

Integration of 5G Networks and Internet of Things for Future Smart City

Lead Guest Editor: Bo Rong

Guest Editors: Shuai Han, Michel Kadoch, Xi Chen, and Antonio Jara





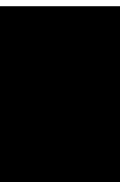
Integration of 5G Networks and Internet of Things for Future Smart City

Wireless Communications and Mobile Computing

Integration of 5G Networks and Internet of Things for Future Smart City

Lead Guest Editor: Bo Rong

Guest Editors: Shuai Han, Michel Kadoch, Xi Chen, and Antonio Jara



Copyright © 2020 Hindawi Limited. All rights reserved.

This is a special issue published in “Wireless Communications and Mobile Computing.” All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Editorial Board

Javier Aguiar, Spain
Ghufran Ahmed, Pakistan
Wessam Ajib, Canada
Muhammad Alam, China
Eva Antonino-Daviu, Spain
Shlomi Arnon, Israel
Leyre Azpilicueta, Mexico
Paolo Barsocchi, Italy
Alessandro Bazzi, Italy
Zdenek Becvar, Czech Republic
Francesco Benedetto, Italy
Olivier Berder, France
Ana M. Bernardos, Spain
Mauro Biagi, Italy
Dario Bruneo, Italy
Zhipeng Cai, USA
Jun Cai, Canada
Claudia Campolo, Italy
Gerardo Canfora, Italy
Rolando Carrasco, United Kingdom
Vicente Casares-Giner, Spain
Luis Castedo, Spain
Ioannis Chatzigiannakis, Italy
Yu Chen, USA
Lin Chen, France
Hui Cheng, United Kingdom
Ernestina Cianca, Italy
Riccardo Colella, Italy
Mario Collotta, Italy
Massimo Condoluci, Sweden
Daniel G. Costa, Brazil
Bernard Cousin, France
Telmo Reis Cunha, Portugal
Laurie Cuthbert, Macau
Donatella Darsena, Italy
Pham Tien Dat, Japan
André L. F. de Almeida, Brazil
Antonio De Domenico, France
Antonio de la Oliva, Spain
Gianluca De Marco, Italy
Luca De Nardis, Italy
Liang Dong, USA





Mohammed El-Hajjar, United Kingdom
Oscar Esparza, Spain
Maria Fazio, Italy
Mauro Femminella, Italy
Manuel Fernandez-Veiga, Spain
Gianluigi Ferrari, Italy
Ilario Filippini, Italy
Jesus Fontecha, Spain
Luca Foschini, Italy
Alexandros G. Fragkiadakis, Greece
Sabrina Gaito, Italy
Óscar García, Spain
Manuel García Sánchez, Spain
L. J. García Villalba, Spain
José A. García-Naya, Spain
Miguel Garcia-Pineda, Spain
Antonio-Javier García-Sánchez, Spain
Piedad Garrido, Spain
Vincent Gauthier, France
Carlo Giannelli, Italy
Carles Gomez, Spain
Juan A. Gómez-Pulido, Spain
Ke Guan, China
Antonio Guerrieri, Italy
Daojing He, China
Paul Honeine, France
Sergio Ilarri, Spain
Antonio Jara, Switzerland
Xiaohong Jiang, Japan
Minho Jo, Republic of Korea
Shigeru Kashihara, Japan
Dimitrios Katsaros, Greece
Minseok Kim, Japan
Mario Kolberg, United Kingdom
Nikos Komninos, United Kingdom
Juan A. L. Riquelme, Spain
Pavlos I. Lazaridis, United Kingdom
Tuan Anh Le, United Kingdom
Xianfu Lei, China
Hoa Le-Minh, United Kingdom
Jaime Lloret, Spain
Miguel López-Benítez, United Kingdom

Martín López-Nores, Spain
Javier D. S. Lorente, Spain
Tony T. Luo, USA
Maode Ma, Singapore
Imadeldin Mahgoub, USA
Pietro Manzoni, Spain
Álvaro Marco, Spain
Gustavo Marfia, Italy
Francisco J. Martinez, Spain
Davide Mattera, Italy
Michael McGuire, Canada
Nathalie Mitton, France
Klaus Moessner, United Kingdom
Antonella Molinaro, Italy
Simone Morosi, Italy
Kumudu S. Munasinghe, Australia
Keivan Navaie, United Kingdom
Thomas Newe, Ireland
Tuan M. Nguyen, Vietnam
Petros Nicopolitidis, Greece
Giovanni Pau, Italy
Rafael Pérez-Jiménez, Spain
Matteo Petracca, Italy
Nada Y. Philip, United Kingdom
Marco Picone, Italy
Daniele Pinchera, Italy
Giuseppe Piro, Italy
Sara Pizzi, Italy
Vicent Pla, Spain
Javier Prieto, Spain
Rüdiger C. Pryss, Germany
Sujan Rajbhandari, United Kingdom
Rajib Rana, Australia
Luca Reggiani, Italy
Daniel G. Reina, Spain
Jose Santa, Spain
Stefano Savazzi, Italy
Hans Schotten, Germany
Patrick Seeling, USA
Muhammad Z. Shakir, United Kingdom
Mohammad Shojafar, Italy
Giovanni Stea, Italy
Enrique Stevens-Navarro, Mexico
Zhou Su, Japan
Ville Syrjälä, Finland
Hwee Pink Tan, Singapore
Pierre-Martin Tardif, Canada




Mauro Tortonesi, Italy
Federico Tramarin, Italy
Reza Monir Vaghefi, USA
Juan F. Valenzuela-Valdés, Spain
Enrico M. Vitucci, Italy
Honggang Wang, USA
Jie Yang, USA
Sherali Zeadally, USA
Jie Zhang, United Kingdom
Meiling Zhu, United Kingdom

Contents




Integration of 5G Networks and Internet of Things for Future Smart City

Bo Rong , Shuai Han , Michel Kadoch , Xi Chen , and Antonio Jara
Editorial (2 pages), Article ID 2903525, Volume 2020 (2020)

A Deep Learning-Aided Detection Method for FTN-Based NOMA

Jianxiong Pan, Neng Ye , Aihua Wang , and Xiangming Li 
Research Article (11 pages), Article ID 5684851, Volume 2020 (2020)

Co-Channel Coexistence Analysis between 5G IoT System and Fixed-Satellite Service at 40 GHz

Xi Meng , Liyuan Zhong , Dong Zhou , and Dacheng Yang
Research Article (9 pages), Article ID 9790219, Volume 2019 (2019)

Practical Aspects for the Integration of 5G Networks and IoT Applications in Smart Cities Environments

Daniel Minoli  and Benedict Occhiogrosso
Review Article (30 pages), Article ID 5710834, Volume 2019 (2019)


Radar-Assisted UAV Detection and Identification Based on 5G in the Internet of Things

Jingcheng Zhao , Xinru Fu , Zongkai Yang, and Fengtong Xu
Research Article (12 pages), Article ID 2850263, Volume 2019 (2019)




