software defined radio may significantly reduce the time for mobile phone detection if there is no coverage in the area of search. Optionally, for the detection of the mobile phone using UA, Wi-Fi or Bluetooth can be utilized.[17] Moreover, if the media access control (MAC) addresses of wireless interfaces are known, the user detection can be performed more efficiently.[18]

The real use case described in Goodrich et al.[19] has shown that resource allocation is a crucial challenge when using UA to support a wilderness search. To address the resource optimization challenge, the methods of using UA for wireless signal detection must take into account a broad range of parameters including the number of UA used, the UA's battery capacity, coverage probability, target mobility, and the area of search size.

The coverage probability problem was partially addressed in Guo et al.[20] by providing a guideline for the design of a wireless network. The issues of target mobility, optimal UA route, and battery use are considered in Mohibullah and Simon.[21]

Through related works, it can be seen that with the use of UAs the efficiency of the search and rescue mission is enhanced. Locating the victim in these cases is one of the critical problems. Based on the above studies, by using UAS, we consider the ability to locate and provide communications for lost people outside the cellular networks.

Methods of searching the wildness for a person using beacon signals from a mobile phone

In this section, we consider methods for searching a lost person based on the detection of SOS signals from the mobile phone. UAs, which fly around the area of the lost person, are used to scan the signals.

Scenario description

One of the options for signaling a request for help may be an application in a mobile phone that generates SOS signals and sends them over cellular and Wi-Fi channels. Suppose that an application has been installed on the mobile phone of the missing person that will help create emergency signals (beacons) with an intensity of once per minute. The proposed solution is, in fact, an analog of the emergency warning system, which is used in the automotive industry to signal an emergence.

It is evident that the signal from the mobile phone must be transmitted to the base station of the telecom operator; however, due to the remoteness of the forest, the power of the mobile phone transceiver may not be enough to receive the signal.

The solution to this problem is to send one or more UA, which surveys the territory in which a person is supposedly lost and the scanning of emergency signals, as Figure 1. On board, the UA is a portable base station, implemented by software-configured radio.

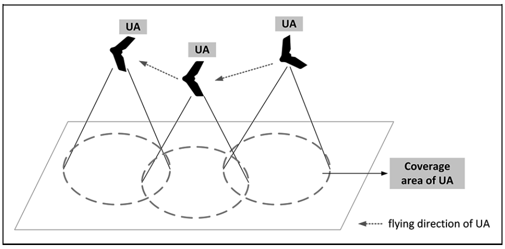


Figure 1. UA overflights areas to scan emergency signals from mobile phones. UA: unmanned aircraft.

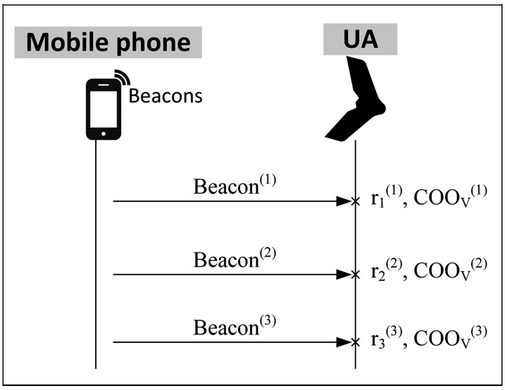


Figure 2. The interaction process between a UA and a mobile phone. UA: unmanned aircraft.

Because the onboard UA base station receives data in a certain radius, depending on the antenna aperture, the number of UA used affects the speed of searching for the coordinates of the missing person.

In the case of detecting emergency signals, UA coordinates are fixed regarding GPS/GLONASS in the place where these signals were detected.

Methods of searching for a person based on known distance to the source of emergency signals

One of the methods for calculating the distance is the energy conversion of signals between the receiver and transmitter—the Received Signal Strength Indicator (RSSI). Because the UA is in the air, it can be assumed that there will be direct visibility between the mobile phone and the UA. Thus, it becomes possible to calculate the distance between the portable base station and the source of emergency signals. Figure 2 shows the interaction of a UA with a mobile phone. The coordinates of the UA are known, but to determine the coordinates of a mobile phone, it must receive at least three emergency signals. It will deliver it using the triangulation algorithms to determine the coordinates of the mobile phone.

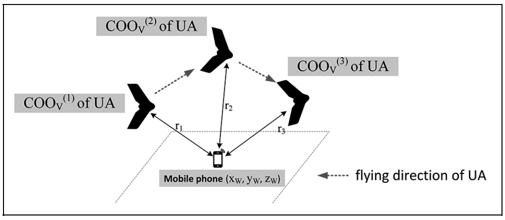


Figure 3. Method for determining the coordinates of a mobile phone using UA. UA: unmanned aircraft.

It is worth noting that the detection zone or radio coverage zone of the UA is limited to radius r, then in three-dimensional space, this zone can be represented as a ball with a known center (coordinates of UA) and radius r. The shape of the ball is represented by the equation

$$(x - x_v)^2 + (y - y_v)^2 + (z - z_v)^2 = r^2$$

where (x_v, y_v, z_v) are coordinates of UA (COO_v—vehicle coordinates).

Figure 3 shows the detection method of a mobile phone that sends emergency signals. According to the Figure 3, the desired point is found at a distance r, that is, it is located on a ball with radius r with three different UA positions and three distances between the UA and the mobile phone.

Therefore, we obtain a system of equations of the form

$$(x_{v_i} - x_W)^2 + (y_{v_i} - y_W)^2 + (z_{z_i} - z_W)^2 = r_i^2 \qquad (1)$$

where $i = 1, 2, ..., K$ are known coordinates of K points; r_i is the distance between the i-th position and the desired point; $(x_{v_i}, y_{v_i}, z_{z_i})$ are coordinates of the UA in the i-th zone; and (x_W, y_W, z_W) are coordinates of the desired point.

In the case of three-dimensional space, the system (1) must contain at least three controls, that is, $k \geqslant 3$, then the system of equations has a solution.

The presented system of equations can also be used for the case of using the UA group with i—the number of UA, then the probability of successfully determining the coordinates of a mobile phone is increased compared with the case where only one UA is considered, which carries out radio scanning for detecting emergency signals, for example, using three UAs that fly over terrain on the same plane and form a triangle, as shown in Figure 4.

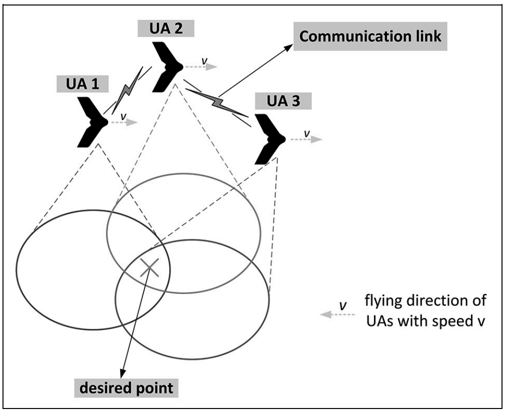


Figure 4. A group of three UA overflights. UA: unmanned aircraft.

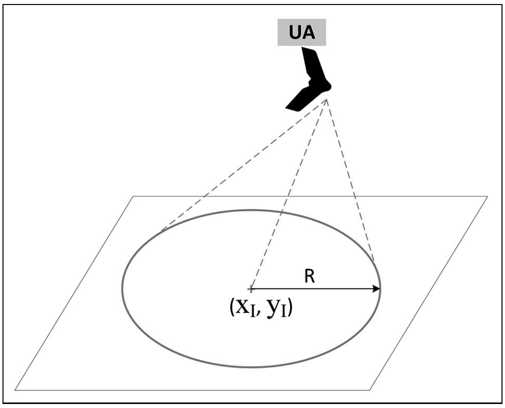


Figure 5. Representation of UA coordinates as a center of a circle with radius *R*. UA: unmanned aircraft.

Besides, the probability also depends on the flight speed of the UA and the frequency of sending emergency signals.

The method of detecting the coordinates of a mobile phone when the distance to the source of emergency signals is unknown

In the first case, the distance between the UA and the mobile phone, which is the source of emergency signals, was calculated based on the RSSI. The accuracy of determining the distance depends on the absence of interference on the path between the source and the receiver, as well as a direct line of sight. In some cases, there may be obstacles in the form of trees, fog, and other interfering environmental influences between the mobile phone and the UA; therefore, it is advisable to consider a method for detecting the coordinates of a mobile phone when the distance from the UA to the source of emergency signals is unknown. Each UA can detect emergency signals within a radius (R_v); therefore, a sphere is formed with a center (center is the UA position) and with a radius (R_v). This sphere intersects with the plane of the earth, so a circle is obtained with a center (x_I, y_I) and with a radius R on the plane of the earth. If the mobile phone (x_W, y_W) is in the circle zone, then the condition

$$(x_{v_i} - x_W)^2 + (y_{v_i} - y_W)^2 \leq R^2$$

is satisfied when considering the two-dimensional space on the earth plane. Figure 5 presents the case when the coordinates of the center of the circle (x_I, y_I) can be considered as the coordinates of UA (x_V, y_V).

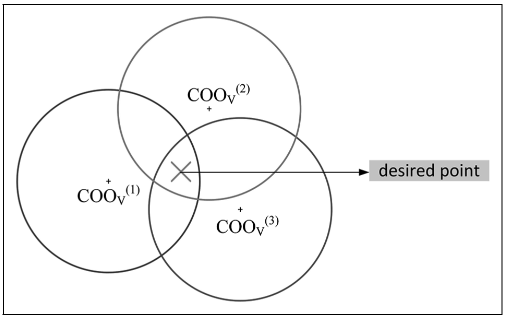


Figure 6. The method of detecting the coordinates of a mobile phone at the intersection of the zones with coordinates $COO_v^{(1)}$, $COO_v^{(2)}$, and $COO_v^{(3)}$.

Figure 6 showed the case when three positions of emergency signals were detected on the earth's surface. According to Figure 6, $COO_v^{(1)}$, $COO_v^{(2)}$, and $COO_v^{(3)}$ correspond to UA coordinates in three different positions.

A mobile phone (desired point) will be located at the intersection of several circles, with the coordinates of the centers of the rings and the radius R known; therefore, the coordinates of the desired point are given by expression (2)

$$\left((x_W - x_{v_i})^2 + (y_W - y_{v_i})^2\right) \leq R^2 \tag{2}$$

where $i = 1, 2, ..., K$ are known coordinates K points of UA; (x_{v_i}, y_{v_i}) are coordinates of UA in the i-th place; and (x_W, y_W) are coordinates of the desired point.

By expression (2), you can find the interval to which the coordinates of the mobile phone belong. When considering the case using the UA group (with i—the number of UA in the group), it will also use expression (2) to find the interval with the coordinates of the mobile

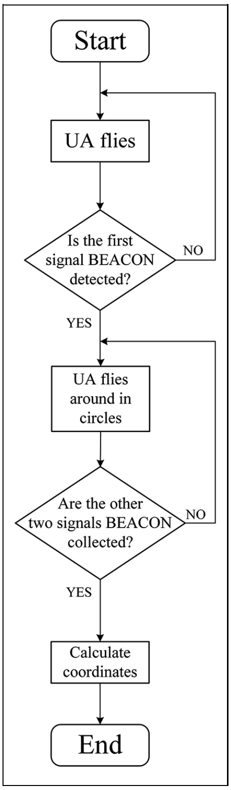

Start
UA flies
Is the first signal BEACON detected?
NO
YES
UA flies around in circles
Are the other two signals BEACON collected?
NO
YES
Calculate coordinates
End

Figure 7. Flow diagram illustrating the process of the search for beacon signals using UA. UA: unmanned aircraft.

phone. To calculate the coordinates of a mobile phone, if at least one UA from the group detects emergency signals, its position is saved and sent to the head node or base station to carry out calculations of the coordinates of the desired point.

Evaluation of searching time

The searching process using UA is represented by two steps:

- UA flies and scans beacon signals.
- When a UA detects the first beacon signal, in order to get enough of the beacon signals, UA begins flying around in a circle of radius R with center coordinates, which are the position of detecting the first beacon signal.

The process of the search for beacon signals using UA is presented in Figure 7.

It is assumed that speed of UA—v_{UA}, generation frequency of beacon signal—f_B, and distance to the place of detection of the first beacon signal in the flight direction of UA—S_{UA}. Searching time (T_{sum}) is the sum of flight time of UA needed to fly to the first place, where the first beacon signal was found (T_1) and flight time of UA needed to fly around in the circle (T_2). The desired point is in the searching zone with radius R. Having the speed of UA—v_{UA}, generation frequency of beacon signal—f_B, and distance to the place of detection of the first beacon signal in the flight direction of UA—S_{UA}, to detect the beacon, the generation frequency of the beacon signal must be satisfied by the formula (3)

$$\frac{1}{f_B} < \frac{2R}{v_{UA}} \Leftrightarrow v_{UA} < 2R \cdot f_B \tag{3}$$

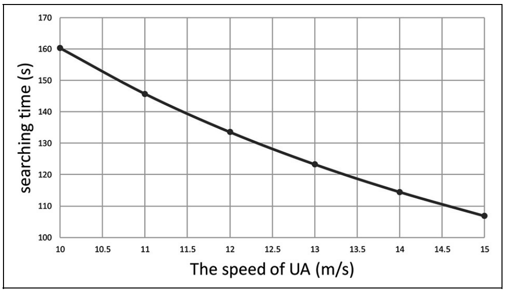


Figure 8. Dependence of searching time on the speed of UA. UA: unmanned aircraft.

To calculate the coordinates of the desired point, UA requires at least 2 more beacon signals after the first signal is detected. Since UA fly within a circle with radius R, the flight distance is $2 \cdot \pi \cdot R$, with k is a number of rounds, then we get the formula (4)

$$\frac{2 \cdot \pi \cdot R}{v_{UA}} \cdot k > \frac{2}{f_B} \Rightarrow k > \frac{2 \cdot v_{UA}}{2 \cdot \pi \cdot R \cdot f_B} \tag{4}$$

Therefore, the searching time is described in the formula (5)

$$T_{sum} = T_1 + T_2 = \frac{S_{UA}}{v_{UA}} + \frac{2 \cdot \pi \cdot R}{v_U A} \cdot k \tag{5}$$

An application, installed on the mobile phone generates beacon signals with a period of 12 s, thus speed of UA should be set to less than $2R \cdot f_B = \frac{2 \cdot 95.91}{12} = 15,98(\text{m/s})$, with $R = 95,91(\text{m})$ calculated in Coverage model for each group of UA. Calculation of the coverage area depending. By formula (4), we receive the number of rounds $k > 0,64$.

When changing the speed of UA from 10 to 15 m/s with $S_{UA} = 1000(\text{m})$, according to the formula (5), the searching time is showed in Figure 8.

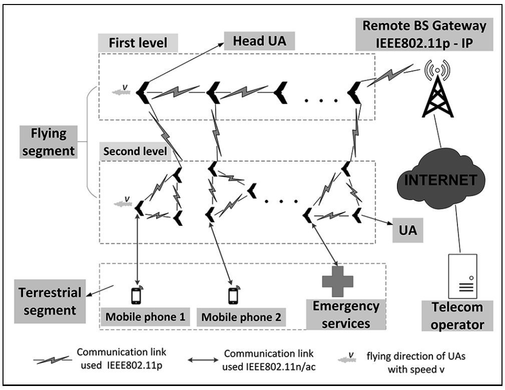


Figure 9. Unmanned aerial system hierarchical structure for providing voice services. UA: unmanned aircraft.

Voice transmission by a flying network

One of the possible scenarios in the search for a missing person is the lack of communication due to the remoteness of the cellular base station, as well as potentially difficult meteorological conditions when the communication range is significantly reduced. This section presents an alternative solution for the provision of communications based on UA. Today, most mobile phones support Wi-Fi technologies of various IEEE 802.11a/n/ac standards. This technology can be used to transmit voice over Wi-Fi based on voice over Wi-Fi (VoWi-Fi)[22,23] applications. In this case, the UA is represented as a mobile node with a base station gateway located on it through which it communicates with other subscribers.

The hierarchical structure of the UAS

To solve this problem, it is supposed to use a flying network consisting of a set of UA, which are mobile access points and relay the received/transmitted data to a stationary base station, which operates in the normal mode.[24] The novelty of this approach is that it is possible to launch the UA quickly, thereby introducing an additional mobile base station to provide communication between the person and the emergency operator. Since the distance between a person in the forest and the position of the emergency services operator is considerable, to ensure communication, many UA are required even within the line of sight. It is also worth considering that the signal is transmitted through several hops and due to an increase in network delay, the quality of voice transmission decreases. To improve the quality of service and also reduce the required number of UA, it is proposed to divide the UAS into two levels. The first level will consist of the head UA and the second level of one of the members of the UA in each group (Figure 9).

Second-level UAS interact with subscribers and search for the shortest route to transfer data to the head UA, which is located on the first level. Next, the first-level UAS transmits data to a stationary base station through other head UAs. Thus, instead of transferring data through all UA, the number of intermediate nodes is reduced by introducing two levels of hierarchy. The role of the mobile base station for subscribers will be performed by a Wi-Fi access point onboard the UA, which supports the IEEE 802.11n, IEEE 802.11ac, IEEE 802.11p standards. Because VoWi-Fi technology is widespread in a large number of mobile phone models, it can be assumed that this approach allows making calls over Wi-Fi while organizing a flying network with support for these technologies. It is also worth noting that all calls are made through the operator with the preservation of numbering and identification of subscribers of the mobile communication network.

We assume that the connection between the UA and the subscriber is based on the IEEE 802.11n/ac standard since currently, most mobile phones support these technologies. Communication between UA members in the group of the first level and between members of the second-level UAS is based on the IEEE 802.11p standard, which was developed for wireless information transfer between vehicles with the support of self-organization. Communication between the head UAs in different groups is performed using IEEE 802.11p* technology in advanced mode, within which data can be transmitted over a distance of 750 m.[25–27]

Coverage model for each group of UA: calculation of the coverage area depending on the height of the UA and the range of interaction

In the first stage, it is necessary to determine the relative position between the UA and the subscriber. To do this, set the source data. Suppose that a UA is flying at a constant speed (v); the distance between the UA and the subscriber is equal to the range of the selected data transfer technology—IEEE 802.11n/ac (d = 40m, since the subscriber is in the forest, the radio propagation distance is organic) . The distance between the UA, as well as between the members of the UA and the head UA, is constant and equal to the range of the selected data transfer technology—IEEE 802.11p in normal mode (D = 100m). The distance between the head UA is constant and equal to the radius of action of the selected data transfer technology—IEEE 802.11p in advanced mode (D_H = 500m).