Dynamic Traffic Prediction with Adaptive Sampling for 5G HetNet IoT Applications

Shuangli Wu , Wei Mao, Cong Liu , and Tao Tang
Research Article (11 pages), Article ID 4687272, Volume 2019 (2019)

A Hybrid Predictive Strategy Carried through Simultaneously from Decision Space and Objective Space for Evolutionary Dynamic Multiobjective Optimization

Peng Xu, Xiaoming Wu, Man Guo, Shuai Wang, Qingya Li , and Weiping Huang
Research Article (17 pages), Article ID 5190879, Volume 2019 (2019)

Performance Analysis for Downlink MIMO-NOMA in Millimeter Wave Cellular Network with D2D Communications

Jianguo Li , Xiangming Li , Aihua Wang, and Neng Ye 
Research Article (11 pages), Article ID 1914762, Volume 2019 (2019)

Editorial

Integration of 5G Networks and Internet of Things for Future Smart City

Bo Rong ¹, **Shuai Han** ², **Michel Kadoch** ³, **Xi Chen** ⁴ and **Antonio Jara**⁵

¹Communications Research Centre Canada, Ottawa, ON, Canada

²Department of Electronics and Communication Engineering, Harbin Institute of Technology, Harbin, Heilongjiang, China

³Department of Electrical Engineering, Ecole de Technologie Supérieure, Montreal, QC, Canada

⁴Flatiron Institute, Simons Foundation, New York, NY, USA

⁵University of Applied Sciences of Western Switzerland-HES SO Valais, Sierre, Switzerland

Correspondence should be addressed to Bo Rong; bo.rong@canada.ca

Received 7 November 2019; Accepted 7 November 2019; Published 12 February 2020

Copyright © 2020 Bo Rong et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The applications of IoT will proliferate in all sorts of devices, contributing to the infrastructure of smart city. These devices from different city sectors will cover every corner of the society, which calls for city-wide networks with ubiquitous accessibility to retrieve and deliver the produced data. Envisioning this demand, 5G has extended its mission to communicate things more than just people. The resulting 5G IoT contributes to the prosperity of the smart city ecosystem by allowing entities, big or small, to set up IoT services without the need to implement their own network facilities, compared with solutions, e.g., LoRa and SigFox. It is foreseeable plenty of smart city services will be running over 5G, pushing forward the integration of 5G and IoT. This trend, in turn, is going to impose unprecedented challenges on the on-building 5G mobile network and impact the normative work of 5G.

This special issue thus aims at bringing together the state-of-the-art innovations, research activities, and standardization updates related to the integration of 5G networks and IoT, in a bid to help both academic and industrial research communities understand the recent research advances and emerging technologies.

In the paper titled “Practical Aspects for the Integration of 5G Networks and IoT Applications in Smart Cities Environments,” the authors investigate a number of practical issues related to 5G-based IoT applications, including the need for small cells, the transmission issues at millimeter wave frequencies, building penetration issues, and the need for distributed antenna systems. To meet the special interest

in smart cities environments, this work also presents a brief introduction to pre-5G IoT technologies, such as NB-IoT and LTE-M.

The paper titled “Co-Channel Coexistence Analysis between 5G IoT System and Fixed-Satellite Service at 40 GHz” presents a promising way to successfully operate the fifth generation (5G) system with Internet of Things (IoT) in potential mmWave spectrum bands. The authors investigate the intelligent cochannel coexistence between the 5G IoT system and the fixed-satellite service (FSS) system at 40 GHz. The simulation results reveal that interference from the 5G IoT system into the FSS ground stations can be kept below the protection threshold by considering different deployment parameters, such as antenna patterns, height of Earth station (ES), and separation distance.

The paper titled “Radar-Assisted UAV Detection and Identification Based on 5G in the Internet of Things” proposes a radar-assisted positioning method for unmanned aerial vehicles (UAVs) based on 5G millimeter waves. The authors employ high-resolution range profile (HRRP), micro-Doppler characteristics, and the sinusoidal frequency modulation (SFM) parameter optimization method, respectively, to obtain the UAV location in the detection zone, identify the UAVs, extract the number and speed information of the UAV rotor, and separate multiple UAVs. The simulation results show that the proposed radar detection method is well suited for UAV detection and identification and provides a valid GPS-independent method for UAV tracking.

The paper titled “Performance Analysis for Downlink MIMO-NOMA in Millimeter Wave Cellular Network with D2D Communications” develops closed-form expressions for the outage probability and ergodic capacity in downlink MIMO-NOMA mmWave cellular network with D2D communications considered. The influencing factors of performance, such as transmission power and antenna number, are analyzed with a conclusion that higher transmission power and more antennas in the base station can decrease the outage probability and enhance the ergodic capacity of NOMA.

The paper titled “Dynamic Traffic Prediction with Adaptive Sampling for 5G HetNet IoT Applications” suggests an improved Call Session Control Function (CSCF) scheme, in which the improved CSCF server contains additional modules to facilitate IoT traffic prediction and resource reservation. The authors develop a compressed sensing based linear predictor to catch the traffic patterns. The proposed CSCF scheme can forecast the traffic load with high accuracy but low sampling overhead.

In the paper titled “A Hybrid Predictive Strategy Carried through Simultaneously from Decision Space and Objective Space for Evolutionary Dynamic Multiobjective Optimization,” the authors propose a hybrid prediction strategy carried through from both decision space and objective space (DOPS), in order to handle all kinds of optimization problems. The prediction in decision space is based on the center point, and the prediction in objective space is based on CTI. In addition, a kind of memory method is added to handle the problems with periodic changes, and a self-adaptive diversity maintenance method is adopted to compensate for the inaccuracy of the prediction in particularly complex problems. The experimental results show that DOPS is effective in dynamic multiobjective optimization.

Conflicts of Interest

The editors declare that they have no conflicts of interest regarding the publication of this special issue.

Bo Rong
Shuai Han
Michel Kadoch
Xi Chen
Antonio Jara

Research Article

A Deep Learning-Aided Detection Method for FTN-Based NOMA

Jianxiong Pan, Neng Ye , Aihua Wang , and Xiangming Li 

School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China

Correspondence should be addressed to Aihua Wang; wah@bit.edu.cn

Received 6 April 2019; Revised 4 June 2019; Accepted 5 September 2019; Published 29 January 2020

Guest Editor: Shuai Han

Copyright © 2020 Jianxiong Pan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The rapid booming of future smart city applications and Internet of things (IoT) has raised higher demands on the next-generation radio access technologies with respect to connection density, spectral efficiency (SE), transmission accuracy, and detection latency. Recently, faster-than-Nyquist (FTN) and nonorthogonal multiple access (NOMA) have been regarded as promising technologies to achieve higher SE and massive connections, respectively. In this paper, we aim to exploit the joint benefits of FTN and NOMA by superimposing multiple FTN-based transmission signals on the same physical resources. Considering the complicated intra- and interuser interferences introduced by the proposed transmission scheme, the conventional detection methods suffer from high computational complexity. To this end, we develop a novel sliding-window detection method by incorporating the state-of-the-art deep learning (DL) technology. The data-driven offline training is first applied to derive a near-optimal receiver for FTN-based NOMA, which is deployed online to achieve high detection accuracy as well as low latency. Monte Carlo simulation results validate that the proposed detector achieves higher detection accuracy than minimum mean squared error-frequency domain equalization (MMSE-FDE) and can even approach the performance of the maximum likelihood-based receiver with greatly reduced computational complexity, which is suitable for IoT applications in smart city with low latency and high reliability requirements.

1. Introduction

With the rapid development of 5G, higher demands have been brought forward for communication systems. The typical usage scenarios, such as smart city, virtual reality (VR), wearable computing, and the transmission of big data [1], will be effectively realized with the support of advanced radio access technologies in 5G. Smart city is a promising scenario which represents a whole new way of production and lifestyle characterized by automation and intelligence. Meanwhile, Internet of things (IoT), which enables human-computer interaction and machine-to-machine (M2M) communications, will be the foundation of the services in smart city [1, 2]. It is estimated that over 50 billion devices will be connected to the internet by 2020. With IoT, numerous devices in future smart city can be closely linked and a more intelligent life will be expected. To this end, the spectral efficiency and transmission latency shall be greatly improved and reduced, respectively, under massive connected devices. To fulfill these requirements, novel radio

access technologies are required, such as novel multiple access (MA) technologies, network architectures, encoding, and modulation methods [3–5].

Faster-than-Nyquist (FTN) is a promising modulation method to achieve high spectral efficiency [6]. In conventional digital communication theory, Nyquist points out that the symbol transmission rates must satisfy the Nyquist criterion in order to enable transmission without intersymbol interference (ISI). However, in 1975, Mazo [6] found that in band-limited additive white Gaussian noise channel, the normalized minimum Euclidean distance does not decrease when the symbol rate exceeds within 25% of the Nyquist rate. FTN signaling can boost up a transmission rate without consuming more bandwidth or increasing the number of transceivers' antennas. The process of FTN shaping destroys the orthogonality among symbols and introduces ISI unavoidably. In 1995, the authors in [7] proposed an achievable FTN transmission method, which uses the iterative method to design the filter. Due to the fact that the sinc pulse is unachievable in practical, the authors in

[8] studied the raised cosine-like pulse and thoroughly analyzed the error patterns at different FTN rates. The usage of constraint coding may further increase the rate of the FTN system at a cost of about 1 dB power and complexity. Recently, FTN has received much more attention owing to high spectral efficiency requirement of 5G. The authors of [9] found the FTN pulses aliased at the same time in the time domain and the frequency domain. Also, the time and frequency domain intervals satisfying the minimum distance are searched. The simulation results of Anderson et al. [10] reveal that when the transmitted pulse has excessive bandwidth, FTN transmission can achieve higher capacity than transmission under Nyquist rate. In other works, the research of FTN has been extended on multicarrier systems [11, 12] and low-complexity demodulation [13].

However, most research studies about FTN only focus on the point-to-point communication, which contradicts to the massive connection requirement in future smart city and IoT. To enable a practical multiuser communication system, efficient resource mapping patterns [14] and new multiple access methods are required. In 5G communication, non-orthogonal multiple access (NOMA) has been recognized as the promising access protocol due to the superior performance in spectral efficiency, connectivity, and flexibility [15]. Nonorthogonal schemes allow overlapping among the signals from different users by exploiting power domain, code domain or interleaver pattern, etc., and thus can provide better performance than the orthogonal counterparts [16]. In NOMA, multiple users share the same physical resource in downlink to increase the system capacity. Successive interference cancellation (SIC) is usually utilized to distinguish signals from multiple users [17]. Recently, many research studies combine grant-free transmission with NOMA scheme to simultaneously accomplish high throughput, large connectivity, and low energy cost [18–20]. FTN transmission is also considered in NOMA scenario to gain superior performance in spectral efficiency [21, 22]. However, the gains are finite and the structure of the receiver is too complicated to be applied in low-latency scenes.

NOMA and FTN are all essentially nonorthogonal signal processing methods. The former is nonorthogonal between users and the latter is in the level of symbols. FTN destroys the orthogonality among symbols and introduces ISI unavoidably. In order to eliminate ISI and ensure accuracy, the detection algorithm at the receiver is usually extremely complicated. Some detection methods, e.g., Viterbi algorithm [8], minimum mean squared error-frequency domain equalization (MMSE-FDE) [23], and Bahl–Cocke–Jelinek–Raviv (BCJR) detection [24], are studied to reduce the impact of ISI. The performance of Viterbi approaches maximum likelihood (ML) detection without ISI at the cost of huge complexity, which is obviously not suitable for the transmission scenarios with low latency requirements (e.g., VR and autopilot). Compared with Viterbi detections, FDE-based algorithm reduces the complexity to a certain extent but the usage of CP decreases the spectral efficiency and the detection accuracy does not meet the ideal requirement. Therefore, this paper aims to design a joint detection algorithm of

FTN-NOMA, where both low computational complexity and high detection accuracy are simultaneously achieved.

While Geoffrey Hinton rediscovered the potential of deep neural network (DNN) [25], deep learning (DL) has recently developed rapidly and has made huge achievements in various fields. It is especially skilled in solving the optimization problems for complex models in a data-driven fashion. Some previous research studies have considered the deployment of deep learning in wireless communication fields [26–29]. By utilizing deep networks, which has strong learning ability and universal approximation characteristic, DL can be deployed to simulate the communication system with complicated structure. Moreover, the data-driven optimization method can achieve the end-to-end optimization of the whole communication system as a whole. As a preliminary research, the authors in [4] exploited the joint benefit of grant-free access and NOMA transmissions to meet the requirements of tactile IoT with the aid of DL.

In this paper, we extend the typical FTN transmission into NOMA in multiuser scenario. To ensure the reliability and reduce detection latency, a DL-aided receiver method of FTN-NOMA named sliding window detection is proposed. First of all, we construct a neural network model of FTN's receiver to acquire the trained weight matrix and bias vector. Then, these matrixes are utilized to directly transform the received signals into transmitted bits with simple matrix multiplication and addition. The proposed scheme avoids the complicated iteration and convolution operations in conventional receivers, such as BCJR and Viterbi receiver. In addition, this scheme can ensure high detection accuracy which is close to the performance of ML.

The rest of this paper is organized as follows. The system model of multiuser FTN including the transmitter and receiver design is described in Section 2. In Section 3, we provide the specific proposed detection method based DNN. Simulation results are presented in Section 4, which validate the performance gain of the proposed detection scheme. Conclusions and future works are provided in Section 5.