According to Figure 10, a subscriber can make calls when a UA is present in the range of selected data standards—IEEE 802.11n/ac (point A) and when the UA

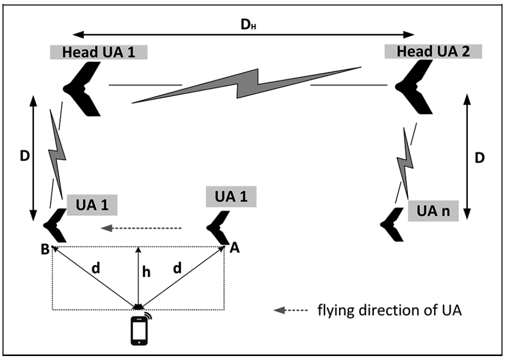


Figure 10. Schematic representation of the distance between all elements of the interaction model under consideration. UA: unmanned aircraft.

moves to point B, another UA from the group moves to point A to provide continuous communication. Points A and B are the most at a considerable distance, the distance between the UA and the subscriber is more (d = 40m). In this case, the altitude UA can be calculated by the formula (6)

$$h = d \cdot \sqrt{2} = 28,28(\mathrm{m}) \tag{6}$$

Now we calculate the distance between the members of UA in one group, which is equal to $2 \cdot h = 56,56(\mathrm{m})$; this distance is considered valid because the radius of action of the members of the UA is 100 m. Because the distance between the head nodes of the UAS is 500 m, then we have a modem to calculate the number of UA in one group, to ensure continuous communication using the formula (7)

$$N = max\left[\frac{D_H}{2 \cdot h}\right] + 1 = 10 \tag{7}$$

According to Figure 11, it can be seen that the maximum number of UA members in a group along a straight line is 10 (5 on the left and 5 on the right) and the maximum number of hops during voice transmission from the subscriber to the head UA is 5. The communication between two cluster heads in the MESH mode is enabled by the IEEE 802.11p standard.

Since the distance between the UA is 56.56 m, and the radius of data transmission at the UA is 100 m, overlapping zones may occur as shown in Figure 13. To calculate the coverage area of UA groups, it is necessary to calculate the coverage area at least between two UA.

The coverage area of a single UA is a circle with a radius $R_{cov} = 95,91(\mathrm{m})$ ($R_{cov} = IH = r^2 - h^2$) (Figure 12), which can be calculated using formula (8)

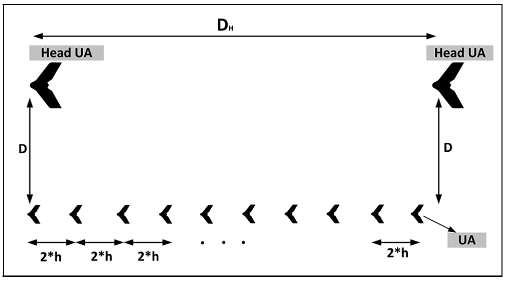


Figure 11. UA interaction in the group to ensure continuous communication for subscribers. UA: unmanned aircraft.

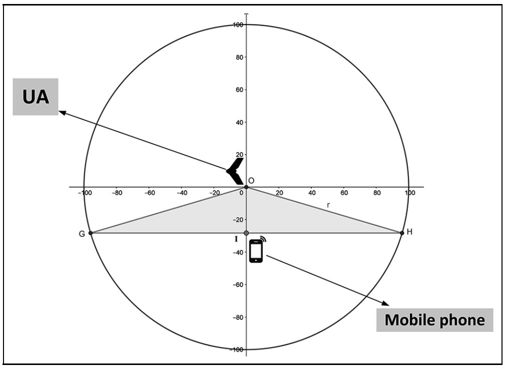


Figure 12. Coverage area of one UA. UA: unmanned aircraft.

$$S_{cov_1} = \pi \cdot R^2_{cov} \approx 28.884(\mathrm{m}^2) \tag{8}$$

In Figure 13, the coverage area between two UA is equal to the total coverage area of two UA minus the overlap of two areas and is equal to

$$S_{cov_2} = S_{cov_1} + S_{ex} \tag{9}$$

$$S_{cov_2} \approx 28.884 + 10.687,21 \approx 39.571,21(\mathrm{m}^2)$$

where S_{ex} is the excess area

$$S_{ex} = S_{cov_1} - 2 \cdot S_{\overset{\frown}{ACD}} + S_{ACBD} \tag{10}$$

As shown in Figure 13

$$S_{cov_1} = \pi \cdot R^2_{cov}$$

$$S_{\overset{\frown}{ACD}} = \pi \cdot R^2_{cov} \cdot \frac{2 \cdot arccos\frac{h}{R_{cov}}}{360}$$

$$S_{ACBD} = 2 \cdot h \cdot \sqrt{R^2_{cov} - h^2)}$$

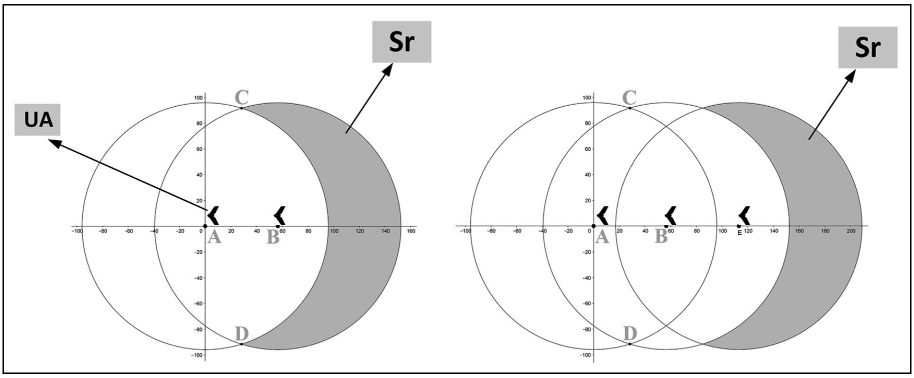


Figure 13. Schematic representation of the area of coverage between two UA, taking into account the overlap of two areas. UA: unmanned aircraft.

Since the distance between the UA in the same group is the same, the excess area (S_{ex}) is also the same; hence, we get the coverage area for one UA group consisting of $n = 10$ UA and equal to

$$S_{cov_{10}} = S_{cov1} + (n - 1) \cdot S_{ex} \tag{11}$$

$$S_{cov_{10}} \approx 28.884 + 9 \cdot 10.687,21 \approx 125.068,89(\mathrm{m}^2)$$

Continuous connection method to ensure guaranteed communication

In Wi-Fi networks, each station (STA) is associated with an access point (AP) located onboard the UA. When the UA moves, the access points move through the zones, respectively, so mobile phones reconnect to the available access point, but the transmission of voice traffic does not stop. This process is called handover. To ensure seamless connectivity, one needs to speed up the handover process. The handover process consists of four main steps:

1. Detection of a possible set of access points to which data can be transmitted;
2. Select the access point (AP) destination;
3. Communication with this access point;
4. (Re)authentication of the mobile station STA in the network.

The article[28] showed that mobile phone authentication is an essential requirement for security on Wi-Fi networks. In particular, due to the lack of a physical connection between the STA and the AP, authentication becomes indispensable for controlling access to the network. However, the authentication mechanisms used in Wi-Fi are quite slow and cannot guarantee low latency to ensure handover.

The handover process has three main elements:

- A mobile station (STA)—mobile phone;
- An access point (AP) located onboard the UAs, which are deployed in the second level (Figure 9);
- Server authentication, authorization, and accounting (AAA) located onboard the UAs, which are deployed in the first level (Figure 9).

The Figure 14 shows the delay time in each handover step.

T_{scan}—time delay in the scan phase,
T_{auth}—open authentication,
T_{asso}—association,
T_{1x}—IEEE 802.1X,[29]
T_{4way}—four-sided handshake.

Typical communications messages protocols involved in the process are represented by arrows between the vertical lines (Figure 14).

The first phase of the handover process is to check whether the conditions need to be changed by the AP and, if so, which AP should be associated with the STA. This phase can last several seconds, but this phase can do this without actually breaking the connection. The next stage of the handover process contains an empty authentication step, which is a legacy of WEP (Wired Equivalent Privacy) security architecture, which takes a very short time. The next step is the association phase, in which the STA establishes a logical connection with the AP. The purpose of this step is to notify the entire network that the STA can now be connected to any other AP. The time required for the association is insignificant, so there is no need to spend any money on accelerating this phase.

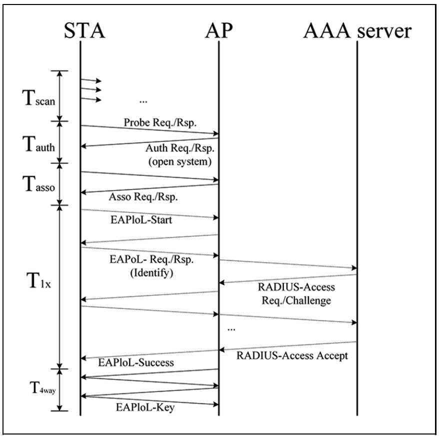


Figure 14. Interaction of elements to support the handover process.[28] STA: mobile station; AP: access point; AAA: authentication, authorization, and accounting.

The actual authentication phase begins after the association phase. In this phase, the STA is authenticated to the AAA server, which helps to configure a shared session between the STA and the AP. As we will see later, this phase can take a significant amount of time, especially if the AAA server is remote.

Finally, STA and AP perform a four-way handshake, as a result of which they agree on the knowledge of the session key with each other, and they also receive new keys from the shared session key for various purposes. A quadrilateral handshake is necessary to comply with the IEEE 802.11i standard; it cannot be shortened.

In Bohák et al.[28] and Aboba and Alimian,[30] it was proved that the authentication phase takes a lot of time, and therefore the right idea is to accelerate this phase.

To achieve this goal, in Bohák et al.[28] the authors propose modifying the EAP-SIM protocol, which is described in RFC 4186.[31] The authentication mechanism of the EAP-SIM protocol is based on the scheme used in GSM networks to authenticate subscribers. In the case of a Wi-Fi network, the STA and the AAA server share a public key K_i. When an STA moves from one base station to another base station, it must re-identify itself on the network. The AAA server sends so-called triplets to the base station, where each triplet contains a random RAND value, an Signed RESponse (SRES), and a session key K_c. The base station challenges the STA with RAND; using unique computed algorithms, the STA can read the SRES itself and send a response to the base station. If the reaction from the STA matches the SRES that the AAA server assigned to the base station earlier, then the STA is received.

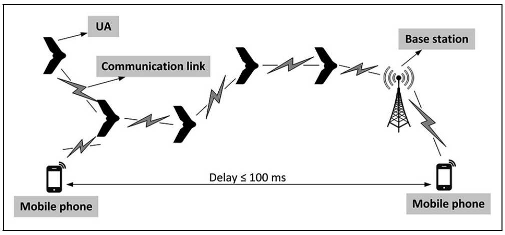


Figure 15. Schematic representation of the interaction of mobile phones through the chain UA. UA: unmanned aircraft.

The K_c session key will support the security of communication between the STA and the base station.

The basic idea of modifying the EAP-SIM protocol is to consider the case when a mobile device that is already connected to the network connects to another access point, in which case the triplets necessary for its authentication will already be available in the new access point and there is no need for remote communication. In the case where the handover process occurs, such triplets should already be possible, since they are preloaded during pre-authorization when the STA is still connected to the old access point. If the triplets are in the database, authentication can be performed locally. This means that the STA does not need to contact the AAA server during the handover process. In the end, the authors found a way to reduce authentication latency below 55 ms.

If we consider the application of the above method to our task, then it is known that mobile phones are stations (STAs), and first-level UAS will play the role of AAA North, and second-level UAS will play the role of an access point (AP). In this case, the handover time will be less than 55 ms, and thus, in the call time between the subscriber and the rescue services, every 13 s, the handover time will be below 55 ms (at a UA speed of 15 m/s). Since the handover time is concise, it can be assumed that it does not significantly affect the quality of the call.

Service quality model

Above, the hierarchical structure of UAS was considered as a flying network model, which is used to provide voice transmission services. In order to evaluate the performance of this structure, we will represent it as a Queuing System (QS), in which each UA is represented by an element that has the function of receiving and transmitting voice traffic.[32]

First, to use this service, a mobile phone must undergo an authentication procedure with a base station. After that, the mobile phone will get access to the service, and the UA nodes will also receive information

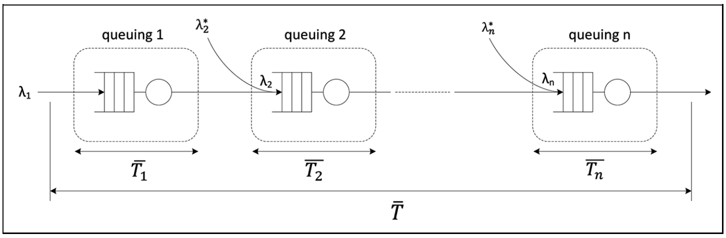


Figure 16. Queuing model of flying network for emergency call.

about mobile phones. A call between two subscribers, for example, between the missing subscriber and the rescue service, will be performed by interacting with each other through the UA chain. The UAS-based flying network model is satisfied with the condition of acceptable quality for voice transmission if the network delay is no more than 100 ms between two subscribers.[33]

Figure 15 shows schematically that the network delay in this scheme consists of the processing time, the waiting time for the queue, the transfer time between the subscriber unit and the UA, between the UA, and the transfer time between the UA and the base station. The average delay can be represented mathematically, as in the formula (12)

$$D_{average} = T_{S_1} + T_V + T_{COO} + T_{S_2} \tag{12}$$

where T_{S_1} is the Subscriber-Vehicle Delay (ms); T_V is the Vehicle-Vehicle Delay in each group (ms); T_{COO} is the Delay between Head Vehicle in each group (ms); and T_{S_2} is the Base Station-Subscriber Delay (ms).

To calculate the delays in this system, consider the multiphase QS model presented in Figure 16. Suppose that the incoming flows to each QS have the same properties. Consequently, it is possible to calculate the average time spent on the delivery of one application in each QS.

The average delivery time in multiphase QS can be represented by the expression (13)

$$T = \sum_{i=1}^{n} T_i = T_{S_1} + m \cdot T_V + (n-1) \cdot T_{COO} + T_{S_2} \tag{13}$$

where m is the maximum number of hops in each UA group and n is the number of head nodes UA.

According to the requirement of voice quality, it follows that the total time T should not exceed 100 ms. According to the formula (13) and the considered architecture of the flying network, it is possible to calculate the number of UA head nodes that are required for the organization of the network to transfer voice between two subscribers. Knowing the number of UA nodes, we will know the maximum possible distance between the base station and subscriber devices in the forest or the area of coverage in which subscriber devices should be located. It is assumed that the time between the subscriber and UA T_{S_1} and the time between the base station and the subscriber of T_{S_2} are the same; therefore, the formula (13) can be reduced by $T = 2 \cdot T_{S_1} + (m-1) \cdot T_V + (n-1) \cdot T_{COO} \leqslant 100.$ In this case, we have the following in equality (14)

$$n \leqslant \frac{100 - 2 \cdot T_{S1} - m \cdot T_V}{T_{COO}} + 1 \tag{14}$$

It is assumed that in Figure 16 each phase of the QS is considered an M/M/1 service model, in which the voice traffic has the properties of the most straightforward flow and the service time obeys an exponential distribution. At each phase or each QS, there is an intensity of loading the system, which can be represented by the formula (15)

$$\gamma_i = \rho_i = \frac{\lambda_i}{\mu_i}(Erl) \tag{15}$$

where λ_i is the intensity of incoming applications (request/ms) and μ_i is the intensity of service requests (requests/ms).

When considering the M/M/1 model, we will use the formula to calculate the time of the application in the Kleinrock[34] system. Accordingly, the average delivery time in each QS can be calculated using equation (16)

$$T_i = w_i + t_i = \frac{t_i}{1 - \rho_i} \tag{16}$$

where T_i is the average delivery time in phase i (ms); w_i is the average waiting time in the queue (ms); and t_i is the average duration of service requests (ms).

As noted above, subscriber devices currently support IEEE 802.11n/ac Wi-Fi technologies, which provide relatively high data transfer rates. According to the IEEE 802.11n standard, the maximum data transfer rate is b_n = 300Mbps, and in the IEEE 802.11n standard, b_{ac} = 650Mbps. Communication between UA is

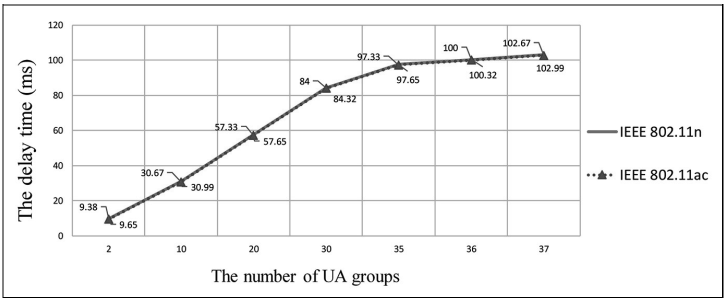


Figure 17. The change in the delay time, depending on the number of UA groups. UA: unmanned aircraft.

carried out according to the IEEE 802.11p standard in two modes. In each group, UA interact with each other, as well as with the head UA based on the IEEE 802.11p standard in the usual mode with a transmission rate of $b_{p_1} = 12\text{Mbps}$. At the same time, head UAs interact with each other by the IEEE 802.11p standard in advanced mode with a transmission speed of $b_{p_2} = 6\text{Mbps}$. In advanced mode, and the data transmission distance can reach up to 740 m; therefore, to ensure the guaranteed transmission speed, we choose the value of 6 Mbps. In normal mode, for IEEE 802.11p, the data transmission distance is about 100 m, and the transmission speed can reach up to 12 Mbps.[35,36]

Consequently, it is possible to calculate the average duration of service requests by the formula (17)

$$t_i = \frac{L}{b_i} \tag{17}$$

where L is the average length of one packet (bit); $L = 1000\text{bytes}$ is allowed; and b_i is the average data transfer rate (bit/ms).

Previously, the hierarchical structure of the UAS organization was considered, based on the introduced assumptions, it was possible to find out the number of nodes (hops) in each UA group, that is, $m = 5$, to satisfy the conditions of full coverage and provision of communication. Therefore, using formulas (13), (16), and (17), we can calculate the average delivery time when changing the number of head nodes (n), respectively, with a different number of UA groups. If we assume that the load factor of each phase of the QS $\rho_i = 0.5$, then the results of the calculation of the delivery time are presented in Table 1.

Table 1 shows the calculation of the time of delivery of voice between subscribers when changing the number of UA for the model M/M/1 when loading the entire system 0.5. According to the table, the required number of UA is established to cover the area and provide communication in the forest. Under this condition, the voice delivery time does not exceed 100 ms; then, according to Table 1, we obtain the maximum number of UA groups—35 UA groups when the IEEE 802.11n standard is used; 36 groups when using the IEEE 802.11ac standard. It is seen that the difference is insignificant because the main delay time for voice transmission is the transmission time through the UA, which is used in the IEEE 802.11p standard. We know that the more UA groups, the more extensive the coverage. Therefore, the results that are presented in Table 1 also show the maximum coverage of the area, provided that the voice delivery time is no more than 100 ms. In a real situation, there is a dependence on the distance between the target zones and the base station, in which case we can determine the number of important UA groups for deployment, but not more than 35 groups. For example, the distance between base stations is usually in a section of 3–5 km; if one needs to transmit a voice (about 5 km), you need about 11 UA groups, then you get a coverage area of $11 \cdot 125.068,89 = 1.375.757,79(m^2)$.