2. The Proposed FTN-NOMA System Model

The paper considers an uplink two-user NOMA in AWGN channel. The system adopts the FTN method to transmit the modulated symbols with the shaping filter of root raised cosine (RRC) pulse, where ISI is introduced unavoidably. Considering the attenuation characteristic of RRC pulse, there exists little interference among symbols far apart. Therefore, we assume that each symbol is only interfered by two adjacent symbols, i.e., the number of detection nodes is equal to 3. Figure 1 describes our model of FTN transmission process in two-user case, in which the conventional demodulation receiver algorithms for FTN, e.g., Viterbi and FDE, are substituted by deep learning module.

2.1. Transmitter. For each user, after baseband modulation, the symbols would pass through the shaping filter $h(t)$. In the FTN method, the waveforms of each symbol are no longer orthogonal in time domain. The symbol time is

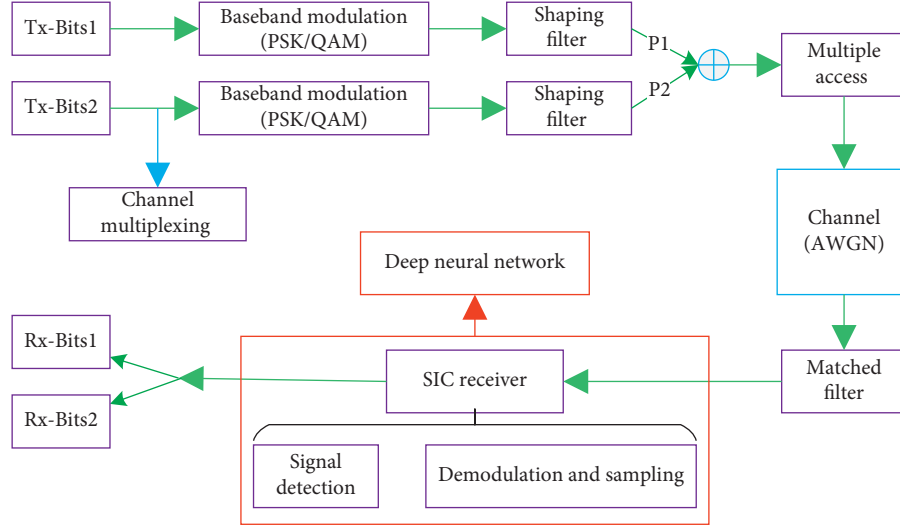


FIGURE 1: Transmission model of the proposed FTN-NOMA scheme.

shortened to αT , $0 < \alpha < 1$. Considering a point-to-point communication system, the general expression of single-user FTN is

$$s_i(t) = \sqrt{E_s} \sum_n a_n h(t - n\alpha T), \quad (1)$$

where E_s is the energy of BPSK symbols, $h(t)$ is the RRC shaping pulse with a roll-off factor $\beta = 0.5$, T is the symbol period satisfying Nyquist principle, a_n is the n -th BPSK symbol, and α is the FTN compression factor. If $\alpha = 1$, the whole process is equal to Nyquist transmission without any ISI. FTN reduces the distance among adjacent symbols in limited bandwidth and thus improves the spectral efficiency.

Figure 2 illustrates the transmitted symbols with Nyquist and FTN signaling, respectively. For each symbol's waveform, there exists the interference from other symbols, which is called ISI. The existence of ISI brings difficulty to the demodulation and detection of symbols.

In uplink NOMA case, the received waveforms at base station (BS) are the superposition of two users' FTN transmission symbols. The system model of uplink NOMA is shown in Figure 3, where there are two users in the cell and the signals from different users are superposed in power domain. Different from conventional multiple access methods, all users in the cell occupy the same time and frequency resources to transmit signals.

We assume that the two users are simultaneously multiplexed and user-2 is in the center of the cell. So the transmission power of user-1 is smaller than user-2. The FTN superposition signal of two users can be described by

$$x(t) = \sum_{i=1}^2 \sqrt{P_i} s_i(t) = \sum_{i=1}^2 \sqrt{P_i} \sum_n a_n h(t - n\alpha T), \quad (2)$$

where P_i represents the transmission power of user- i .

2.2. Receiver. Passing through AWGN channel, the received signal at BS becomes

$$\begin{aligned} y_k &= y(k\alpha T) \\ &= \sum_i h_i \times \sqrt{P_i} \sum_{n=1}^N a_{i,n} g(k\alpha T - n\alpha T) + \eta(k\alpha T) \\ &= \underbrace{h_1 \times \sqrt{P_1} \times a_k g(0)}_{\text{Expected Signal}} + \underbrace{h_1 \times \sqrt{P_1} \times \sum_{n \neq k} a_n g(k\alpha T - n\alpha T)}_{\text{ISI}} \\ &\quad + \underbrace{h_2 \times \sqrt{P_2} \sum_{n=1}^N a_{2,n} g(k\alpha T - n\alpha T)}_{\text{Interference among users}} + \underbrace{\eta(k\alpha T)}_{\text{Channel Noise}}, \end{aligned} \quad (3)$$

where h_i denotes channel response from user- i to BS and additive white Gaussian noise is denoted by η . And $g(t) = f(t) * f(t)$ is from the process of matched filtering.

Successive interference cancellation (SIC) is the general receiver algorithm of NOMA, which is performed at user- i to eliminate the multiple access interference and improve SINR of the desired signal. The basic operation of SIC is shown in Figure 4. For the signal from user-2 (higher transmission power), there is no SIC processing. The base station directly conducts signal detection to detect the desired signal regarding the signal from user-1 as noise. For user-1's signal, BS detects the signal from user-2 at first and then SIC is conducted to eliminate this interference. Finally, BS conducts signal detection to detect user-1's desired signal.

For single user, the process of signal detection is equal to the demodulation of FTN received symbols. In this paper, we consider two conventional detection algorithms, i.e., Viterbi decoding and FDE, to compare their performances with the proposed scheme.

2.2.1. Viterbi Algorithm. Viterbi algorithm is essentially the maximum likelihood (ML) detection method. However, different from the typical ML, Viterbi algorithm disperses the complexity into each symbol detection period. However,

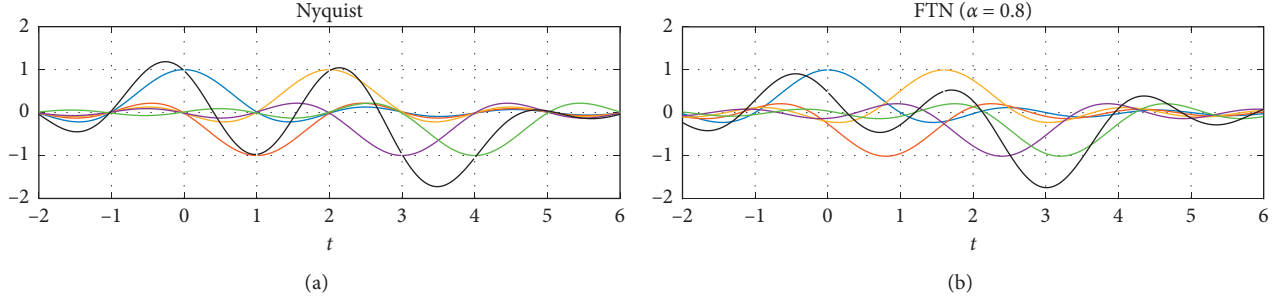


FIGURE 2: Transmitted waveforms with Nyquist and FTN signaling.

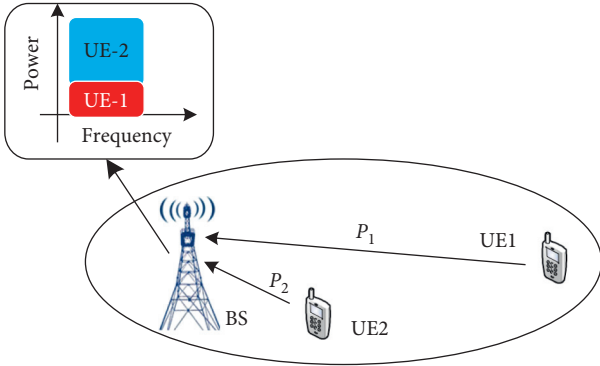


FIGURE 3: Uplink NOMA model in a cell with two users.

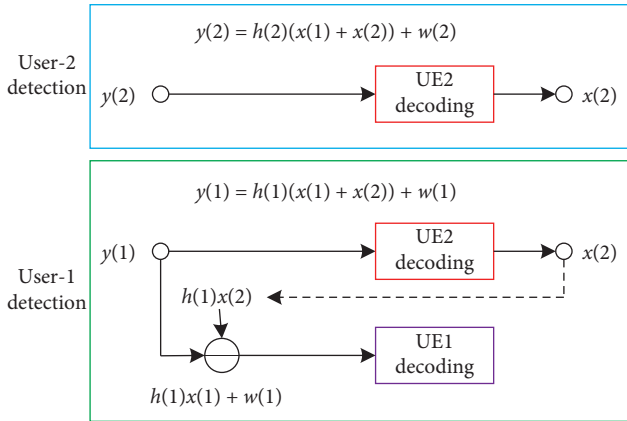


FIGURE 4: The detection process of the SIC receiver in two-user case.

the total complexity of Viterbi algorithm has no descent compared with the ML detection method. For FTN transmission, the Viterbi receiver has the optimum BER performance.

The process of FTN's Viterbi algorithm (BPSK modulation) is as follows:

- (i) Assume that each FTN symbol is interfered by the adjacent two symbols. The Viterbi trellis, as demonstrated in Figure 5, shall contain 4 state nodes which are denoted as a , b , c , and d , respectively. Each state node represents two source bits:

$$a - 00 \quad b - 01 \quad c - 10 \quad d - 11. \quad (4)$$

- (ii) For each received symbol, list all possible combination of transmitted bits. Each node has two possible paths. For example (the underline parts represent the possible symbol/bit at this moment):

$$\begin{aligned} a: \dots, 00 \underline{0}, \dots & \quad b: \dots, 00 \underline{1}, \dots \\ \dots, 10 \underline{0}, \dots & \quad \dots, 10 \underline{1}, \dots \\ c: \dots, 01 \underline{0}, \dots & \quad d: \dots, 01 \underline{1}, \dots \\ \dots, 11 \underline{0}, \dots & \quad \dots, 11 \underline{1}, \dots \end{aligned} \quad (5)$$

- (iii) Calculate the least Euclidean distance between all possible paths and the receive symbol at each moment:

$$\min_m \int_0^{NT+\Delta} |y(t) - \tilde{s}(t)|^2 dt \quad m = 1, 2, \dots, 8, \quad (6)$$

where $y(t)$ is the received symbol and $\tilde{s}(t)$ represents the possible path.

- (iv) For each node, choose the path with smaller Euclidean distance from the two possible paths and note down its distance. Thus, 4 paths survive from the possible 8 paths. For example, the remaining paths are

$$\begin{aligned} a: \dots, 000, \dots & \quad b: \dots, 101, \dots \\ c: \dots, 110, \dots & \quad d: \dots, 011, \dots \end{aligned} \quad (7)$$

- (v) Repeat the above process until the last received symbol is detected.

- (vi) Choose the shortest one from the last 4 paths as the optimum detection result.

The whole process can be described by Figure 5. It is worth mentioning that when the acceleration factor is no less than Mazo bound, the accuracy rate of Viterbi detection approaches the ML receiver without any ISI ($\alpha = 1$).

2.2.2. MMSE-FDE. FDE algorithm adds a short cyclic prefix into each transmission block to carry out fast Fourier transform- (FFT-) based low-complexity minimum-mean square error (MMSE) demodulation at the receiver. This

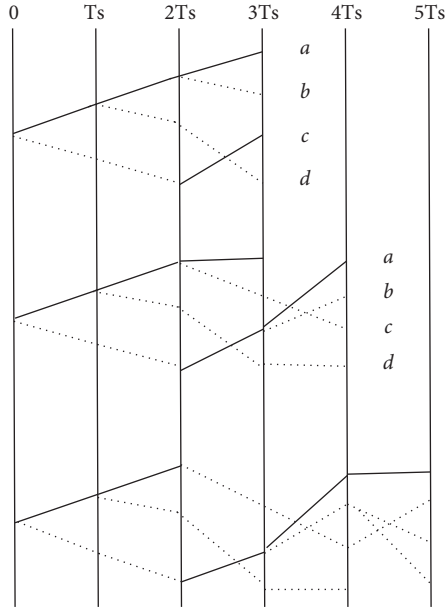


FIGURE 5: The process of Viterbi algorithm.

scheme is especially beneficial for a long-tap FTN case, where a delay spread associated with ISI is substantially large. The basic structure of FDE detection algorithm is illustrated in Figure 6.

At the transmitter, MMSE-FDE scheme divides the symbols into many blocks. Then, cycle prefix (CP) is added into the head of each block to reduce the interference among blocks. At the receiver, CP is first removed and then FFT transforms the received signal in time domain into frequency domain. ISI is eliminated by the process of frequency-domain equalization (FDE), the objective of which is to minimize the mean square error (MSE). Finally, the signal is transformed back into time domain by IFFT, and thus the interference among symbols is reduced. Compared with Viterbi algorithm, FDE can apparently reduce the complexity of detection process owing to the easy operation of FFT/IFFT. However, its accuracy rate is about 3~4 dB lower than the ML method. And the usage of CP reduces the spectral efficiency (SE) to a certain extent.

3. DL-Aided Sliding Window Detection Scheme

In this section, we will present the “sliding window detection” method based on deep learning. Figure 7 shows a set of consecutive symbols shaped by RRC pulses when the FTN acceleration factor is 0.8 and roll-off factor is 0.5. As can be seen, for each received symbol, the greatest interference is from the pulses of two adjacent symbols while the interference from other symbols is not in the same order of magnitude as the adjacent symbols. So, the other interferences could be neglected. Based on this observation, we think that each received symbol stores the information of three transmitted bits (itself and two adjacent bits). In other words, every three received symbols can establish a mapping to a two-dimensional transmitted bit vector (in two-user case).