Table 1. Delivery time depends on the number of unmanned aircraft (UA) groups.

Number of UA groups	Delivery time (ms)	
	IEEE 802.11n	IEEE 802.11ac
2	9.65	9.38
10	30.99	30.67
20	57.65	57.33
30	84.32	84.00
35	**97.65**	97.33
36	100.32	**100.00**
37	102.97	102.67

Bold values highlight the maximum number of UA groups where delivery time does not exceed 100 ms.

According to the data in Table 1, we also get graphs of the changes in the delay time, depending on the number of UA groups (Figure 17). As can be seen from the

Table 2. Delivery time and the number of unmanned aircraft (UA) groups, depending on the load factor.

Load factor, ρ	IEEE 802.11n		IEEE 802.11ac	
	Delivery time (ms)	Number of UA groups	Delivery time (ms)	Number of UA groups
0.1	98.57	65	98.54	65
0.2	99.23	58	99.19	58
0.3	98.17	50	98.13	50
0.4	98.98	43	98.93	43
0.5	97.44	35	97.38	35
0.6	98.47	28	98.39	28
0.7	95.73	20	95.64	20
0.8	96.93	13	96.79	13
0.9	87.20	5	86.91	5

graph, the voice transmission latency for the mobile phones supported by either IEEE 802.11n or IEEE 802.11ac are not significantly different. However, if the voice transmission data are transported across multiple UA groups, the latency will naturally increase. Most of this delay will depend on the number of head nodes (head UA), through which packets are passed, and the technology used to communicate between UAs.

Furthermore, the model using UA is considered when changing the load factor of each phase of the QS. Consequently, we obtain the average delivery time and, accordingly, the number of UA groups depending on the load intensity. According to the formulas (13), (14), (16) and (17), the results of calculating the delivery time and the number of UA groups are presented in Table 2.

It is seen that the higher the number of UA groups, the greater the delivery time between the two subscribers. Under Table 2, which shows the dependence of the number of UA groups, which can provide the required quality of service ($delay \leqslant 100ms$), on the load factor in the service system. As mentioned above, the larger the UA group, the larger the coverage area. Therefore, if the load factor reached 0.9, respectively, 5 UA groups can provide a transmission delay of less than 100 ms; then the coverage area is reduced, and 5 UA groups can provide a coverage area of $(5 \cdot 125.068, 89 = 625.344, 45(m^2))$.

According to the data in Table 2, we also get graphs of changes in the number of UA groups, depending on the load factor; with the maximum delivery time may not exceed 100 ms (Figure 18). Load factor increases corresponding to the growth of requests sent to the head nodes. This suggests that many subscribers requiring a connection are not in the same UA group, but in other groups on the link. The higher the load factor is, the narrower the radius will be for guaranteed voice transmission, for example, when $loadfactor = 0.9$, the QoS (quality of service) of the voice transmission is only guaranteed in the radius of 5 UA group.

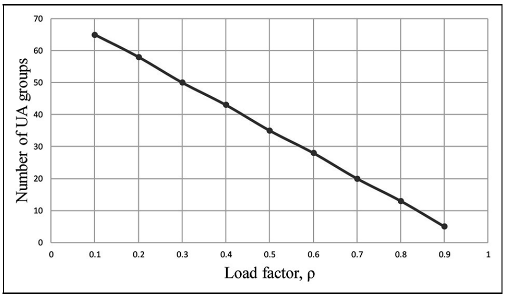


Figure 18. The changes in the number of UA groups, depending on the load factor. UA: unmanned aircraft.

Conclusion

The article reviewed the methods wilderness searches for a person using beacon signals from a mobile phone by using UAS for two situations: when the distance to the source of emergency signals is known and when the distance is unknown. We considered the simultaneous use of UAs involved in the search mission to support the voice communication among the search and rescue team on the ground using Wi-Fi technology. The architecture of a network model for connecting mobile subscribers was presented, which is organized on the basis of a flying network, in which IEEE 802.11p technology is used for UAs communication, and IEEE 802.11n/ac technology for communication between UAs and mobile phones. We also considered the impact of handover on quality of service of calls, as UAs become mobile base station. Quantitative and qualitative values are obtained, which allowed us to use equipment more efficiently, as well as minimize network latency. The article analyzed the model of multiphase queuing system type M/M/1 which are considered for UAs communication, as well as for communication between mobile phones and UAs. We calculated the voice transmission delay and the number of UAs. The results showed that the use of a two-tier architecture has reduced number of UAs, at which permissible quality of service of calls in the disaster zone can be guaranteed and the maximum number of UAs can be achieved while still guarantying the quality of the call. It also means that the coverage of the area can be achieved with this amount of UAs. As further research, an experiment is planned to confirm the results.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The publication has been prepared with the support of the "RUDN University Program 5–100."

ORCID iDs

Truong Duy Dinh https://orcid.org/0000-0002-9993-9792
Andrei Vladyko https://orcid.org/0000-0002-8852-5607

References

1. IAMSAR. *IAMSAR—international aeronautical and maritime search and rescue manual*, vol. I–III. London: International Maritime Organization and International Civil Aviation Organization, 2016.
2. Kirichek R, Vladyko A, Paramonov A, et al. Software-defined architecture for flying ubiquitous sensor networking. In: *Proceedings of the international conference on advanced communication technology (ICACT)*, Bongpyeong, South Korea, 19–22 February 2017, pp.158–162. New York: IEEE.
3. Sharma V, Song F, You I, et al. Energy efficient device discovery for reliable communication in 5G-based IoT and BSNs using unmanned aerial vehicles. *J Netw Comput Appl* 2017; 97: 79–95.
4. Sakano T, Fadlullah ZM, Ngo T, et al. Disaster-resilient networking: a new vision based on movable and deployable resource units. *IEEE Netw* 2013; 27(4): 40–46.
5. Motlagh NH, Taleb T and Arouk O. Low-altitude unmanned aerial vehicles-based internet of things services: comprehensive survey and future perspectives. *IEEE Internet Things J* 2016; 3(6): 899–922.
6. Sánchez-García J, Reina D and Toral S. A distributed PSO-based exploration algorithm for a UAV network assisting a disaster scenario. *Future Gener Comput Syst* 2019; 90: 129–148.
7. Niedzielski T, Jurecka M, Mizinski B, et al. A real-time field experiment on search and rescue operations assisted by unmanned aerial vehicles. *J Field Robot* 2018; 35(6): 906–920.
8. Taleb T, Bottazzi D, Guizani M, et al. Angelah: a framework for assisting elders at home. *IEEE J Select Area Commun* 2009; 27(4): 480–494.
9. Taleb T, Bottazzi D and Nasser N. A novel middleware solution to improve ubiquitous healthcare systems aided by affective information. *IEEE T Inform Technol Biomed* 2010; 14(2): 335–349.
10. Hanna D, Ferworn A, Lukaczyn M, et al. Using unmanned aerial vehicles (UAVs) in locating wandering patients with dementia. In: *Proceedings of the IEEE/ION position, location and navigation symposium (PLANS)*, Monterey, CA, 23–26 April 2018, pp.809–815. New York: IEEE.
11. Rasmussen ND, Morse BS, Goodrich MA, et al. Fused visible and infrared video for use in wilderness search and rescue. In: *Proceedings of the workshop on applications of computer vision (WACV)*, Snowbird, UT, 7–8 December 2009, pp.1–8. New York: IEEE.
12. Munoz-Castaner J, Soto PC, Gil-Castineira F, et al. Your phone as a personal emergency beacon: a portable GSM base station to locate lost persons. *IEEE Ind Electron Mag* 2015; 9(4): 49–57.
13. Bull JF. Wireless geolocation. *IEEE Veh Technol Mag* 2009; 4(4): 45–53.
14. Mendelson E. *System and method utilizing integral wireless protocols of a mobile phone as an emergency beacon to aid first responders in locating people*. US9374673 Patent, 2016.
15. Wolfe V, Frobe W, Shrinivasan V, et al. Detecting and locating cell phone signals from avalanche victims using unmanned aerial vehicles. In: *Proceedings of the international conference on unmanned aircraft systems (ICUAS)*, Denver, CO, 9–12 June 2015, pp.704–713. New York: IEEE.
16. Apodaca V, McConnell C, Maguire D, et al. *Emergency beacon for cell phone or the like*. Patent application 11/755533, USA, 2008.
17. Ho YH, Chen YR and Chen LJ. Krypto: assisting search and rescue operations using Wi-Fi signal with UAV. In: *Proceedings of the first workshop on micro aerial vehicle networks, systems, and applications for civilian use*, Florence, 18 May 2015, pp.3–8. New York: ACM.
18. Wang W, Joshi R, Kulkarni A, et al. Feasibility study of mobile phone WiFi detection in aerial search and rescue operations. In: *Proceedings of the 4th Asia-Pacific workshop on systems*, Singapore, 29–30 July 2013, p.7. New York: ACM.
19. Goodrich MA, Cooper JL, Adams JA, et al. Using a mini-UAV to support wilderness search and rescue: practices for human-robot teaming. In: *Proceedings of the international workshop on safety, security and rescue robotics*, Rome, 27–29 September 2007, pp.1–6. New York: IEEE.
20. Guo Z, Wei Z, Feng Z, et al. Coverage probability of multiple UAVs supported ground network. *Electron Lett* 2017; 53(13): 885–887.
21. Mohibullah W and Simon J. Stigmergic search for a lost target in wilderness. In: *Proceedings of the sensor*

signal processing for defence (SSPD 2011), London, 27–29 September 2011, pp.31–35. New York: IEEE.

22. Chagh Y, Guennoun Z and Jouihri Y. Voice service in 5G network: towards an edge-computing enhancement of voice over Wi-Fi. In: *Proceedings of the 39th international conference on telecommunications and signal processing (TSP)*, Vienna, 27–29 June 2016, pp.116–120. New York: IEEE.
23. Ngongang SFM, Tadayon N and Kaddoum G (2016) Voice over Wi-Fi: feasibility analysis. In: *Proceedings of the advances in wireless and optical communications (RTUWO)*, Riga, 3–4 November 2016, pp.133–138. New York: IEEE.
24. Dinh TD, Pham VD, Kirichek R, et al. Flying network for emergencies. In: *Proceedings of the international conference on distributed computer and communication networks*, Moscow, 17–21 September 2018, pp.58–70. Cham: Springer.
25. Anwer MS and Guy C. A survey of VANET technologies. *J Emerg Trend Comput Inform Sci* 2014; 5(9): 661–671.
26. Cecchini G, Bazzi A, Masini BM, et al. Performance comparison between IEEE 802.11p and LTE-V2V in-coverage and out-of-coverage for cooperative awareness. In: *Proceedings of the vehicular networking conference (VNC)*, Torino, 27–29 November 2007, pp.109–114. New York: IEEE.
27. Wang Q, Leng S, Fu H, et al. An IEEE 802.11p-based multichannel MAC scheme with channel coordination for vehicular ad hoc networks. *IEEE T Intell Transp Syst* 2012; 13(2): 449–458.
28. Bohák A, Buttyán L and Dóra L. An authentication scheme for fast handover between WiFi access points. In: *Proceedings of the ACM wireless internet conference (WICON)*, Austin, TX, 22–24 October 2007. New York: ACM.
29. IEEE 802.1X:2004. IEEE standard for local and metropolitan area networks: port-based network access control.
30. Aboba B and Alimian A. *Analysis of roaming techniques* (IEEE 802.11-04/0377r1). New York: IEEE, 2004.
31. Haverinen H and Salowey J. Extensible authentication protocol method for global system for mobile communications (GSM) subscriber identity modules (EAP-SIM) (No. RFC 4186), 2005, https://tools.ietf.org/html/rfc4186
32. Koucheryavy A, Vladyko A and Kirichek R. State of the art and research challenges for public flying ubiquitous sensor networks. *Lect Note Comput Sci* 2015; 9247: 299–308.
33. ITU-T Recommendation G.114. *One-way transmission time*. Geneva: International Telecommunication Union, 2003.
34. Kleinrock L. *Queueing systems, volume 2: computer applications*, vol. 66. New York: Wiley, 1976.
35. Jiang D, Chen Q and Delgrossi L. Optimal data rate selection for vehicle safety communications. In: *Proceedings of the fifth ACM international workshop on vehicular inter-networking*, San Francisco, CA, 15 September 2008, pp.30–38. New York: ACM.
36. Li YJ. An overview of the DSRC/WAVE technology. In: *Proceedings of the international conference on heterogeneous networking for quality, reliability, security and robustness*, Houston, TX, 17–19 November 2010, pp.544–558. Berlin: Springer.

PUBLICATION

V

Performance of mmWave-Based Mesh Networks in Indoor Environments with Dynamic Blockage

R. Pirmagomedov, D. Moltchanov, V. Ustinov, M. N. Saqib and S. Andreev

International Conference on Wired/Wireless Internet Communication. Ed. by D. F. M., N. E., B. R. and K. A. 2019, 129–140

DOI: 10.1007/978-3-030-30523-9_11

Performance of mmWave-based Mesh Networks in Indoor Environments with Dynamic Blockage

Rustam Pirmagomedov[1], Dmitri Moltchanov[2], Viktor Ustinov[3], Md Nazmus Saqib[2], and Sergey Andreev[2]

[1] Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation
[2] Tampere University, Korkeakoulunkatu 6, Tampere, 33720, Finland
[3] St.Petersburg State University of Telecommunication, Bolshevikov 22/1, St.Petersburg, 190000, Russian Federation

Abstract. Due to the growing throughput demands driven by innovative media applications (e.g., streaming 360° video, augmented and virtual reality), millimeter wave (mmWave) wireless access is considered as a promising technological enabler for emerging mobile networks. One of the critical uses for such systems is indoor public protection and disaster relief (PPDR) applications which may greatly benefit from new high bandwidth applications. In this paper, we evaluate the performance of on-demand mmWave mesh systems in indoor environments with dynamic blockage conditions, 3GPP propagation model, mobile nodes, and multi-connectivity operation. To evaluate the performance of the mesh we have developed a system-level simulation framework based on a realistic floor layout. Our numerical assessment has revealed that the use of a multi-connectivity operation in indoor deployments allows for efficiently improving connectivity while slowly improving the associated per-node throughput. The latter implies that mmWave systems in indoor deployments operate in a blockage-rich environment which differs from outdoor environments. Furthermore, the number of simultaneously supported links at each node, required to improve the system performance, can be greater than two imposing significant control overheads.

Keywords: Milliliter wave mesh · 5G· PPDR · Emergency response · Indoor environments

1 Introduction

Due to the growing capacity demands on the air interface driven by innovative media applications, e.g., 360° HD streaming, augmented and virtual reality (AR/VR), the millimeter wave (mmWave) wireless access is considered as a promising technology for future mobile networks.

In addition to extraordinary promises, the mmWave systems bring new challenges to systems designers including high propagation losses, sensitivity to blockage by obstacles, beamsteering functionality for highly directive transmission [1]. Particularly, the free space propagation loss at 60 GHz, is 28 decibels greater than that at 2.4 GHz [2]. For these reasons, mmWave links are highly directional, and they use steerable antenna arrays with sufficient gain to compensate for these extreme losses. Moreover, due to short wavelength (approximately, 5 mm. at 60 GHz), even small objects may block the mmWave link [3]. Finally, the mmWave links are also highly affected by atmospheric and molecular absorption [4]. Altogether these three factors significantly affect the reliability of mmWave communication.

In addition to conventional scenarios such as cellular systems, where 3GPP is currently at the completion phase of New Radio standardization, the extreme throughput makes mmWave communications technology appealing for many other use cases. Mainly, the natural extension of the mmWave mobile access is in wireless mesh networking. The mmWave meshes may extend the coverage of mmWave access points by utilizing multi-hop links and simultaneously enhance the reliability of communications using multi-connectivity operation [5]. One of the critical use cases in this context is public protection and disaster relief (PPDR) missions which may greatly benefit from using 360° HD streaming and AR/VR applications enabled by bandwidth-rich mmWave communications technology.

Performance of mmWave systems with multi-connectivity capabilities in the outdoor environment has been thoroughly investigated. An upper bound for the capacity of mmWave systems with multi-connectivity operation has been obtained in [6]. This bound has been refined in [7, 8]. Recently, engineering studies addressing multi-connectivity aspects of mmWave networks have started to appear [9, 10]. These studies reveal that multi-connectivity operation improves both outage probability and system throughput exponentially. These results have been extended to mmWave mesh deployments in the outdoor environment in [11–13], where the multi-connectivity operation has also been shown to improve network connectivity and throughput greatly.

Compared to the outdoor environment, indoor mmWave mesh deployments brings additional challenges that are primarily caused by an extremely complex propagation process. The use of the mmWave meshes in dynamic PPDR use cases, such as fire suppression missions where the environment may dynamically change, adds another level complexity. Notably, the blockage in such indoor use cases is induced by the interior of the building (e.g., partitions, walls), and dynamic moving objects (e.g., people, moving equipment). Taking into account the potential mobility of nodes, indoor mmWave mesh deployments are expected to be highly dynamic with constantly changing connectivity patterns between mesh nodes. In these conditions, the multi-connectivity can be considered as one of the vital options not only to maintain network connectivity at all times but to improve throughput.

In this paper, we evaluate the performance of mmWave mesh systems in a realistic indoor environment with the mobility of nodes, dynamic blockers, 3GPP propagation, and multi-connectivity operation. Particularly, we consider the PPDR scenario as an illustrative use case and investigate network connectivity and throughput characteristics. Our main conclusions are:

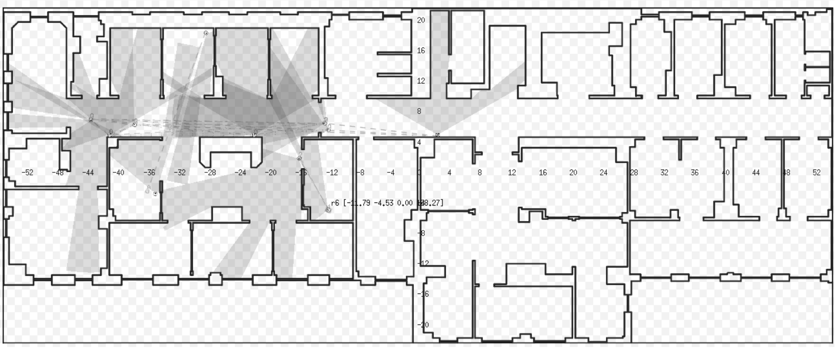


Fig. 1. Nodes on the layout during the simulation 2D view.