We can utilize deep neural network to model the mapping between three received symbols and one transmitted bit. Therefore, we create a symbol window whose size is three. As shown in Figure 8, the window moves by one symbol once until all received symbols are detected. Then, the complete mapping of all symbols is established. The whole process is like the sliding window so we name this scheme as “sliding window detection.” At each sliding point, the real and imaginary parts of the three received symbols in the window are extracted. Thus, a matrix of six elements is formed as the input layer of the neural network:

$$y = \begin{bmatrix} 0 & y_{1,r} & y_{2,r} & \cdots & y_{N-1,r} \\ 0 & y_{1,i} & y_{2,i} & \cdots & y_{N-1,i} \\ y_{1,r} & y_{2,r} & y_{3,r} & \cdots & y_{N,r} \\ y_{1,i} & y_{2,i} & y_{3,i} & \cdots & y_{N,i} \\ y_{2,r} & y_{3,r} & y_{4,r} & \cdots & 0 \\ y_{2,i} & y_{3,i} & y_{4,i} & \cdots & 0 \end{bmatrix}, \quad (8)$$

and the output layer is a two-dimensional matrix including the transmitted bits from two users:

$$x = \begin{bmatrix} x_{1,1} & x_{1,2} & x_{1,3} & \cdots & x_{1,N} \\ x_{2,1} & x_{2,2} & x_{2,3} & \cdots & x_{2,N} \end{bmatrix}. \quad (9)$$

The first transmitted bit vector \vec{x}_1 and the last one \vec{x}_N map with only two received symbols. To keep the format of input layer, the head and tail of received symbol sequence are filled by zero. With the help of DNN, a mapping from the received symbol to the transmitted bit sequence is established. After a large amount of training steps, we can acquire the weight matrix \mathbf{W} and bias matrix \mathbf{B} , which can be utilized to directly transform the received symbols into transmitted bits. In the condition of two-user case and BPSK modulation, the above is essentially equal to a four-category classification method.

The proposed network as shown in Figure 9 consists of input layer, output layer, and three hidden layers. Sigmoid function is used as activation function of the neural network, which can establish a nonlinear mapping for neurons. The expression of sigmoid function is

$$f(x) = \frac{1}{1 + e^{-x}}. \quad (10)$$

In our proposed network, loss function is expressed by cross entropy. In the process of training steps, the loss function first maps the output values from the last layer to the interval of (0, 1) with sigmoid function and then calculates the cross entropy between the output value and training target:

$$\text{loss}_{ij} = -[x_{ij} * \ln p_{ij} + (1 - x_{ij}) \ln(1 - p_{ij})], \quad (11)$$

in which the expected output distribution is expressed by x_{ij} while p_{ij} represents the practical output distribution. In sigmoid function, the slope of the upper and lower boundary

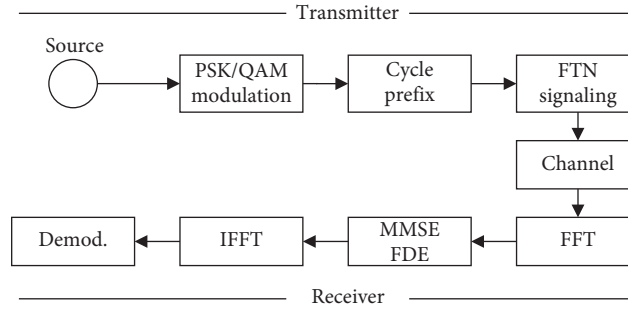


FIGURE 6: The process of MMSE-FDE algorithm.

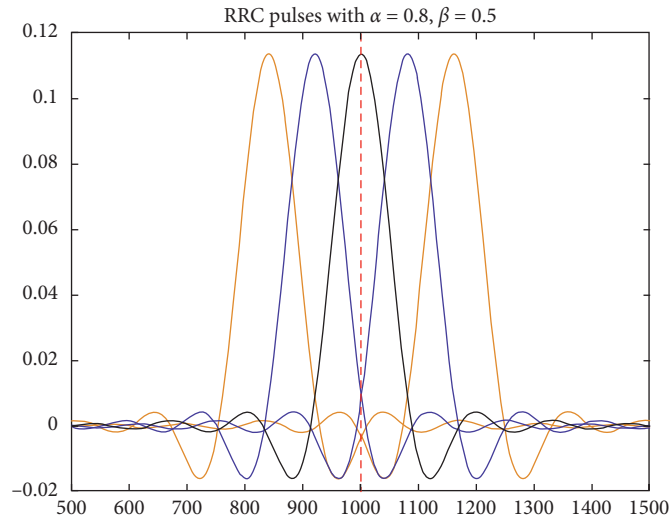


FIGURE 7: Superposition RRC pulses with FTN transmission.

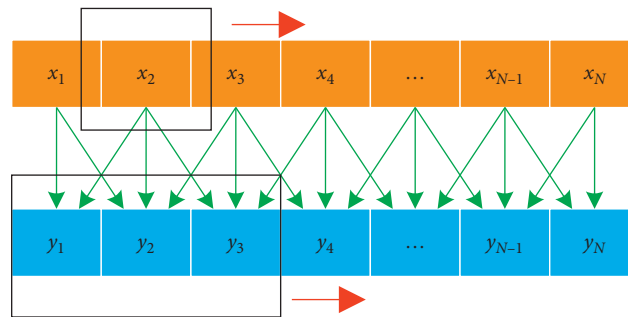


FIGURE 8: Illustration of the proposed sliding window detection method.

decreases rapidly; thus, the gradient value is extremely small when the training result is close to the real value, which slows down the convergence speed of the model. Cross entropy function is a logarithmic function which makes it maintain high gradient value when approaching the upper boundary. As a result, the convergence speed of the model will not be affected. So, the cross entropy can effectively show the similarity between the output value and the target distribution.

The objective of the training steps is to minimize the loss function. Here, we use adaptive moment estimation method (Adam) which is essentially a gradient descent method to

optimize the parameter. The training steps of gradient descent are given in Algorithm 1. Different from the typical gradient descent method, Adam has special way to update the parameters. Assume that the random variable X obeys a certain distribution; Adam constantly adjusts the learning rate α for each parameter according to the estimation of the first moment (i.e., the mean value of the sample) and the second moment (i.e., variance) of each parameter's gradient based on the loss function. Therefore, the learning speed of Adam can be controlled. In addition, the range of learning rate is limited, which can avoid big fluctuations of the network parameters, thus the value of the parameter in

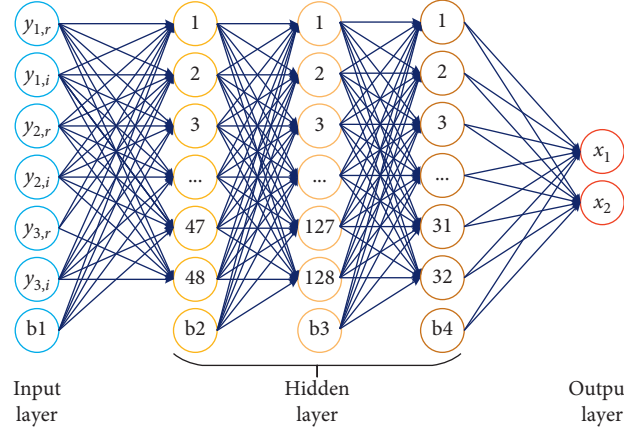


FIGURE 9: Structure of neural network for FTN-NOMA detection.

Input: \mathbf{y} : training data; \mathbf{x} : training labels; α : the learning rate;
 Output: $\mathbf{W}^{(l)}$: the connection weight matrix between layer l and layer $l+1$
 $\mathbf{b}^{(l)}$: the bias matrix between layer l and layer $l+1$
 $\mathbf{pred}^{(l)}$: output of layer l

(1) Set $\Delta \mathbf{W}^{(l)} = 0, \Delta \mathbf{b}^{(l)} = 0$ (matrix/vector of zeros) for all l .

(2) For $i = 1$ to m ,

(a) Use backpropagation to compute the partial derivatives: $\nabla_{\mathbf{W}^{(l)}} \mathbf{J}(\mathbf{W}, \mathbf{b}; \mathbf{y}, \mathbf{x})$ and $\nabla_{\mathbf{b}^{(l)}} \mathbf{J}(\mathbf{W}, \mathbf{b}; \mathbf{y}, \mathbf{x})$.

(i) Perform a feedforward pass, computing the activations for layers 1 to n_l ($n_l = 5$) and using the equation defining the forward propagation steps.

(ii) For the layer n_l , set $\delta^{(l)} = \mathbf{pred}^{(l)} - \mathbf{y}^{(l)}$.

(iii) For the layer $l = n_l - 1, n_l - 2, \dots, 3, 2$ set
 $\delta^{(l)} = (\mathbf{W}^{(l+1)})^T \delta^{(l+1)} \mathbf{f}'(\mathbf{z}^{(l)})$.

(iv) Compute the desired partial derivatives:

$$\nabla_{\mathbf{W}^{(l)}} \mathbf{J}(\mathbf{W}, \mathbf{b}; \mathbf{y}, \mathbf{x}) = \delta^{(l+1)} (\mathbf{pred}^{(l)})^T$$

$$\nabla_{\mathbf{b}^{(l)}} \mathbf{J}(\mathbf{W}, \mathbf{b}; \mathbf{y}, \mathbf{x}) = \delta^{(l+1)}$$

(b) Set $\Delta \mathbf{W}^{(l)} = \Delta \mathbf{W}^{(l)} + \nabla_{\mathbf{W}^{(l)}} \mathbf{J}(\mathbf{W}, \mathbf{b}; \mathbf{y}, \mathbf{x})$.

(c) Set $\Delta \mathbf{b}^{(l)} = \Delta \mathbf{b}^{(l)} + \nabla_{\mathbf{b}^{(l)}} \mathbf{J}(\mathbf{W}, \mathbf{b}; \mathbf{y}, \mathbf{x})$.

(3) Update the parameters:

$$\mathbf{W}^{(l)} = \mathbf{W}^{(l)} - \alpha [(1/m) \Delta \mathbf{W}^{(l)} + \lambda \mathbf{W}^{(l)}].$$

$$\mathbf{b}^{(l)} = \mathbf{b}^{(l)} - \alpha [(1/m) \Delta \mathbf{b}^{(l)}].$$

ALGORITHM 1: Gradient descent training algorithm of neural network.

AdamOptimizer is relatively stable. Adam optimizer improves the performance of typical gradient descent and promotes the dynamic adjustment of hyperparametric.

4. Simulation Results

In this section, the performances of the proposed sliding window detection scheme based on DL are presented. We consider single-user and two-user uplink NOMA cases where the ratio of transmitter power is 0.9:0.1. The root raised cosine (RRC) pulse whose roll-off factors are $\beta = 0.3$ and $\beta = 0.5$ is employed for shaping filter $h(t)$. BPSK symbols are adopted which are conveyed over the additive Gaussian white noise (AWGN) channel. For the process of FTN shaping, we consider two acceleration factors of $\alpha = 0.8$ and $\alpha = 0.5$. In addition to the proposed scheme, two conventional receiver algorithms, i.e., MMSE-FDE and Viterbi, are simulated as the references.

The training network consists of input layer, three hidden layers, and output layer with 6, 48, 128, 32, and 2 neurons, respectively. The proposed network is constructed and trained based on Tensorflow framework. The module of loss function, i.e., cross entropy in Tensorflow, is sigmoid_cross_entropy_with_logits, in which output values are mapped to the (0, 1) interval with sigmoid function at first, and then the cross entropy between the practical output value and training target is calculated. The AdamOptimizer module which can control the learning speed provided by Tensorflow environment is utilized to minimize the loss function.

In the following part of this section, we present the proposed scheme's performances of single-user and two-user cases.

4.1. Single-User FTN Case. For single-user case, the size of training set is 143360 (14 * 1024 * 10). The input layer of the network which represents the received signals is a matrix of

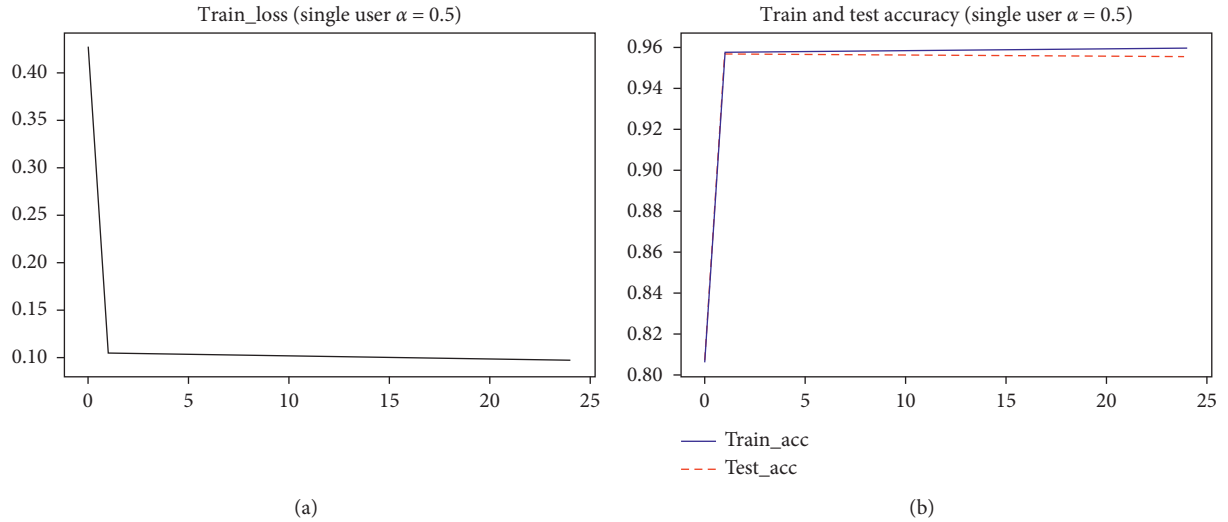


FIGURE 10: Training loss and accuracy versus training epoch. The unit of the x -coordinate is 10^3 epochs. The DNN-based receiver is trained with $\alpha = 0.5$ and $\beta = 0.5$ in the single-user case.