- the use of multi-connectivity operation exponentially improves mmWave mesh connectivity in indoor deployments, but its effect on per-node throughput is linear implying that indoor systems operate in blockage-rich environments;
- to improve connectivity and throughput in dense indoor mmWave meshes the number of simultaneously supported links is greater than two implying significant control signaling overheads.

The rest of the paper is organized as follows. In Section 2 we introduce the system model and its components. The system-level simulation framework and data collection/analysis procedure are described in Section 3. We report our results in Section 4. Conclusions are provided in the last section.

2 System model

In this section, we introduce our system model by first defining the scenario of interest and briefly specifying its submodules including mobility, propagation, beamforming, and dynamic blockage models. Finally, we introduce the connectivity process and define metrics on interest.

Illustrative scenario. We consider a fire suppression mission as an illustrative example for evaluation of the mmWave mesh network performance in a realistic single-floor indoor deployment, see Fig. 1. Specifically, we consider a team of firefighters operating on the floor of office building. The team utilizes assisting media applications enabled by mmWave mesh. The application acquires the information streams from the firefighters' on-body video cameras and transmits it to the command center. The command center processes the video streams and develops an optimal team operation strategy, providing digital assistance and guidance to the firefighters. In this scenario, we assume that the telecommunication infrastructure inside the building is not working. The communication with the command center is provided using a relay node (access point) installed by the rescue team near a window.

Nodes mobility model. To capture the mobility of nodes, we assume that they move according to random direction movement model (RDM, [14]) since it captures the essentials of random movement and still allows for analytical tractability. According to this model, a node first randomly chooses the direction of movement uniformly in $(0, 2\pi)$ and then moves in this direction at constant speed v_B for exponentially distributed time with parameter $\gamma = 1/E[\tau]$, where τ is the mean movement duration. The process is restarted at the stopping point. If during the movement an obstacle is reached, the node chooses a new direction of movement.

Propagation model. The received signal power at a mesh node is given by

$$P_R(x) = P_T G_T G_R - PL, \tag{1}$$

where P_T is the transmitter power, G_T and G_R are the antenna gains at the transmitter and receiver sides, respectively, which depend on the antenna array (these parameters can be obtained from the beamforming model introduced in the next subsection), PL - path loss. Following 3GPP TR 38.901, the mmWave path loss in dB for the line-of-sight (LoS) link is given by

$$PL_{InH-LOS} = 32.4 + 17.3\,\mathrm{lg}(d_{3D}) + 20\,\mathrm{lg}(f_c), \tag{2}$$

where f_c is the centre frequency, d_{3D} is the three-dimensional distance between wireless interfaces of two communicating nodes.

For the non-LoS (NLoS) link the path loss defined as

$$PL_{InH-NLOS} = \max(PL_{InH-LOS},\ PL'_{InH-NLOS}), \tag{3}$$

where

$$PL'_{InH-NLOS} = 38.3lg(d_{3D}) + 17.3 + 24.9\,\mathrm{lg}(f_c). \tag{4}$$

Beamforming model. Following [15], we assume linear antenna arrays at both the transmitter and receiver sides. Half-power beamwidth (HPBW) of the array, α, is assumed to be proportional to the number of elements as [16]

$$\alpha = 2|\theta_m - \theta_{3db}|, \tag{5}$$

where θ_{3db} is the 3-dB point and θ_m is the array maximum. Note that $\theta_m = \arccos(-\beta/\pi)$, where β is the array direction angle.

Assuming $\beta = 0$, we have $\theta_m = \pi/2$. The upper and lower 3-dB points are

$$\theta^{\pm}_{3db} = \arccos[-\beta \pm 2.782/(N\pi)], \tag{6}$$

where N is the number of antenna elements.

For $\beta = 0$, the mean antenna gain over HPBW is computed as [16]

$$G = \frac{1}{\theta^{+}_{3db} - \theta^{-}_{3db}} \int_{\theta^{-}_{3db}}^{\theta^{+}_{3db}} \frac{\sin(N\pi\cos(\theta)/2)}{\sin(\pi\cos(\theta)/2)} d\theta. \tag{7}$$

Dynamic Blockage Model. In this paper, we consider three types of blockages: (i) blockage by inherent indoor constructions, e.g., walls, furniture, (ii) self-blockage, and (iii) dynamic blockage by environmental objects. The former type of blockage is captured by the considered propagation model introduced in the previous subsection. Self-blockage refers to special positioning on a node such that it may no longer beamform its antenna towards the intended recipient.

We also assume a dynamic spatially-temporal blockage model. According to this model, blockers appear at a randomly chosen position uniformly distributed over the area of a floor according to homogeneous temporal Poisson process with intensity λ. Each blocker is assumed to exist for an exponentially distributed period of time with mean $1/\mu$. Observe that this stochastic process is inherent of M/M/∞ type and the number of active blockers given by Poisson distribution with the parameter λ/μ. Given the radius of the blocker r_B, the fraction of a floor covered by this type of blockers can be obtained using the integral geometry as follows [17]

$$p_C = (1 - f_{C,1})^{\lambda/\mu},\ p_{C,1} = \frac{2\pi S_B}{2\pi(S_A + S_B) + L_A L_B}, \tag{8}$$

where S_A is the floor area, $S_B = \pi r_B^2$ is the blocker's radius.

Connectivity and Metrics of Interest. To improve mmWave performance, we assume that a single node supports 3GPP multi-connectivity option described in Rel. 15 NR specification TS 37.340 [5]. According to it, UE simultaneously supports multiple connections to adjacent systems and may dynamically switch between them in case the current connection is lost. In our study, the number of simultaneously supported connections, known as the degree of multi-connectivity, is assumed to be M. It should be noted that depending on the node locations forming the mesh, the actual number of connections at any given moment of time can be less than M.

In our study, we address connectivity and throughput performance metrics. These are: (i) the fraction of time at least one node is disconnected from the mesh network, (ii) the mean number of disconnected nodes at the arbitrary instant of time and (iii) mean per-node throughput.

3 Simulation Framework

3.1 Simulator Design

For numerical evaluation of the mmWave mesh performance within the considered scenario, we developed a custom simulator based on the Stage simulator code [18, 19].

The developed framework is based on discrete-event simulation (DES) techniques. The simulation procedure comprises two phases: DES simulations and data analysis. The developed DES framework implements the system model described in Section 2. New arrival events are generated according to a spatial

Poisson process where the positions are distributed uniformly within the area of interest. When processing the arrival events, the session completion event and the subsequent arrival event are scheduled.

The primary part of the simulator is a 3D model of an office floor, see Fig.2. The 3D model of the floor allows for evaluation if there is a LoS between two points on the coordinate plane. The DES simulation process starts with the coordinate simulation of the nodes movements and dynamic blockages. For each iteration, the simulator checks the LoS condition between all the nodes of the mesh. When checking LoS condition, the antenna directivity diagram is also taken into account. The results are stored in the SQL database.

The second phase utilizes results obtained during the first phase for evaluating parameters of interests. The second phase starts with the calculation of path losses between all the nodes using the system model described in Section 2. If the signal strength between two nodes (on the receiver side) is lower than the established threshold, the simulator assumes that there is no direct connection between these nodes. If there is a connection between two nodes, the simulator calculates the throughput using Shannon–Hartley theorem for all pairs of nodes in the mesh where direct connection is available. This part represents the physical level of the network. In the next step, the simulator considers channels between nodes, including medium access control. This level delivers a channel topology graph. The third step represents addressing and routing within the mesh topology delivered by the second step.

The developed framework is very flexible from the point of modification. The modifications (e.g., implementation of different protocols or applications) can be performed by applying modified SQL scripts to the coordinate simulation traces stored in the SQL database, without launching new simulation.

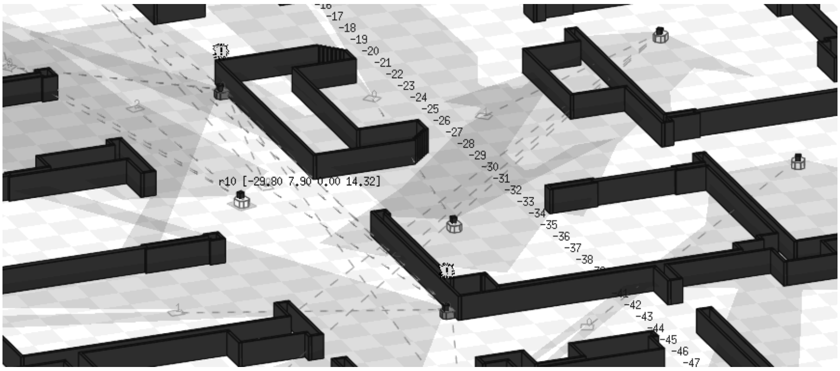


Fig. 2. Three-dimensional view of mesh nodes in the simulated layout.

3.2 Data Collection and Analysis

A simulation campaign has also been carried out to obtain the metrics of interest by relying on the following procedure. For each considered set of input parameters (simulation round), simulations were set to run for 1200 seconds of the system time, with a 0.25 change of system time on each iteration of simulation. The chosen duration of simulation round approximately corresponds to the time required for checking the floor of considered size by rescue team (e.g., firefighters).

It is assumed that all of the involved processes (arrival, service, blockage) are stationary; the steady-state always exists in our system. The starting point of the steady-state period has been detected by utilizing the exponentially-weighted moving average (EWMA) statistics with the weighting parameter set to 0.05 and employing the procedure in [20].

The statistical data has been collected only during the steady-state period. To remove residual correlations in the statistical data, we have used the batch means strategy. Accordingly, the entire steady-state period duration has been divided into 1000 data blocks. The metrics of interest computed for these periods became inputs to the individual statistical samples. The final values for the metrics of interest have been estimated by processing these samples. Due to the large size of statistical samples associated with experiments, only the point estimates are shown. The interval estimates computed for the selected input parameters do not deviate by more than ±0.001 from the point estimates under the level of significance set to $\alpha = 0.05$ and thus are not plotted in the presented graphs.

4 Numerical Results

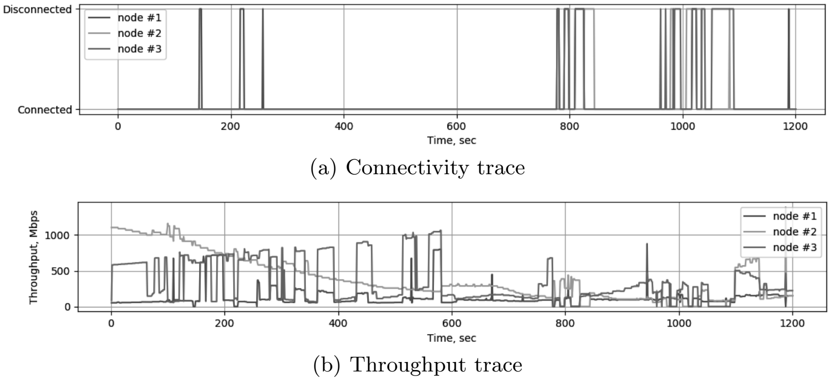
In this section, we report numerical results of mesh performance in indoor environments. The default system parameters are reported in Table 1.

The indoor deployments of mmWave systems are characterized by much greater complexity compared to widely considered outdoor scenarios. Thus, to obtain intuition about the system under investigation we start our analysis assessing the time-dependent behavior of connectivity and throughput processes three randomly selected mesh nodes illustrated in Fig. 3. Observing the connectivity process shown in Fig. 3(a), where 0 indicates connectivity periods and 1 implies the absence of an active connection, one may conclude that connectivity intervals are rather long compared to outages. Outage intervals are rather short, but their frequency is relatively high. This behavior is a consequence of the realistic indoor deployment and mobility model, where layout specifics and dynamic blockage result in many short-lasting outage events.

The associated throughput obtained by nodes is illustrated in Fig. 3(b), where the throughput averaging interval was set to 1 s. As one may observe, the throughput may drastically deviate even during the connectivity intervals. For some nodes, these deviations are rather smooth, but one may also observe many drastic jumps in obtained throughput. Note that the mobility of nodes mainly

Table 1. Simulation parameters.

Parameter	**Value**
Operational frequency, f_c	28 GHz
Antenna array	16×16 el. (planar array)
Channel model	3GPP InH
Emitted power	1 W
Receiver Sensitivity	-91 dBm
Fraction of floor covered by blockers, p_C	0.15
Number of fire crew members	$\{8, 10, 12, 14, 16\}$
Velocity of crew members	1 m/s
Mobility model of crew members	Random direction movement model
Number of simultaneously supported links	$\{2, 3, 4, \infty\}$
Number of iterations per simulation round	4800



(a) Connectivity trace

(b) Throughput trace

Fig. 3. Time-dependent behavior of connectivity and throughput processes.

causes the smooth deviations while quick jumps are associated with blockage events floor layout and dynamic blockage process.

Having observed the time-dependent behavior of the system, we now proceed analyzing stationary state metrics. We start with the time fraction that at least one node is disconnected from the network illustrated in Fig. 4 as a function of the number of nodes in a mesh network and the number of simultaneously supported links, M. Note that this metric can be considered as an integral measure of mesh network connectivity characterizing the fraction of time at least one node does not have access to the gateway. As one may observe, when the number of nodes increases, depending on the degree of multi-connectivity, the analyzed metric is characterized by principally different behavior. For $M = 2$ and $M = 3$ the time fraction that at least one node is disconnected increases as the number of nodes in a mesh increases. The rationale behind this behavior is straightforward: as the number of nodes increases the probability that at least one node finds itself in unfavorable position gets higher, and the degree of multi-connectivity is insufficient to overcome this. However, as the degree of multi-connectivity increases further, the effects of diversity start to dominate when the number of nodes increases. Thus, we may conclude that the multi-connectivity operation of mmWave systems may drastically increase mesh connectivity in indoor deployments. However, the number of simultaneously supported links might be rather high, resulting in significant control overhead.

We are now in a position to quantitatively characterize the mesh network connectivity in the stationary state. Fig. 5 shows the mean number of disconnected nodes as a function of the number of nodes in a mesh and the degree of multi-connectivity. Similarly to Fig. 4 we may observe that the degree of multi-connectivity of $M = 2$ and $M = 3$ does not allow for a network to scale appropriately as the mean number of disconnected nodes starts to increase. However, increasing M further to 4 allows this metric to stay well below one node. One may also notice that if we do not limit the number of simultaneously supported links, the mean number of nodes actually decreases as the number of disconnected nodes in a mesh network increases.

Finally, in Fig. 6 we study the mean per-node throughput as a function of the number of nodes in a mesh and the degree of multi-connectivity. As one may observe, this metric exhibits qualitatively similar behavior for all values of M. Comparing the mean throughput values corresponding to $M = \infty$ and $M = 2$ one of the critical observations of the illustrated results is that the system operates in blockage-limited conditional for realistic values of M. Indeed, the gains of imposing no restrictions on the degree of multi-connectivity the per-node

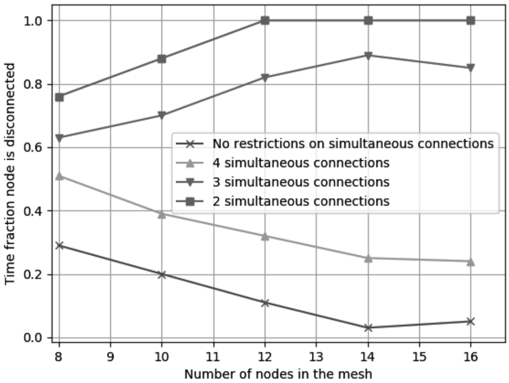


Fig. 4. Time fraction that at least one node disconnected.

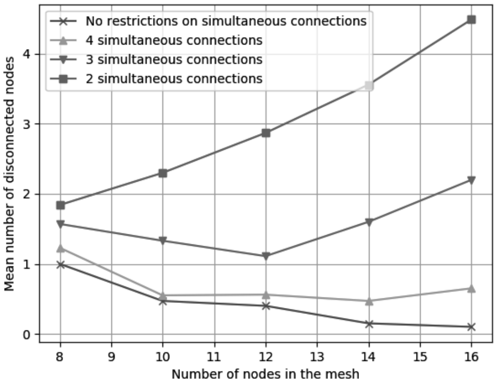


Fig. 5. Mean number of disconnected nodes.

throughput is $3 - 4$ times higher compared to $M = 2$. It is also important that increasing the degree of multi-connectivity, the system approaches this regime rather slowly, e.g., the mean per-node throughput obtained with $M = 4$ is still approximately half of $M = \infty$. This behavior of principally different from that reported for outdoor scenarios in, e.g., [8, 7], where both capacity and outage probabilities improve exponentially with M.

5 Conclusion

Motivated by the need for on-demand high-throughput mesh networking for indoor PPDR use cases, such as a fire suppression mission, in this paper, we investigated the capabilities of mmWave technology for these types of applications. The developed model is based on system-level simulations of the realistic indoor floor deployment of mesh system with nodes mobility, 3GPP indoor propagation, dynamical blockages, and multi-connectivity operation. The metrics of interest in this study are related to network connectivity and throughput.

The simulation campaign revealed that individual node connectivity in an indoor environment is characterized by frequent short-lasting outage time intervals and associated jumps in node throughput even in the presence of multi-connectivity capabilities. Nevertheless, the multi-connectivity capability of end systems may significantly improve the overall network connectivity properties in terms of the fraction of time at least one node is disconnected and the mean number of disconnected nodes. However, the associated increase in per-node performance is less noticeable implying that indoor mmWave mesh deployments mainly operate in a blockage-rich environment which is different from outdoor

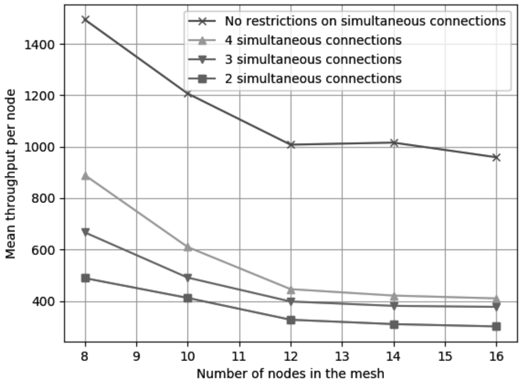


Fig. 6. Mean per-node throughput in a mesh network.

deployments, where multi-connectivity improves both connectivity and throughput performance exponentially [8, 7].

References

1. A. M. Al-samman, M. H. Azmi, and T. A. Rahman, “A survey of millimeter wave (mm-wave) communications for 5g: Channel measurement below and above 6 ghz,” in *Recent Trends in Data Science and Soft Computing* (F. Saeed, N. Gazem, F. Mohammed, and A. Busalim, eds.), (Cham), pp. 451–463, Springer International Publishing, 2019.
2. S. Singh, R. Mudumbai, and U. Madhow, “Interference analysis for highly directional 60-ghz mesh networks: The case for rethinking medium access control,” *IEEE/ACM Transactions on Networking (TON)*, vol. 19, no. 5, pp. 1513–1527, 2011.
3. M. Cheffena, “Industrial wireless communications over the millimeter wave spectrum: opportunities and challenges,” *IEEE Communications Magazine*, vol. 54, no. 9, pp. 66–72, 2016.
4. R. Humpleman and P. Watson, “Investigation of attenuation by rainfall at 60 ghz,” in *Proceedings of the Institution of Electrical Engineers*, vol. 125, pp. 85–91, IET, 1978.
5. 3GPP, “NR; Multi-connectivity; Overall description (Release 15),” 3GPP TS 37.340 V15.2.0, June 2018.
6. D. Moltchanov, A. Ometov, S. Andreev, and Y. Koucheryavy, “Upper bound on capacity of 5g mmwave cellular with multi-connectivity capabilities,” *Electronics Letters*, vol. 54, no. 11, pp. 724–726, 2018.
7. M. Gapeyenko, V. Petrov, D. Moltchanov, M. R. Akdeniz, S. Andreev, N. Himayat, and Y. Koucheryavy, “On the degree of multi-connectivity in 5g millimeter-wave cellular urban deployments,” *IEEE Transactions on Vehicular Technology*, vol. 68, no. 2, pp. 1973–1978, 2019.

8. M. Gerasimenko, D. Moltchanov, M. Gapeyenko, S. Andreev, and Y. Koucheryavy, "Capacity of multi-connectivity mmwave systems with dynamic blockage and directional antennas," *IEEE Transactions on Vehicular Technology*, 2019.
9. V. Petrov, D. Solomitckii, A. Samuylov, M. A. Lema, M. Gapeyenko, D. Moltchanov, S. Andreev, V. Naumov, K. Samouylov, M. Dohler, *et al.*, "Dynamic Multi-connectivity Performance in Ultra-dense Urban mmWave Deployments," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 9, pp. 2038–2055, 2017.
10. M. Polese, M. Giordani, M. Mezzavilla, S. Rangan, and M. Zorzi, "Improved Handover through Dual Connectivity in 5G mmWave Mobile Networks," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 9, pp. 2069–2084, 2017.
11. Y. Niu, Y. Li, D. Jin, L. Su, and A. V. Vasilakos, "A survey of millimeter wave communications (mmwave) for 5g: opportunities and challenges," *Wireless Networks*, vol. 21, no. 8, pp. 2657–2676, 2015.
12. A. Thornburg, T. Bai, and R. W. Heath Jr, "Performance analysis of outdoor mmwave ad hoc networks.," *IEEE Trans. Signal Processing*, vol. 64, no. 15, pp. 4065–4079, 2016.
13. J. Qiao, X. S. Shen, J. W. Mark, Q. Shen, Y. He, and L. Lei, "Enabling device-to-device communications in millimeter-wave 5g cellular networks," *IEEE Communications Magazine*, vol. 53, no. 1, pp. 209–215, 2015.
14. P. Nain, D. Towsley, B. Liu, and Z. Liu, "Properties of Random Direction Models," in *Proc. of 24th Annual Joint Conference of the IEEE Computer and Communications Societies*, vol. 3, pp. 1897–1907, IEEE, 2005.
15. V. Petrov, M. Komarov, D. Moltchanov, J. M. Jornet, and Y. Koucheryavy, "Interference and SINR in Millimeter Wave and Terahertz Communication Systems With Blocking and Directional Antennas," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1791–1808, 2017.
16. C. A. Balanis, *Antenna theory: analysis and design.* John wiley & sons, 2016.
17. V. Petrov, D. Moltchanov, P. Kustarev, J. M. Jornet, and Y. Koucheryavy, "On the use of integral geometry for interference modeling and analysis in wireless networks," *IEEE Communications Letters*, vol. 20, no. 12, pp. 2530–2533, 2016.
18. R. Vaughan, "Massively multi-robot simulation in stage," *Swarm intelligence*, vol. 2, no. 2-4, pp. 189–208, 2008.
19. B. Gerkey, R. T. Vaughan, and A. Howard, "The player/stage project: Tools for multi-robot and distributed sensor systems," in *Proceedings of the 11th international conference on advanced robotics*, vol. 1, pp. 317–323, 2003.
20. H. Perros, "Computer simulation techniques," *The definitive introduction. North Carolina State University*, 2009.

PUBLICATION

VI

Facilitating mmWave Mesh Reliability in PPDR Scenarios Utilizing Artificial Intelligence

R. Pirmagomedov, D. Moltchanov, A. Ometov, K. Muhammad, S. Andreev and Y. Koucheryavy

IEEE Access (2019)
DOI: 10.1109/ACCESS.2019.2958426

IEEE Access
Multidisciplinary ⁝ Rapid Review ⁝ Open Access Journal

Received October 28, 2019, accepted November 5, 2019, date of publication December 9, 2019

Digital Object Identifier 10.1109/ACCESS.2019.2958426

Facilitating mmWave Mesh Reliability in PPDR Scenarios Utilizing Artificial Intelligence

RUSTAM PIRMAGOMEDOV[1], DMITRI MOLTCHANOV[1], ALEKSANDR OMETOV[1], (Member, IEEE), KHAN MUHAMMAD[2], SERGEY ANDREEV[1], (Senior Member, IEEE), AND YEVGENI KOUCHERYAVY[1]

[1]Unit of Electrical Engineering, Tampere University, FI-33720 Tampere, Finland
[2]Department of Software, Sejong University, Seoul 05006, South Korea

Corresponding author: Rustam Pirmagomedov (rustam.pirmagomedov@tuni.fi)

This work was supported in part by the Academy of Finland (Projects PRISMA, WiFiUS, and RADIANT) as well as by the Project 5G-FORCE, and in part by the National Research Foundation of Korea (NRF) Grant funded by the Korean Government through the Ministry of Science, ICT and Future Planning (MSIT) under Grant 2019R1A2B5B01070067.

ABSTRACT The use of advanced AR/VR applications may benefit the efficiency of collaborative public protection and disaster relief (PPDR) missions by providing better situational awareness and deeper real-time immersion. The resultant bandwidth-hungry traffic calls for the use of capable millimeter-wave (mmWave) radio technologies, which are however susceptible to link blockage phenomena. The latter may significantly reduce the network reliability and thus degrade the performance of PPDR applications. Efficient mmWave-based mesh topologies need to, therefore, be constructed, which employ advanced multi-connectivity mechanisms to improve the levels of connectivity. This work conceptualizes predictive blockage avoidance by leveraging emerging artificial intelligence (AI) capabilities. In particular, AI-aided blockage prediction permits the mesh network to reconfigure itself by establishing alternative connections proactively, thus reducing the chances of a harmful link interruption. An illustrative scenario related to a fire suppression mission is then addressed by demonstrating that the proposed approach dramatically improves the connection reliability in dynamic mmWave-based deployments.

INDEX TERMS Mesh networks, millimeter wave communication, artificial intelligence (AI), wireless communication, public protection and disaster relief (PPDR).

I. INTRODUCTION

Wireless communication technologies are an essential enabler in Public Protection and Disaster Relief (PPDR) situations [1], [2]. They were historically utilized to provide sustainable voice communication services for public safety agencies [3]. Today, the cutting-edge PPDR applications include a variety of multimedia services [4] complemented with artificial intelligence (AI) capabilities [5]. This decisive transformation promises advanced situational and contextual awareness as well as enables event prediction and prevention in critical missions. The use of AI in PPDR contexts may lead to an upgrade of mission-critical communication to mission-critical assistance.

For its efficient operation, AI-based technology requires real-time information about both the problem and the context [6]. Computer vision systems and related video analytics can be efficiently employed to collect it. For instance, video information about a specific PPDR event may be obtained with on-body cameras of the rescue crew members (police officers, firefighters, ambulance doctors, etc.) [7], video surveillance cameras deployed across the area of the PPDR mission [8], or with the aid of robots (e.g., unmanned aerial and ground vehicles) [9], [10]. Since data mining from the video stream requires relatively powerful computation capabilities, it can be performed in a remote processing center or in a distributed fashion via edge/fog computing. Both approaches require high-throughput radio access networks.

The emerging millimeter-wave (mmWave) technologies, such as IEEE 802.11ad/ay [11] and 3GPP New Radio (NR) [12], [13], offer the throughputs on the level demanded by traffic-hungry PPDR applications [14]. However, their utilization in PPDR scenarios is hampered by a

The associate editor coordinating the review of this manuscript and approving it for publication was Derek Abbott.

number of challenges. First, there is uncertainty about the existing communication infrastructure, which can be partially or completely unavailable. Second, mmWave propagation is sensitive to atmospheric, weather, and other conditions, which may eventually cause severe link quality degradation. These adverse effects may drastically reduce the reliability of mmWave connectivity.

The PPDR applications are expected to operate in dangerous conditions, such as fire, smoke, gas, or water vapor [9], which may affect the propagation of the mmWave signal [15]. Furthermore, certain PPDR situations may inherently deteriorate the propagation conditions; for example, substances applied by firefighters may occlude the direct path between the communicating nodes. The use of mmWave *mesh* topologies can provide diverse paths between a source and a destination, and thus partially address these challenges by enabling proximate communication when the network infrastructure is unavailable.

The problem of link blockage has been identified as one of the most challenging for mmWave communications. Particularly, it affects connection reliability in both single-hop and multi-hop topologies. In recent literature, this issue is primarily tackled by employing reactive techniques along with standardized capabilities, such as multi-connectivity operation [16], when the association point is only changed when the current links experience outage conditions as the result of a blockage situation [17]–[20].

Hence, contemporary mmWave solutions rely on inherently reactive techniques to mitigate dynamic link blockage events [16], [20]. This approach may introduce harmful delays in data transmission, thus hampering the use of real-time PPDR applications, since it permits radio connections to become interrupted and then recover later. Alternatively, one may use *proactive* mechanisms by predicting blockage situations and taking action in advance [21]–[23]. Possible measures may include altering the trajectory of movement to avoid blockage or utilizing alternative data routes, e.g., via peers that establish backup links. To efficiently enable this functionality, one has to employ advanced prediction techniques.

In this work, we analyze the use of AI methods to enable uninterrupted mmWave mesh connectivity in PPDR scenarios. Consequently, we contribute a novel approach to mitigate dynamic link blockage in mmWave mesh systems. It utilizes AI-aided prediction of blockage situations and helps establish alternative connections via peer relays before the blockage has actually occurred. Numerical results reported in what follows demonstrate the feasibility of our proposal. Particularly, the outlined approach considerably reduces the fraction of time when at least one node of the mmWave mesh in question is disconnected from the rest.

The remainder of this paper is organized as follows. In Section II, we discuss the use of mmWave technologies in PPDR situations. In Section III, we review the use of AI in the context of mmWave mesh technologies for PPDR. Our illustrative scenario is then studied in Section IV. The conclusions are drawn in the last section.

II. MILLIMETER-WAVE TECHNOLOGIES FOR PPDR

In this section, we elaborate on the utilization of mmWave systems in mission-aware PPDR scenarios. Particularly, we discuss the technology aspects of mmWave communications and identify the key challenges of using mmWave radios for PPDR.

A. FEATURES OF MILLIMETER-WAVE COMMUNICATIONS

The recent standardization activities behind 5G NR and WiGig systems are aiming to enable novel technology layout for real-time heavy-traffic applications, such as ultra-high definition video streaming [24], augmented and virtual reality (AR/VR) broadcasting [25], and proximate gaming [26], [27]. These solutions adequately address the bandwidth demands by utilizing the more abundant mmWave spectrum, primarily in 28, 60, and 73 GHz bands [28]–[30].

Radio propagation properties at mmWave frequencies are fundamentally different as compared to microwave setups. This is primarily due to the effects of link blockage, inherent directionality, and complex multi-path propagation, where various obstacles may occlude, reflect, or scatter the narrower mmWave beams [31]. The latter poses numerous challenges related to communication reliability and service continuity that need to be resolved comprehensively [17], [32]. The use of mmWave communications in indoor environments further complicates propagation because of multiple obstacles (e.g., walls, furniture, people) [33]–[35], which lead to more complex and dynamic propagation. In addition to blockage caused by other objects, there are *self-blockages* where a person blocks own links [36], [37].

The use of mobile access points, such as cells-on-wheels (CoWs, [38], [39]) and aerial access points (AAPs, [40], [41]), may enhance the performance of mmWave access technologies by maintaining line-of-sight (LoS) communications for users who are currently blocked or outside the base station coverage. However, these solutions are featured by relatively long deployment times and require additional resources, such as maintenance expenses for an unmanned aerial system. Hence, the use of such access points may not be suitable in all contexts.

To alleviate the effects of blockage and improve the reliability of mmWave connections, 3GPP has recently outlined multi-connectivity features [16]. Accordingly, a device may establish links to multiple access points (APs) in its proximity and dynamically change the serving AP if the current link experiences a blockage. Such an approach yields a dramatic decrease in the outage probability levels [19]. In the absence of network infrastructure, this concept can be enabled via device-to-device (D2D) communications [42], [43]. Direct connectivity between user devices allows for establishing a mesh network topology [44], which expands the service area of the mmWave APs. D2D-based mesh topologies naturally offer multi-connectivity opportunities for the partnering

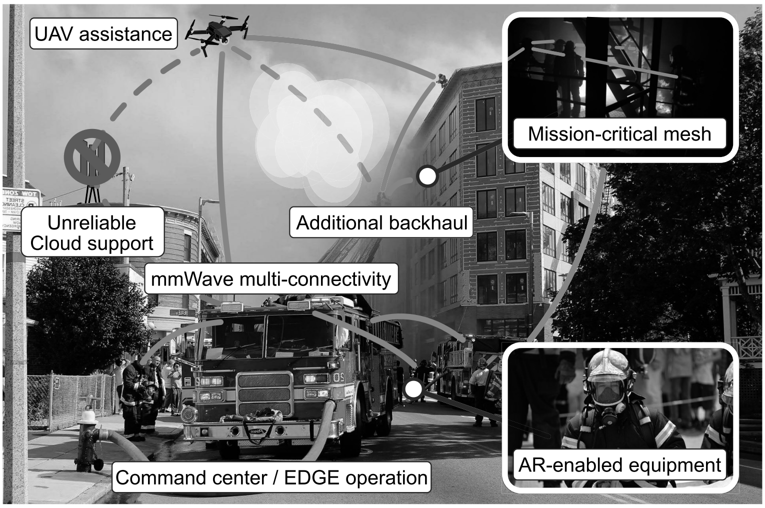


FIGURE 1. mmWave-enabled PPDR operation with a heavy-traffic application.

devices. Multi-connectivity naturally improves the resilience of a communications session to blockage because if a single link is occluded, the device can reroute its traffic via other connections.

In summary, it is known that link blockage may considerably limit the performance of mmWave-based systems. However, the situation can be notably improved with mobile APs arranged in a mesh topology.

B. PUBLIC PROTECTION SCENARIOS

Unexpected natural or human-made disasters require the safety agencies to always be prepared for PPDR situations in uncertain environments. The goal of PPDR missions is to reduce the risk to people's lives and property damage. Generally, a PPDR situation can be regarded as a multi-agent system comprising of intelligent entities, such as human rescuers and autonomous robots. Successful accomplishment of a PPDR mission hinges upon (i) continuous situational awareness, (ii) fast and reliable analysis of data and subsequent decision-making, and (iii) efficient coordination and cooperation between the rescue team members to eliminate task conflicts and duplication.

Innovative assisting technologies are extensively utilized to facilitate various PPDR missions. Computer vision systems have been proposed to obtain holistic information about the problem and its context for improved situational awareness [45]. These allow for monitoring the affected area in order to detect victims, assess damage, and evaluate hazards. The information analysis and decision-making processes can then be supported by AI-based applications, which may operate in a distributed or centralized manner [46]. Finally, dedicated radio technologies enable the coordination and cooperation inside the rescue team.

Previously, voice communications featured as the primary service supported by the PPDR systems. These are now expected to facilitate multiple additional applications that integrate voice, data, video, and image transmission as part of their multimedia capability to enable smooth coordination [47]. The latter requires more throughput and thus higher frequency bands where sufficient spectrum is available. Hence, mmWave communications technologies operating over the rich amounts of bandwidth can presently be considered as the key enabler for the emerging multimedia-ready PPDR applications.

C. APPLYING mmWave TECHNOLOGIES FOR PPDR

In addition to link throughput, there are further specific requirements pertaining to contemporary PPDR communications technologies [47]. Notably, those need to provide uninterrupted services irrespective of the current availability of the static network infrastructure. As long as cellular connectivity remains operational, PPDR applications can also exploit it.

Alternatively, other means of communication have to be deployed, e.g., in tunnels, inside buildings, or wherever the network infrastructure has (partially) collapsed [48]. These may rely on proximity-based D2D mesh operation, see Fig. 1, wherein the devices acting in close proximity (within the reach of a short-range wireless radio) may initialize direct links instead of utilizing network infrastructure. Therefore, the load on the cellular network may decrease, the operation without it might become possible, and better energy efficiency can be achieved. Therefore, the use of proximity-based direct communications is one of the promising solutions for beyond-5G connectivity.

Mesh-based mmWave solutions are expected to be utilized in collaborative PPDR missions to ensure robust connectivity between the rescue team members and enable traffic-hungry applications even when the cellular infrastructure is unavailable [49]. However, the use of advanced mmWave mesh topologies introduces additional challenges that relate to complex blockage dynamics [50]. Indeed, PPDR applications are expected to operate in hazardous environments, which may affect the propagation of mmWave beams. Furthermore, the context of the PPDR mission itself may deteriorate the radio conditions.

As a result, the mmWave mesh system reliability in PPDR situations depends on the mission type as well as on multiple environmental factors. Based on that, it is essential to not only provide high capacity during PPDR operation but also to develop reliable connectivity mechanisms and ensure resilience to various environmental conditions in mmWave-based mesh systems.

III. AI-AIDED MILLIMETER-WAVE MESH SYSTEMS

In this section, we discuss the application of AI methods in the context of mmWave mesh operation. We begin with a brief introduction and then review the use of AI in self-organizing mesh systems. Finally, we conceptualize AI-aided blockage prediction for a mmWave PPDR mesh.

A. DIVERSITY OF AI METHODS

AI techniques have entirely changed human life, from home appliances to automobiles, where every device or a piece of machinery is using some sort of an AI method. Since the beginning of AI evolution, researchers have introduced many AI practices including knowledge representation, expert systems, machine learning, neural networks, multi-agent systems, genetic algorithms, fuzzy logic, neuro-fuzzy, etc. However, based on the state-of-the-art achievements by machine learning and neural networks-based methods, most of today's AI techniques belong to either of the three major types: supervised, unsupervised, or reinforcement learning [51].

The former addresses the problems relying on labeled data (ground truth) or prior knowledge about the expected output. Typically, these tools are used in the context of classification and regression. In classification, the output is acquired in the form of labels or discrete values, whereas in regression, it is obtained as continuous values. Supervised learning algorithms include neural networks, convolutional neural networks, support vector machines, decision trees, naive Bayes, and linear regression. These techniques are widely applicable in many areas including object detection, pattern recognition, speech analysis, human activity recognition, and bio-informatics [52]–[54].

Unsupervised learning methods deal with the problems having unlabeled data, i.e., input with no corresponding output [55]. These automatically establish various patterns in the input data to learn its structure and make decisions based on similar patterns. Most of the corresponding algorithms are used for clustering, association rule learning, and data compression/generation in autoencoders. The typical unsupervised learning algorithms are K-means clustering and principal component analysis. Unsupervised learning tools are widely used for image segmentation, anomaly detection, and association mining.

Reinforcement learning employs reverse dynamics, such as reward and punishment to ''reinforce'' the knowledge for learning [56]. Unlike classical approaches, reinforcement learning exploits the concept of interacting with the environment based on trial and error. In reinforcement learning, the problem can be solved by performing two types of tasks, continuous and episodic. Continuous tasks persist (like forex/stock trading), while episodic tasks have the starting and ending points, which delimit an episode (like playing a game to complete a mission and move to the next level). The well-known algorithms of reinforcement learning are Q-Learning and State-Action-Reward-State-Action (SARSA, [57]). The reinforcement learning algorithms are widely utilized in robotics, web system configuration, advertising, and gaming.

B. AI IN SELF-ORGANIZING NETWORKS

''Brains exist because of the distribution of resources necessary for survival and the hazards that threaten survival vary in space and time'' [58]. This statement is equally applicable to AI used in self-organizing networks since the very utilization of AI aims at efficient management of resources and avoidance of hazards. Here, the role of resources is featured by connectivity and throughput, whereas blockage, interference, and technological incompatibility between the nodes of a mesh can be interpreted as hazards.

The three major sub-functional groups of AI for the emerging mesh networks are self-configuration, self-optimization, and self-healing [59]. The former is required to enable network association simplicity regardless of the employed radio interface or device capabilities. During the configuration stage, the network needs to invoke an authentication procedure and set up the radio interfaces of its nodes, e.g., transmit power, data, and control plane protocols. In the context of self-configuration, the AI can be used for recognizing new users, configuring wireless interfaces, predicting events when the current network state changes, etc.

Mesh networks are highly dynamic systems; hence, their management has to be adaptive, enabled by continuous self-optimization. This includes monitoring of the network state and subsequent adjustment of the network and interface parameters to reach high efficiency of resource utilization. Self-optimization covers a number of aspects including power efficiency, mobility of users, quality of links, and traffic dynamics. It may be empowered by the AI methods, which are utilized for the prediction of user mobility by choosing reliable connections between the nodes of a mesh, predicting link quality and traffic flow structure based on previous experience, and tracing network users. Due to AI, the network may reduce the risk of failures and wastage of resources. As a result, the quality of user experience becomes higher. Hence, self-optimization aims to enable low latency, high bandwidth,

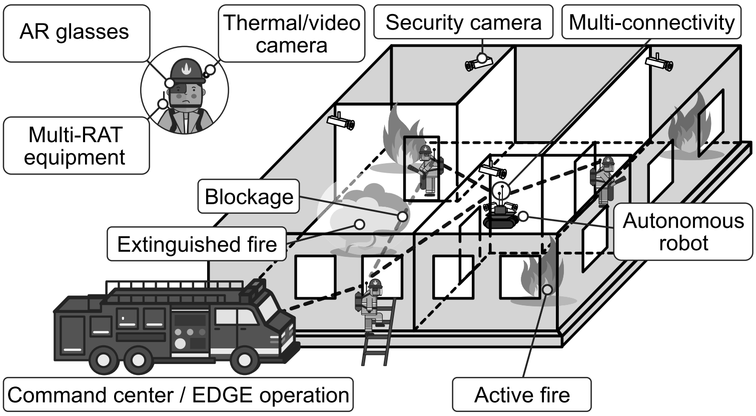


FIGURE 2. Multi-connectivity mmWave mesh setup of interest.

and better connectivity within a mesh network with higher degrees of temporal and spatial variation in user demands.

Self-healing of a network is related to recovering its functionality after failures. With respect to mesh systems, a failure can be defined as a state of the network where communications between two or more nodes is interrupted. A significant proportion of such failures is caused by a lack of connectivity between the devices. The AI methods used for self-healing aim to mitigate the failure events and automatically recover the network performance [59]. They may include but are not limited to automatic failure detection and diagnostics, reconfiguration of nodes in real-time (e.g., increased transmit power to extend coverage of certain nodes), and rerouting. In severe cases (e.g., an essential link is faulty), the original network can split into two or more isolated parts.

C. USING AI FOR BLOCKAGE PREDICTION

Dissimilar static and dynamic objects (e.g., people, buildings, vehicles) may cause link blockage in mmWave mesh systems, which are thus characterized by a high degree of temporal and spatial variability. The objects in question may not only occlude the direct path but also block the reflected paths by disrupting communications between the nodes of a mesh for prohibitive periods of time. As a result, the performance of PPDR applications utilizing mmWave mesh capabilities may degrade considerably.

Recent developments in AI techniques are capable of anticipating the blockage situations in mmWave mesh networks. Particularly, AI-based algorithms may employ computer vision and sensory data to acquire the indicators of an imminent blockage. For example, AI-aided systems can predict link occlusions caused by people or static objects (such as trees, buildings, and landscape) by utilizing the data about (i) their trajectory and speed, (ii) trajectory and speed of the mmWave mesh nodes, and (iii) location of static obstacles.

The blockage prediction systems are potentially able to improve the sustainability of a mmWave mesh layout. If the latter is made aware of a probable blockage, the loss of the radio connectivity can be prevented by relocating the nodes or resorting to D2D technologies, such as peer relaying. Moreover, reliance upon blockage prediction mechanisms potentially requires fewer resources as compared to the use of assisting technologies, such as COWs and AAPs. Hence, AI-enabled blockage prediction can become an attractive solution for improving communications reliability in mmWave mesh systems.

IV. AN ILLUSTRATIVE SCENARIO

In this section, we consider a fire suppression mission as an illustrative example to assess the gains from the use of the AI-aided blockage prediction in mmWave mesh systems.

A. FIRE SUPPRESSION MISSIONS

In the addressed scenario, we assume that a fire spreads dynamically in a particular area of interest, while the involved firefighters lack awareness about the spots of fire across this area, see Fig. 2.

To enhance the efficiency of a fire suppression mission, the collaborating team members may employ AR-based applications [60] and advanced sensory equipment, which require high throughput and network availability to support effective teamwork. The team is also supplemented by autonomous robots aiming to improve the probability of mission success. The said devices utilize multiple cameras and sensors to detect fire and determine their appropriate locations for serving as relay nodes for communicating with a potentially blocked device (if such locations exist), and move in the selected direction. The media-related equipment relies on 3D HD 360° video streaming, which requires approximately 100 Mbit/s of bandwidth per user [61] for the upload link (toward the processing server).

Moreover, about 5 Mbit/s may be demanded by the advanced sensory systems, including on-body health monitoring devices, sensitive smoke analysis sensors (for recognizing which materials are burning), thermal sensors, etc. Additionally, an AR-based assisting application may require up to 15 Mbit/s, e.g., for building navigation, environmental awareness, and command center notifications. In total, the utilized applications call for about 120 Mbit/s of bandwidth per one user. These demands are expected to be satisfied by a mmWave proximity-based mesh system. The latter is maintained between all of the participants of the firefighting team. The information transfer between the remote nodes and the gateway is multi-hop. To improve the reliability of this network, each participant supports multi-connectivity of M simultaneous links to its neighbors, referred to as the ''degree of multi-connectivity''. Further, if/when all M links are blocked, the network management layer may employ robot-based relays to establish alternative data routes.

A characteristic feature of the considered scenario is dynamic link blockage caused by water vapor from the fire extinguishing process. Recent AI methods may detect the fire by using video cameras, e.g., a video surveillance system deployed in the area and wearable video cameras of the firefighting team members. Knowing the exact location of the fire, the considered system becomes more aware of the spots where connectivity disruption chances are high. Using this information, the mesh network can improve its reliability by establishing an alternative connection proactively.

B. AI FOR DYNAMIC BLOCKAGE DETECTION

Accidents involving fire directed the attention of researchers to the development of new fire detection systems [62]. Presently, these follow either of the two general approaches: traditional and vision-based detection. Traditional fire detection systems utilize sensors, which rely upon temperature measurements, particle sampling, smoke analysis, and relative humidity sampling [63]. However, these sensors are mostly applicable for indoor environments, and remain unable to provide required details about the fire (e.g., burning degree, location, size). Vision-based systems utilize computer vision techniques and can overcome the limitations of the traditional systems [64], [65].

Recently, vision-based systems attracted significant research attention in the field of early fire detection due to their efficient response. These systems are attractive due to various advantages including (i) larger covered regions, (ii) lower costs, (iii) detection of fire without visiting the scene, (iv) providing the fire details such as location, burning degree, and size. Due to these features, vision-based systems may significantly enhance the efficiency of the traditional fire alarm applications.

The vision-based systems rely on static or adaptive (learned) methods for fire recognition purposes. The methods belonging to the first category use color and shape features for detecting the flame on an image (e.g., RGB, HIS, YUV, YUC, and YCbCr models). The main drawback of these methods is in their high false alarm rate [66]. Several tools based on motion features were developed to cope with this issue. However, these solutions are limited to shorter distances. Adaptive methods rely on convolutional neural networks (CNNs) for efficient fire detection [67]. The CNN-based approach enables fire detection over longer distances and with higher accuracy.

CNN is one of the essential types of neural networks initially designed for 2D image data, but presently its variants can also handle 1D and 3D data. A CNN is typically composed of convolutional, pooling, activation, and fully connected layers that are stacked in a hierarchical way. The convolutional and fully connected layers contain a number of kernels that are also known as neurons or trainable parameters, while the pooling and activation layers are functions without trainable parameters [68]. The parameters of these layers are learned via backpropagation techniques over numerous iterations to fit a particular task.

The convolution is a linear operation, which convolves a kernel over the entire image to extract the needed patterns from it. The pooling layer of a CNN is responsible for reducing the dimensionality of features. The success of the CNNs is not only in the field of object detection and image classification, but also in more complex problems, such as smoke and fire scene analysis [69]–[71], image, and video retrieval, medical image analysis, action, and activity recognition [72], scene parsing, and movie analysis [73]. Over the past few years, CNN-based methods became popular for feature extraction from videos as well as the image data. Moreover, the feature extraction techniques confirmed that the initial layers of a CNN may extract local image features, while its deeper layers provide a global representation of the image data.

In this paper, we focus on the CNNs that demonstrate state-of-the-art performance in image classification and other computer vision tasks. CNNs are deep learning frameworks that are inspired by the mechanism of visual perception of living creatures [74]. Their application in fire detection systems will substantially improve the detection accuracy, which will eventually minimize fire damage while reducing the ecological and social consequences. However, a major concern related to CNN-based fire detection systems is their implementation in the real-world surveillance networks due to the high memory and computation requirements for inference.

We further advocate the use of proactive approaches to avoid blockage situations in PPDR environments by assuming that AI is utilized to detect fire locations and provide information about a potential blockage situation that may occur in the future. When the fire location is detected, the multi-connectivity mmWave functionality is employed to avoid link blockage. Two potential situations are considered. If there are other connections available at a node whose link is going to be occluded soon, the traffic is rerouted via these alternative connections. If the node in question does not have backup connections to other nodes, a robot (if there is one available) is steered to establish a backup connection for the considered node. Otherwise, if the link quality is deteriorated due to blockage, the subject node becomes disconnected from the network. We assess these options below.

To this aim, we conduct a performance evaluation campaign based on two datasets. The first one comprises of a relatively small number of 226 photos, where 119 are with fire and 107 are without [74]. The second one is more informative and corresponds to 31 videos captured in both indoor and outdoor environments, where 14 videos contain fire and 17 videos belong to the non-fire class [70]. These sets were selected specifically with respect to two scenarios: (i) low-quality connection (the first set where the frames are delivered to the CNN with low rate of around 2 FPS assuming poor link quality), and (ii) high-throughput connection potentially provided by the mmWave links (the second set where FPS equals 25 for any resolution).

We further apply our pre-trained GoogleNet algorithm to both sets and calculate the false alarm rate (FAR) as well as the accuracy for both datasets under different resolution constraints (from 640 × 480 to 4K quality). Our results indicate that for the small and infrequent frame rate of the first dataset the accuracy is kept approximately at the level of 89%, while the FAR value fluctuates around 18% even when utilizing the CNN. When the overall system is operating with higher FPS and/or resolution, e.g., utilizing mmWave connections, the accuracy reaches 98.5% and FAR drops to near zero.

C. METHODOLOGY AND SIMULATOR DESCRIPTION

Our approach is based on a computationally efficient CNN implementation inspired by GoogleNet architecture, with its reasonable computational complexity and suitability for the intended problem as compared to other computationally expensive networks, such as AlexNet. It is utilized for fire and blockage detection, localization, and semantic understanding of the scene of the fire. This solution is based on a paradigm that classifies the input video frames into their respective class, i.e., ''Fire'' and ''Non-Fire''.

For the classification of videos, we employ a pre-trained GoogleNet architecture with further modifications according to our target problem. A simplified algorithm for fire detection utilized in the proposed solution is shown in Fig. 3. There are several reasons behind preferring this option for the detection of fire in our illustrative use case. The first one is in its high performance during fire detection. The second reason is the small size of the model, which allows for deploying the system on resource-constrained edge/fog devices. Finally, the proposed solution outperforms other state-of-the-art CNN models and fire detection methods in terms of its FAR rate and accuracy [70].

The proposed system includes two network overlays that function cooperatively. The first one utilizes mmWave radio technology for enabling high throughput among the users, which is required for the heavy traffic of media-centric applications. The second overlay relies upon a long-distance wireless technology (IEEE 802.11ah nicknamed Wi-Fi HaLow), which provides reliable albeit low throughput connections among all the nodes of a network [75]. These low throughput connections carry signaling for managing the mmWave mesh operation.

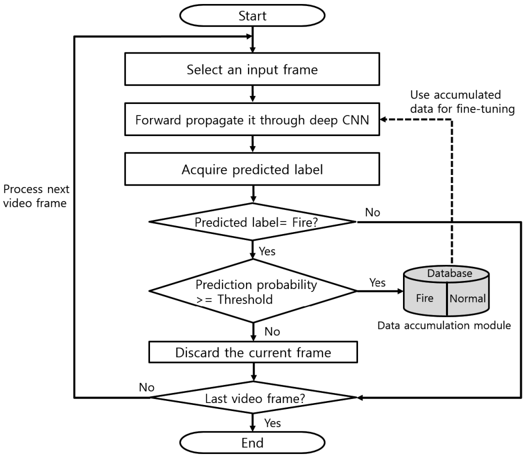


FIGURE 3. Simplified algorithm for efficient fire detection.

The proposed system operation comprises of continuously repeating cycles as shown in Fig. 4. Repeating the cycle allows for timely updates of the information about the mmWave mesh status. The frequency of updates depends on the operation dynamics as regulated by the command center. Such dynamics includes the number and density of nodes in the network, the intensity of blockage situations, etc. Scenarios with higher levels of dynamics require a higher frequency of updates. Every update cycle starts by determining the current topology of the mmWave mesh.

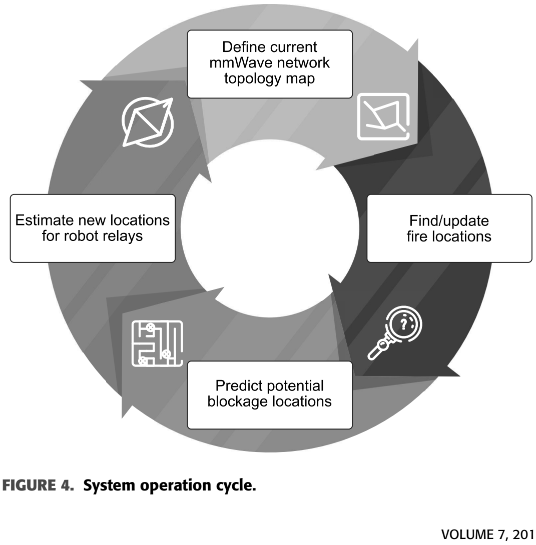

Define current mmWave network topology map
Find/update fire locations
Predict potential blockage locations
Estimate new locations for robot relays

FIGURE 4. System operation cycle.

For this purpose, every node sends its coordinates and mmWave link-state advertisements (LSAs) via the long-range HaLow connections. Using LSAs, a processing unit located in the command center acquires the current mmWave mesh topology. At the same time, the command center recognizes the fire zones and updates their locations by utilizing the available media and sensory information obtained from the fire suppression team members via the mmWave mesh. At the next step, the system provides a mapping of the fire locations, the building plan, and the mesh topology to predict the potential blockages as illustrated in Fig. 5. Finally, using the information about the potential blockages, the system estimates where to move the robot relays for reducing the risk of disconnecting mesh nodes from the gateway.

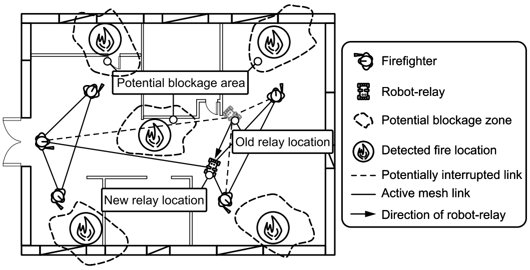


FIGURE 5. Blockage prediction scenario: top view.

For the numerical assessment, we employ our advanced "large-scale" system-level simulator (SLS), which takes into account all of the relevant details of the mmWave system operation and has been thoroughly calibrated in our past publications [20], [76]–[81]. This SLS tool is capable of mimicking large-scale environments together with the underlying wireless technologies, such as LTE, WiFi, and mmWave-based RATs: IEEE 802.11ad and 3GPP NR.

The tool is based on a flexible event-driven architecture, which allows decreasing the computation time in the low-load scenarios. For all the considered technologies, PHY and MAC layers are implemented in detail, based on the appropriate IEEE and 3GPP specifications, whereas the higher layers are generally simplified to abstract away the traffic models represented by analytical approximations. Regarding the environment generation, our SLS tool supports 3D geographical models, which take into account time- and location-based interference, antenna configurations, and UE mobility models.

With respect to mmWave communications, the SLS implements the propagation models specified by 3GPP in [82], with dynamic blockage extensions from [80], [81] and further advanced functions, such as multi-connectivity [16]. The D2D and multi-hop functionalities in mmWave bands are currently under specification by both IEEE and 3GPP for WiGig and NR technologies (under IEEE 802.11ay and 5G NR standards, respectively). Hence, to assess the performance of multi-hop relaying solutions, we rely on the current work-in-progress 3GPP documents (R1-1812199, R1-1812982, and R1-1813418), along with the NR relaying capabilities discussed in TR 38.874 – initially planned for Rel. 15 and now continued with the focus on Rel. 16. An open-source version of our SLS is made available at [83].

D. RELIABILITY ASSESSMENT OF AI-AIDED MESH

We consider an area of 100 × 100 m with 10 fire crew members, each equipped with mmWave-based radios for communications and cameras for video streaming. The crew is accompanied by K autonomous robots also supplied with mmWave radios and cameras suitable for 4K video transmission. A mesh network is constructed between the firefighting crew members to enable uninterrupted video delivery to the remote cloud. In order to capture the dynamics of the fire suppression process, we employ a spatially-temporal Poisson process [84] that is built on three parameters: (i) the temporal intensity of fire locations, (ii) the mean duration of evaporated water after the fire suppression, and (iii) the radius of the evaporated water. As confirmed by the measurements in [85], attenuation caused by water is sufficient to occlude the propagation of mmWave signals.

To improve the system performance in dynamic blockage-prone environments, devices carried by the crew members implement multi-connectivity functionality; hence, they establish multiple links to the neighboring nodes and switch over to non-blocked connections whenever the current link is disrupted. The considered mmWave technology is IEEE 802.11ay operating in the 60 GHz band [11]. To approximate the coverage of a single mmWave radio, we utilize the InH propagation model and 0.2 W of transmit power [82]. Other system parameters are summarized in Table 1. We specifically note that as compared to

TABLE 1. Default system parameters.

Parameter	Value
Operating frequency, f_c	28 GHz
Antenna array	16 × 16 el. (planar array)
Channel model	3GPP InH
Transmit power	0.2 W
Area of interest	100×100 m
Number of static blockers	10
Radius of static blockers	5 m
Attenuation by static blockers	40 dB
Temporal intensity of fire locations	0.1 events/s
Mean duration of suppressed fire location	120 s
Radius of suppressed fire location	3 m
Attenuation by suppressed fire location	20 dB
Number of firefighting crew members	$\{10, 20, 50\}$
Moving speed of crew members	3 m/s
Mobility model of crew members	Random direction model
Number of autonomous robots	$\{1, 2 \ldots 10\}$
Number of simultaneously supported links	$\{1, 2 \ldots 10\}$
Number of executions per setup	10e5

microwave technology, user devices greatly benefit from operating in the mmWave band. In particular, having 16 × 16 linear antenna arrays leads to approximately 6.5° half-power beamwidth (HBPW) at both the transmit and the receive ends.

Further, we evaluate and compare two representative scenarios. In the *baseline* setup, no fire detection assistance is provided, and the firefighting crew members move randomly across the area of interest at the speed of v in their search of a fire location for suppression. Note that the baseline scenario includes the state-of-the-art blockage avoidance techniques, such as multi-connectivity [86] – the devices are allowed to support multiple links to improve network connectivity. In the *AI-aided* scenario, fire detection capabilities are used to guide the firefighters toward the actual fire locations. To improve the levels of connectivity, in addition to supporting multiple simultaneous links, devices may rely on a fleet of autonomous robots moving in the environment with the speed of v_R. The AI algorithms not only allow to detect the locations of fire but also employ cameras for predicting the blockage situations. If a blockage is expected to occur, the respective crew member connects to an autonomous robot, whose aim is to support additional relay links to maintain uninterrupted connectivity.

For the *baseline* and the *AI-aided* scenarios, we consider the following performance metrics related to mmWave system reliability in dynamic blockage environments: (i) fraction of time when a randomly chosen node in the mesh is disconnected, (ii) probability that a certain number of nodes are disconnected at a randomly chosen instant of time, (iii) intensity of node disconnections from the mesh, and (iv) data rate at the access gateway. Note that these parameters depend on the considered setup, e.g., the number of nodes, the degree of multi-connectivity, the use of AI-based proactive blockage detection, and the number of autonomous robots. A system-level performance assessment is then conducted within our simulation environment that integrates the main functionality of the mmWave system and extends it to support the multi-connectivity operation. The core of this modeler is based on a discrete-even simulation framework. The statistics were collected via the method of replications in the steady-state period. The beginning of this period was determined with an exponentially-weighted moving average (EWMA) filter [87].

First, in Fig. 6 we study the fraction of time an arbitrarily chosen node is disconnected from the network for both scenarios of interest as a function of the temporal intensity of fire locations. Analyzing these results, one may observe that the spatial diversity made available via multi-connectivity allows to drastically reduce the parameter under investigation over the entire range of considered intensities. However, the use of AI for fire detection and autonomous robot relaying results in even more profound positive effects. Particularly, the system with the multi-connectivity degree of 3 and no AI support performs worse than the system with AI assistance and $M = 1$. The use of multi-connectivity and AI together allows to dramatically improve the performance by efficiently avoiding blockage even in extremely dynamic conditions where the fire location intensity reaches significant values of $0.4 - 0.5$ events/s.

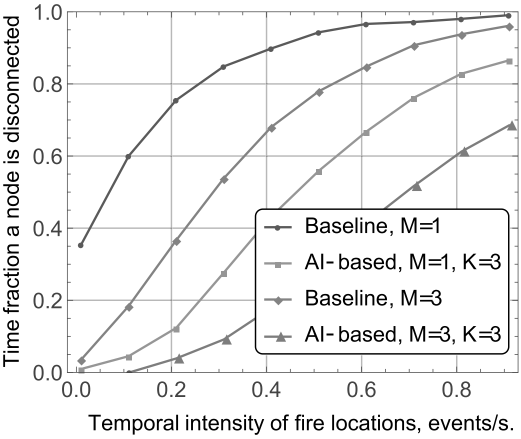


FIGURE 6. Fraction of disconnect time from a mesh.

Another parameter of interest that characterizes the reliability of a mesh is the number of disconnected nodes at an arbitrarily chosen instant of time. Recall that this value characterizes the ability of the network to support the ongoing mission. Here, the more nodes are disconnected, the less information is available for coordinating a mission, which may eventually lead to additional nodes disconnecting from the network. Fig. 7 illustrates this behavior for the degree of multi-connectivity $M = 3$ and the temporal intensity of fire locations of 0.1 events/s as a function of the mean water vapor duration in the suppressed fire locations for both scenarios. As one may observe, the effect of AI assistance is visible across the entire considered range of the mean durations of the suppressed fire location. For $M = 3$ and $K = 6$,

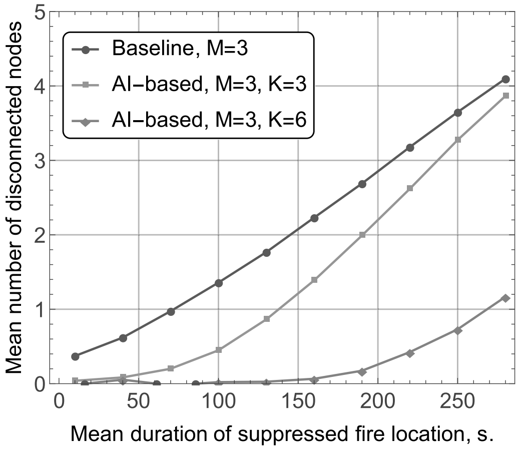


FIGURE 7. Mean number of disconnected nodes.

the mean number of disconnected nodes remains close to zero for up to the duration of approximately 150 s. However, as the mean duration increases, the system is no longer capable of maintaining uninterrupted mesh connectivity even with the autonomous robot relays, and the mean number of disconnected nodes increases.

We further characterize time-dependent performance – the intensity of node disconnections from a mesh in Fig. 8 as a function of the degree of multi-connectivity, M, the number of autonomous robot relays, K, and the mean water vapor duration in suppressed fire locations, $1/\theta$. As we learn, the response of the system is qualitatively similar for both the baseline and the AI-aided scenarios. An initial increase in the intensity of node disconnections is explained by the fact that the temporal intensity of node locations adds to the blockage dynamics as moving firefighters begin to experience link interruptions more frequently. However, when the intensity of dynamic blockers exceeds a certain value that generally depends on the type of the scenario and the selected system parameters, the intensity of node disconnections begins to decrease. The reason is that in this regime the number of static and dynamic blockers becomes so high that individual blockage periods merge into the longer ones, thus forcing a node to spend more time in the disconnected state, see Fig. 6.

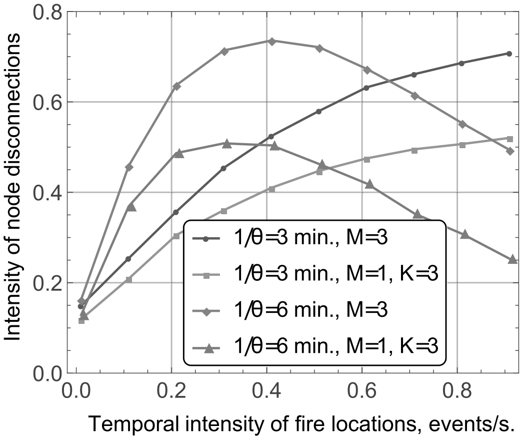


FIGURE 8. Intensity of node disconnections from a mesh.

Finally, in Fig. 9 we assess the maximum aggregate data rate of the mesh network at the access gateway for $M = 3$ and $K = 3$. Observe that an upper bound on the radio access level data rate is provided by a zero intensity of fire locations, which results in approximately 5.4, 3.7, and 2.0 Gbps for $N = 50$, $N = 20$, and $N = 10$, respectively. These values can be used for choosing the appropriate mmWave technology in the overlay. Particularly, the IEEE 802.11ad solution theoretically provides up to 6 Gbps and may thus support the fire suppression crews of up to 50 persons. For higher values of N, the emerging IEEE 802.11ay technology can be preferred.

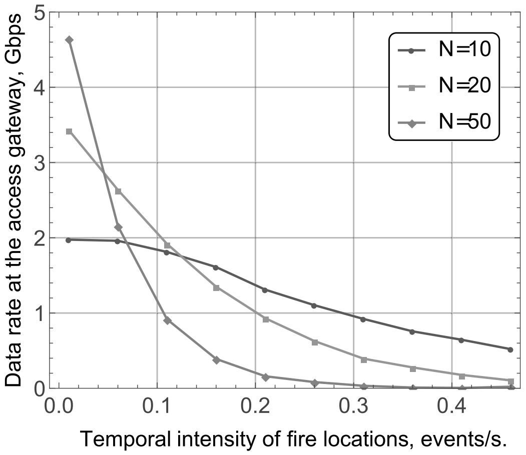


FIGURE 9. Data rate at the access gateway.

One may also notice that for all the values of N, the data rate decreases as the temporal intensity of fire locations grows. The reason is that the time fraction when a node is disconnected from the network rapidly increases, as shown in Fig. 6, which reduces the number of nodes delivering their traffic to the gateway. This process is characterized by an avalanche-like trend, since a lower fraction of the delivered data leads to fewer fire locations detected, which, in its turn, increases the fraction of disconnect time. Hence, for all the values of N, the data rate drops to zero. In this regime, the system no longer maintains its intended functionality, and additional robot relays are needed for improved mesh operation.

V. CONCLUSION

The contemporary PPDR requirements go far beyond conventional voice services. The use of advanced applications like AR/VR may drastically improve the efficiency of collaborative PPDR missions by providing real-time 3D information about the environmental conditions. These new requirements naturally call for the use of mmWave radio technologies that offer extensive bandwidths at the air interface. To maintain uninterrupted connectivity of mmWave-based mesh layouts in challenging environments with both natural and artificial obstacles, one has to rely upon advanced techniques to intelligently predict the potential blockage situations and effectively mitigate them in real-time.

In this work, we considered the use of AI-aided techniques to improve the performance of the mmWave-based mesh systems in the representative firefighting scenarios. We employed computer vision to detect the areas with potential blockage situations and further predict the chances of losing connectivity in dynamic self-organizing mmWave deployments. In this case, either the user itself or a remote control center may take preventive measures to avoid potential node disconnects by, e.g., utilizing proximate robot-based relaying. Our numerical results indicate that the proposed approach significantly enhances the reliability of the mmWave mesh operation by substantially reducing the fraction of disconnect time.

The results of this study are relevant beyond the considered fire suppression scenarios. Particularly, AI-based solutions can be utilized for predicting blockage situations in 5G NR systems by improving link reliability and thus augmenting session continuity.

REFERENCES

[1] S. A. R. Naqvi, S. A. Hassan, H. Pervaiz, and Q. Ni, "Drone-aided communication as a key enabler for 5G and resilient public safety networks," *IEEE Commun. Mag.*, vol. 56, no. 1, pp. 36–42, Jan. 2018.

[2] M. Mezzavilla, M. Polese, A. Zanella, A. Dhananjay, S. Rangan, C. Kessler, T. S. Rappaport, and M. Zorzi, "Public safety communications above 6 GHz: Challenges and opportunities," *IEEE Access*, vol. 6, pp. 316–329, 2017.

[3] R. Fantacci, F. Gei, D. Marabissi, and L. Micciullo, "Public safety networks evolution toward broadband: Sharing infrastructures and spectrum with commercial systems," *IEEE Commun. Mag.*, vol. 54, no. 4, pp. 24–30, Apr. 2016.

[4] A. Jarwan, A. Sabbah, M. Ibnkahla, and O. Issa, "LTE-based public safety networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 2, pp. 1165–1187, 2nd Quart., 2019.

[5] P. Borges, B. Sousa, L. Ferreira, F. B. Saghezchi, G. Mantas, J. Ribeiro, J. Rodriguez, L. Cordeiro, and P. Simoes, "Towards a hybrid intrusion detection system for android-based PPDR terminals," in *Proc. IFIP/IEEE Symp. Integr. Netw. Service Manage. (IM)*, May 2017, pp. 1034–1039.

[6] M. Chen, U. Challita, W. Saad, C. Yin, and M. Debbah, "Artificial neural networks-based machine learning for wireless networks: A tutorial," 2017, *arXiv:1710.02913*. [Online]. Available: https://arxiv.org/abs/1710.02913

[7] A. Ometov, E. Sopin, I. Gudkova, S. Andreev, Y. V. Gaidamaka, and Y. Koucheryavy, "Modeling unreliable operation of mmWave-based data sessions in mission-critical PPDR services," *IEEE Access*, vol. 5, pp. 20536–20544, 2017.

[8] A. Ometov, V. Petrov, S. Bezzateev, S. Andreev, Y. Koucheryavy, and M. Gerla, "Challenges of multi-factor authentication for securing advanced IoT applications," *IEEE Netw.*, vol. 33, no. 2, pp. 82–88, Mar./Apr. 2019.

[9] M. Półka, S. Ptak, and Ł. Kuziora, "The use of UAV's for search and rescue operations," *Procedia Eng.*, vol. 192, pp. 748–752, Jun. 2017.

[10] T. D. Dinh, R. Pirmagomedov, V. D. Pham, A. A. Ahmed, R. Kirichek, R. Glushakov, and A. Vladyko, "Unmanned aerial system–assisted wilderness search and rescue mission," *Int. J. Distrib. Sensor Netw.*, vol. 15, no. 6, 2019, Art. no. 1550147719850719.

[11] Y. Ghasempour, C. R. da Silva, C. Cordeiro, and E. W. Knightly, "IEEE 802.11 ay: Next-generation 60 GHz communication for 100 Gb/s Wi-Fi," *IEEE Commun. Mag.*, vol. 55, no. 12, pp. 186–192, Dec. 2017.

[12] S. Parkvall, E. Dahlman, A. Furuskar, and M. Frenne, "NR: The new 5G radio access technology," *IEEE Commun. Standards Mag.*, vol. 1, no. 4, pp. 24–30, Dec. 2017.

[13] S.-Y. Lien, S.-L. Shieh, Y. Huang, B. Su, Y.-L. Hsu, and H.-Y. Wei, "5G new radio: Waveform, frame structure, multiple access, and initial access," *IEEE Commun. Mag.*, vol. 55, no. 6, pp. 64–71, Jun. 2017.

[14] W. Guo, M. Fuentes, L. Christodoulou, and B. Mouhouche, "Roads to multimedia broadcast multicast services in 5G new radio," in *Proc. IEEE Int. Symp. Broadband Multimedia Syst. Broadcast. (BMSB)*, Jun. 2018, pp. 1–5.

[15] A. Ometov, D. Moltchanov, M. Komarov, S. V. Volvenko, and Y. Koucheryavy, "Packet level performance assessment of mmWave backhauling technology for 3GPP NR systems," *IEEE Access*, vol. 7, pp. 9860–9871, 2019.

[16] *NR; Multi-Connectivity; Overall Description: Stage 2*, document V15.2.0, 3GPP, Jul. 2019.

[17] V. Petrov, D. Solomitckii, A. Samuylov, M. A. Lema, M. Gapeyenko, D. Moltchanov, S. Andreev, V. Naumov, K. Samouylov, M. Dohler, and Y. Koucheryavy, "Dynamic multi-connectivity performance in ultra-dense urban mmWave deployments," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 2038–2055, Sep. 2017.

[18] M. Polese, M. Giordani, M. Mezzavilla, S. Rangan, and M. Zorzi, "Improved handover through dual connectivity in 5G mmWave mobile networks," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 2069–2084, Sep. 2017.

[19] M. Gapeyenko, V. Petrov, D. Moltchanov, M. R. Akdeniz, S. Andreev, N. Himayat, and Y. Koucheryavy, "On the degree of multi-connectivity in 5G millimeter-wave cellular urban deployments," *IEEE Trans. Veh. Technol.*, vol. 68, no. 2, pp. 1973–1978, Feb. 2019.

[20] M. Gerasimenko, D. Moltchanov, M. Gapeyenko, S. Andreev, and Y. Koucheryavy, "Capacity of multiconnectivity mmWave systems with dynamic blockage and directional antennas," *IEEE Trans. Veh. Technol.*, vol. 68, no. 4, pp. 3534–3549, Apr. 2019.

[21] T. Nishio, R. Arai, K. Yamamoto, and M. Morikura, "Proactive traffic control based on human blockage prediction using RGBD cameras for millimeter-wave communications," in *Proc. 12th Annu. IEEE Consum. Commun. Netw. Conf. (CCNC)*, Jan. 2015, pp. 152–153.

[22] A. Orsino, R. Kovalchukov, A. Samuylov, D. Moltchanov, S. Andreev, Y. Koucheryavy, and M. Valkama, "Caching-aided collaborative D2D operation for predictive data dissemination in industrial IoT," *IEEE Wireless Commun.*, vol. 25, no. 3, pp. 50–57, Jul. 2018.

[23] D. Moltchanov, R. Kovalchukov, M. Gerasimenko, S. Andreev, Y. Koucheryavy, and M. Gerla, "Socially inspired relaying and proactive mode selection in mmWave vehicular communications," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 5172–5183, Jun. 2019.

[24] K. Sakaguchi, T. Haustein, S. Barbarossa, E. C. Strinati, A. Clemente, G. Destino, and A. Pärssinen, I. Kim, H. Chung, J. Kim, "Where, when, and how mmWave is used in 5G and beyond," *IEICE Trans. Electron.*, vol. E100-C, no. 10, pp. 790–808, 2017.

[25] K. Zeman, M. Stusek, J. Pokorny, P. Masek, J. Hosek, S. Andreev, P. Dvorak, and R. Josth, "Emerging 5G applications over mmWave: Hands-on assessment of WiGig radios," in *Proc. 40th Int. Conf. Telecommun. Signal Process. (TSP)*, Jul. 2017, pp. 86–90.

[26] S. Paasovaara and T. Olsson, "Proximity-based automatic exchange of data in mobile gaming: Studying the experiences of streetpass users," in *Proc. 9th Nordic Conf. Hum.-Comput. Interact.*, 2016, p. 26.

[27] S. Paasovaara, E. Olshannikova, P. Jarusriboonchai, A. Malapaschas, and T. Olsson, "Next2You: A proximity-based social application aiming to encourage interaction between nearby people," in *Proc. 15th Int. Conf. Mobile Ubiquitous Multimedia*, 2016, pp. 81–90.

[28] A. I. Sulyman, A. T. Nassar, M. K. Samimi, G. R. MacCartney, Jr., T. S. Rappaport, and A. Alsanie, "Radio propagation path loss models for 5G cellular networks in the 28 GHz and 38 GHz millimeter-wave bands," *IEEE Commun. Mag.*, vol. 52, no. 9, pp. 78–86, Sep. 2014.

[29] M. Elkashlan, T. Q. Duong, and H. Chen, "Millimeter-wave communications for 5G—Part 2: Applications [guest editorial]," *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 166–167, Jan. 2015.

[30] S. Sur, X. Zhang, P. Ramanathan, and R. Chandra, "BeamSpy: Enabling robust 60 GHz links under blockage," in *Proc. 13th USENIX Symp. Netw. Syst. Design Implement. (NSDI)*, 2016, pp. 193–206.

[31] A. Thornburg, T. Bai, and R. W. Heath, Jr., "Performance analysis of outdoor mmWave ad hoc networks," *IEEE Trans. Signal Process.*, vol. 64, no. 15, pp. 4065–4079, Aug. 2016.

[32] D. Moltchanov, A. Samuylov, V. Petrov, M. Gapeyenko, N. Himayat, S. Andreev, and Y. Koucheryavy, "Improving session continuity with bandwidth reservation in mmWave communications," *IEEE Wireless Commun. Lett.*, vol. 8, no. 1, pp. 105–108, Feb. 2019.

[33] K. Venugopal, M. C. Valenti, and R. W. Heath, Jr., "Device-to-device millimeter wave communications: Interference, coverage, rate, and finite topologies," *IEEE Trans. Wireless Commun.*, vol. 15, no. 9, pp. 6175–6188, Sep. 2016.

[34] K. Venugopal and R. W. Heath, Jr., "Location based performance model for indoor mmWave wearable communication," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–6.

[35] C. Slezak, V. Semkin, S. Andreev, Y. Koucheryavy, and S. Rangan, "Empirical effects of dynamic human-body blockage in 60 GHz communications," *IEEE Commun. Mag.*, vol. 56, no. 12, pp. 60–66, Dec. 2018.

[36] T. Bai and R. W. Heath, Jr., "Analysis of self-body blocking effects in millimeter wave cellular networks," in *Proc. 48th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2014, pp. 1921–1925.

[37] B. Han, L. Wang, and H. D. Schotten, "A 3D human body blockage model for outdoor millimeter-wave cellular communication," *Phys. Commun.*, vol. 25, no. P2, pp. 502–510, May 2017.

[38] L. Rabieekenari, K. Sayrafian, and J. S. Baras, "Autonomous relocation strategies for cells on wheels in environments with prohibited areas," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–6.

[39] S. Sekander, H. Tabassum, and E. Hossain, "Multi-tier drone architecture for 5G/B5G cellular networks: Challenges, trends, and prospects," *IEEE Commun. Mag.*, vol. 56, no. 3, pp. 96–103, Mar. 2018.

[40] M. Alzenad, A. El-Keyi, F. Lagum, and H. Yanikomeroglu, "3-D placement of an unmanned aerial vehicle base station (UAV-BS) for energy-efficient maximal coverage," *IEEE Wireless Commun. Lett.*, vol. 6, no. 4, pp. 434–437, Aug. 2017.

[41] J. Pokorny, A. Ometov, P. Pascual, C. Baquero, P. Masek, A. Pyattaev, A. Garcia, C. Castillo, S. Andreev, and J. Hosek, "Concept design and performance evaluation of UAV-based backhaul link with antenna steering," *J. Commun. Netw.*, vol. 20, no. 5, pp. 473–483, 2018.

[42] A. Ometov, E. Olshannikova, P. Masek, T. Olsson, J. Hosek, S. Andreev, and Y. Koucheryavy, "Dynamic trust associations over socially-aware D2D technology: A practical implementation perspective," *IEEE Access*, vol. 4, pp. 7692–7702, 2016.

[43] A. Ometov, P. Masek, J. Urama, J. Hosek, S. Andreev, and Y. Koucheryavy, "Implementing secure network-assisted D2D framework in live 3GPP LTE deployment," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC)*, May 2016, pp. 749–754.

[44] C. Huang, B. Zhai, A. Tang, and X. Wang, "Virtual mesh networking for achieving multi-hop D2D communications in 5G networks," *Ad Hoc Netw.*, vol. 94, Nov. 2019, Art. no. 101936.

[45] J. Simola and J. Rajamäki, "How a real-time video solution can affect to the level of preparedness in situation centers," in *Proc. 2nd Int. Conf. Comput. Sci., Comput. Eng., Social Media (CSCESM)*, Sep. 2015, pp. 31–36.

[46] E. A. Yfantis and S. L. Harris, "An autonomous UAS with AI for forest fire prevention, detection, and real time advice and communication to and among firefighters," *J. Comput. Sci. Appl. Inf. Technol.*, vol. 2, no. 3, p. 5, 2017.

[47] *Radiocommunication Objectives and Requirements for Public Protection and Disaster Relief*, Standard ITU-R M.2377-1, International Telecommunication Union, Geneva, Switzerland, Nov. 2017.

[48] A. Alnoman and A. Anpalagan, "On D2D communications for public safety applications," in *Proc. Canada Int. Hum. Technol. Conf. (IHTC)*, Jul. 2017, pp. 124–127.

[49] A. Kumbhar, F. Koohifar, I. Güvenç, and B. Mueller, "A survey on legacy and emerging technologies for public safety communications," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 1, pp. 97–124, 1st Quart., 2016.

[50] K. Kairbek, "TCP performance in 5G mmWave systems with dynamic blockage," M.S. thesis, Dept. Electron. Commun. Eng., Tampere Univ. Technol., Tampere, Finland, 2018.

[51] K. Zhou, Y. Qiao, and T. Xiang, "Deep reinforcement learning for unsupervised video summarization with diversity-representativeness reward," in *Proc. 32nd AAAI Conf. Artif. Intell.*, 2018, pp. 7582–7589.

[52] K. Lan, D.-T. Wang, S. Fong, L.-S. Liu, K. K. L. Wong, and N. Dey, "A survey of data mining and deep learning in bioinformatics," *J. Med. Syst.*, vol. 42, no. 8, p. 139, Aug. 2018.

[53] M. De Marsico, A. Petrosino, and S. Ricciardi, "Iris recognition through machine learning techniques: A survey," *Pattern Recognit. Lett.*, vol. 82, pp. 106–115, Oct. 2016.

[54] J. Padmanabhan and M. J. J. Premkumar, "Machine learning in automatic speech recognition: A survey," *IETE Tech. Rev.*, vol. 32, no. 2, pp. 240–251, 2015.

[55] M. Weber, M. Welling, and P. Perona, "Unsupervised learning of models for recognition," in *Proc. Eur. Conf. Comput. Vis.* Berlin, Germany: Springer, 2000, pp. 18–32.

[56] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: MIT Press, 2018.

[57] Y.-P. Lin and X.-Y. Li, "Reinforcement learning based on local state feature learning and policy adjustment," *Inf. Sci.*, vol. 154, nos. 1–2, pp. 59–70, 2003.

[58] J. M. Allman, *Evolving Brains* (Scientific American Library), no. 68. New York, NY, USA: Scientific American Library, 1999.

[59] X. X. Wang Li and V. C. M. Leung, "Artificial intelligence-based techniques for emerging heterogeneous network: State of the arts, opportunities, and challenges," *IEEE Access*, vol. 3, pp. 1379–1391, 2015.

[60] K. A. Kapalo, P. Bockelman, and J. J. LaViola, Jr., "'Sizing up' emerging technology for firefighting: Augmented reality for incident assessment," in *Proc. Hum. Factors Ergonom. Soc. Annu. Meeting*, vol. 62. Los Angeles, CA, USA: SAGE, 2018, pp. 1464–1468.

[61] K. Doppler, E. Torkildson, and J. Bouwen, "On wireless networks for the era of mixed reality," in *Proc. Eur. Conf. Netw. Commun. (EuCNC)*, 2017, pp. 1–5.

[62] K. Muhammad, S. Khan, M. Elhoseny, S. H. Ahmed, and S. W. Baik, "Efficient fire detection for uncertain surveillance environment," *IEEE Trans. Ind. Informat.*, vol. 15, no. 5, pp. 3113–3122, May 2019.

[63] T.-H. Chen, P.-H. Wu, and Y.-C. Chiou, "An early fire-detection method based on image processing," in *Proc. Int. Conf. Image Process.*, vol. 3, Oct. 2004, pp. 1707–1710.

[64] C. Yuan, Z. Liu, and Y. Zhang, "UAV-based forest fire detection and tracking using image processing techniques," in *Proc. Int. Conf. Unmanned Aircraft Syst. (ICUAS)*, Jun. 2015, pp. 639–643.

[65] C. Yuan, Z. Liu, and Y. Zhang, "Vision-based forest fire detection in aerial images for firefighting using UAVs," in *Proc. Int. Conf. Unmanned Aircr. Syst. (ICUAS)*, Jun. 2016, pp. 1200–1205.

[66] O. Giwa and A. Benkrid, "Fire detection in a still image using colour information," 2018, *arXiv:1803.03828*. [Online]. Available: https://arxiv.org/abs/1803.03828

[67] S. Frizzi, R. Kaabi, M. Bouchouicha, J.-M. Ginoux, E. Moreau, and F. Fnaiech, "Convolutional neural network for video fire and smoke detection," in *Proc. of 42nd Annu. Conf. IEEE Ind. Electron. Soc.*, Oct. 2016, pp. 877–882.

[68] M. Sajjad, S. Khan, K. Muhammad, W. Wu, A. Ullah, and S. W. Baik, "Multi-grade brain tumor classification using deep CNN with extensive data augmentation," *J. Comput. Sci.*, vol. 30, pp. 174–182, Jan. 2019.

[69] K. Muhammad, J. Ahmad, Z. Lv, P. Bellavista, P. Yang, and S. W. Baik, "Efficient deep CNN-based fire detection and localization in video surveillance applications," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 49, no. 7, pp. 1419–1434, Jul. 2019.

[70] K. Muhammad, J. Ahmad, and S. W. Baik, "Early fire detection using convolutional neural networks during surveillance for effective disaster management," *Neurocomputing*, vol. 288, pp. 30–42, May 2018.

[71] K. Muhammad, S. Khan, V. Palade, I. Mehmood, and V. H. C. De Albuquerque, "Edge intelligence-assisted smoke detection in foggy surveillance environments," *IEEE Trans. Ind. Informat.*, to be published.

[72] A. Ullah, K. Muhammad, J. Del Ser, S. W. Baik, and V. H. C. de Albuquerque, "Activity recognition using temporal optical flow convolutional features and multilayer LSTM," *IEEE Trans. Ind. Electron.*, vol. 66, no. 12, pp. 9692–9702, Dec. 2019.

[73] I. U. Haq, K. Muhammad, A. Ullah, and S. W. Baik, "DeepStar: Detecting starring characters in movies," *IEEE Access*, vol. 7, pp. 9265–9272, 2019.

[74] K. Muhammad, J. Ahmad, I. Mehmood, S. Rho, and S. W. Baik, "Convolutional neural networks based fire detection in surveillance videos," *IEEE Access*, vol. 6, pp. 18174–18183, 2018.

[75] P. Di Marco, R. Chirikov, P. Amin, and F. Militano, "Coverage analysis of Bluetooth low energy and IEEE 802.11ah for office scenario," in *Proc. 26th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Aug./Sep. 2015, pp. 2283–2287.

[76] O. Galinina, S. Andreev, A. Pyattaev, M. Gerasimenko, Y. Koucheryavy, N. Himayat, K. Johnsson, and S.-P. Yeh, "Modeling multi-radio coordination and integration in converged heterogeneous networks," in *Proc. Towards 5G, Appl., Requirements Candidate Technol.*, 2016, pp. 99–128.

[77] N. Himayat, S.-P. Yeh, A. Y. Panah, S. Talwar, M. Gerasimenko, S. Andreev, and Y. Koucheryavy, "Multi-radio heterogeneous networks: Architectures and performance," in *Proc. Int. Conf. Comput., Netw. Commun. (ICNC)*, Feb. 2014, pp. 252–258.

[78] M. Gerasimenko, D. Moltchanov, R. Florea, S. Andreev, Y. Koucheryavy, N. Himayat, S.-P. Yeh, and S. Talwar, "Cooperative radio resource management in heterogeneous cloud radio access networks," *IEEE Access*, vol. 3, pp. 397–406, 2015.

[79] M. Gerasimenko, D. Moltchanov, S. Andreev, Y. Koucheryavy, N. Himayat, S.-P. Yeh, and S. Talwar, "Adaptive resource management strategy in practical multi-radio heterogeneous networks," *IEEE Access*, vol. 5, pp. 219–235, 2016.

[80] A. Samuylov, M. Gapeyenko, D. Moltchanov, M. Gerasimenko, S. Singh, N. Himayat, S. Andreev, and Y. Koucheryavy, "Characterizing spatial correlation of blockage statistics in urban mmWave systems," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2016, pp. 1–7.

[81] M. Gapeyenko, A. Samuylov, M. Gerasimenko, D. Moltchanov, S. Singh, M. R. Akdeniz, E. Aryafar, N. Himayat, S. Andreev, and Y. Koucheryavy, "On the temporal effects of mobile blockers in urban millimeter-wave cellular scenarios," *IEEE Trans. Veh. Technol.*, vol. 66, no. 11, pp. 10124–10138, Nov. 2017.

[82] *Study on Channel Model for Frequencies From 0.5 to 100 GHz (Release 14), V14.1.1*, document TR 38.901, 3GPP, Jul. 2017.

[83] (Jul. 2019). *WINTERsim System-Level Simulator Description*. [Online]. Available: http://winter-group.net/downloads/

[84] P. J. Diggle, *Statistical Analysis of Spatial and Spatio-Temporal Point Patterns*. Boca Raton, FL, USA: CRC Press, 2013.

[85] T. Ismail, E. Leitgeb, and T. Plank, "Free space optic and mmWave communications: Technologies, challenges and applications," *IEICE Trans. Commun.*, vol. E99-B, no. 6, pp. 1243–1254, Jun. 2016.

[86] S. A. Busari, K. M. S. Huq, S. Mumtaz, L. Dai, and J. Rodriguez, "Millimeter-wave massive MIMO communication for future wireless systems: A survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 2, pp. 836–869, 2nd Quart., 2018.

[87] B. P. Zeigler, T. G. Kim, and H. Praehofer, *Theory of Modeling and Simulation*. New York, NY, USA: Academic, 2000.

RUSTAM PIRMAGOMEDOV received the M.Sc. and Cand.Sc. degrees from the St. Petersburg State University of Telecommunications, Russia, in 2010 and 2014, respectively. From 2010 to 2018, he worked in engineering companies focused on the Industrial IoT. Since 2018, he has been serving as an Associate Rapporteur of Q9/11 at the International Telecommunication Union. He is currently with Tampere University (TAU), Finland. His research interests include 5G/5G+ systems, blockchain, communication aspects of augmented human, and mission-critical applications of the IoT.

DMITRI MOLTCHANOV received the M.Sc. and Cand.Sc. degrees from the St. Petersburg State University of Telecommunications, Russia, in 2000 and 2003, respectively, and the Ph.D. degree from the Tampere University of Technology, in 2006. He is currently a University Lecturer with the Faculty of Information Technology and Communication Sciences, Tampere University, Finland. He has (co-)authored over 150 publications. His current research interests include 5G/5G+ systems, ultrareliable low-latency service, the industrial IoT applications, mission-critical V2V/V2X systems, and blockchain technologies.

ALEKSANDR OMETOV received the D.Sc. (Tech.) degree in telecommunications and the M.Sc. in information technology from the Tampere University of Technology (TUT), Finland, in 2016 and 2018, respectively, and the Specialist degree in information security from the Saint Petersburg State University of Aerospace Instrumentation (SUAI), Russia, in 2013. He is currently a Postdoctoral Researcher with Tampere University, Finland, managing H2020 A-WEAR EJD/ITN Project. His research interests are wireless communications, information security, blockchain, and wearable applications.

KHAN MUHAMMAD is currently an Assistant Professor with the Department of Software, Sejong University, South Korea. He has authored over 40 articles in peer-reviewed international journals, such as the IEEE TII, TIE, IoTJ, and TSMC-Systems. His research interests include information security, video summarization, computer vision, and video surveillance. He is a member of the ACM. He is also a reviewer of over 30 SCI/SCIE journals, including the *IEEE Communications Magazine*, IEEE NETWORK, IEEE INTERNET OF THINGS JOURNAL, TIP, TII, TCYB, and IEEE ACCESS.

SERGEY ANDREEV received the Cand.Sc. degree from SUAI, in 2009, and the Ph.D. degree from TUT, in 2012, as well as the Specialist degree, in 2006. He is currently an Assistant Professor of communications engineering and an Academy Research Fellow with Tampere University, Finland. Since 2018, he has also been a Visiting Senior Research Fellow with the Centre for Telecommunications Research, King's College London, U.K. He (co-)authored more than 200 published research works on intelligent IoT, mobile communications, and heterogeneous networking. He has been serving as an Editor of the IEEE WIRELESS COMMUNICATIONS LETTERS, since 2016, and as a Lead Series Editor of the IoT Series for *IEEE Communications Magazine*, since 2018.

YEVGENI KOUCHERYAVY received the Ph.D. degree from TUT, in 2004. He is currently a Full Professor at TAU. He is the author of numerous publications in the field of advanced wired and wireless networking and communications. His current research interests include various aspects in heterogeneous wireless communication networks and systems, the Internet of Things and its standardization, and nanocommunications. He is an Associate Technical Editor of the *IEEE Communications Magazine* and an Editor of the IEEE COMMUNICATIONS SURVEYS AND TUTORIALS.