143360 * 6 while the vector of output layer has 143360 elements. In addition, we generalize a test set which is the same size as training set to detect the accuracy of training steps. To prevent overfitting, we cut the training set into several blocks in the size of 100. The learning rate of single user is set to 0.0001. The number of the whole training step is 25000. To timely acquire the training effect, the network prints the loss values and accuracy rate of training set and test set per 1000 steps.

Figure 10 shows the trends of loss values and accuracy rate of $\alpha = 0.5$. The total training steps are set to 25000. From the figure, we can find that the loss values and accuracy rate tend to be convergent after 10000 steps of training, where the convergence value of accuracy is about 0.96, which approaches the accuracy rate of Viterbi detection algorithm when SNR is 4~6 dB, i.e., the SNR point of dataset. The above results verify that DL-aided detection scheme is feasible to reach the ideal performance.

Figure 11 illustrates the BER performance of our proposed sliding window detection scheme compared with Viterbi and FDE when $\alpha = 0.8$. Here, we adopt two RRC shaping filters of $\beta = 0.3$ and $\beta = 0.5$. For the FDE method, two kinds of CP lengths of 2 and 144 are considered. The baseband modulation scheme is BPSK.

As can be seen, our proposed scheme achieves a better performance than the FDE method. Especially when $\beta = 0.5$, the proposed scheme can approach Viterbi algorithm with the gap of about only 1 dB. As noted above, Viterbi algorithm can achieve the best BER performance of FTN at the cost of high complexity. Compared with the Viterbi receiver, our scheme has much lower computational complexity which can be obviously observed in the simulation process. Different from FDE, it does not use CP in AWGN channel and thus avoids the loss of spectral efficiency. Therefore, our proposed DL-aided scheme can make a good trade-off among the performances of detection accuracy, spectral efficiency, and computation complexity.

The BER performances of the above three schemes with $\alpha = 0.5$ are illustrated in Figure 12. In this case, the accelerate factor α is lower than the Mazo limit and thus the BER performance of the FDE receiver has error floor. The Viterbi receiver can still achieve good detection performance at the cost of high computational complexity owing to the increase of state nodes. The BER performance of our proposed scheme can approach that of the Viterbi receiver while maintaining low computation complexity. Above results verify that our proposed scheme based on DL algorithm can combat strong ISI with low computational complexity in single-user case.

4.2. Multiuser Uplink FTN-NOMA Scene. In this section, we consider a two-user uplink FTN-NOMA case where the ratio of transmitter power is 0.9:0.1. To match the characteristics of two users simultaneously, the training set is collected from several signal to noise ratio (SNR) points. Considering the discontinuities of head and tail, we reduce the number of symbols. The size of training set is 84000 (14 * 60 * 100). The input layer of the network which represents the received signals is a matrix of 84000 * 6 while the vector of output layer has 84000 elements. In addition, we generalize a test set which is the same size as training set to detect the accuracy of training steps. In order to prevent overfitting, we cut the training set into several blocks with a size of 128. The learning rate of two-user is set to 0.0005. The number of the whole training step is 25000.

The performances of two-user FTN-NOMA case are illustrated in Figure 13. The proposed scheme, i.e., sliding window detection, is compared with the conventional SIC receiver where each user's signals are demodulated with Viterbi algorithm. For the SIC receiver, it is clear that BER performance of the user with higher transmitter power approaches single-user scenario with the Viterbi receiver. And the performance of lower powered user is about 10 dB

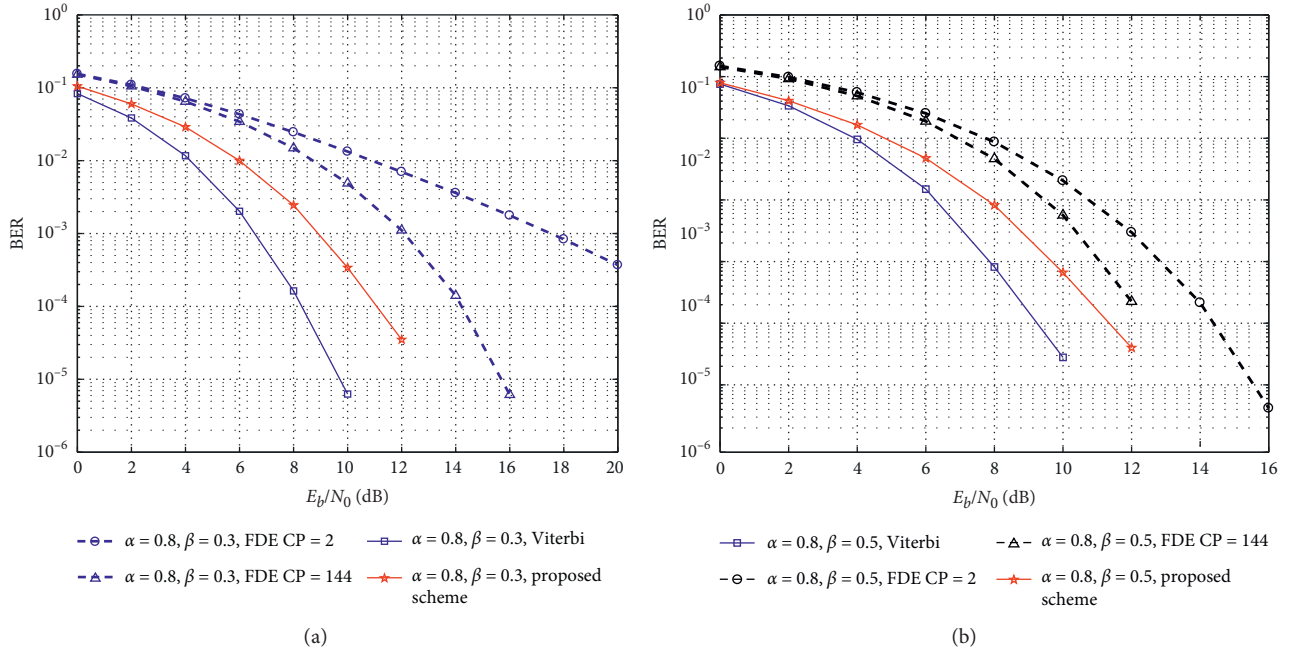


FIGURE 11: (a) BER performance of the proposed scheme, Viterbi, and FDE algorithms in single-user case at $\alpha = 0.8$ and $\beta = 0.3$. (b) BER performance of the proposed scheme, Viterbi, and FDE algorithms in single-user case at $\alpha = 0.8$ and $\beta = 0.5$.

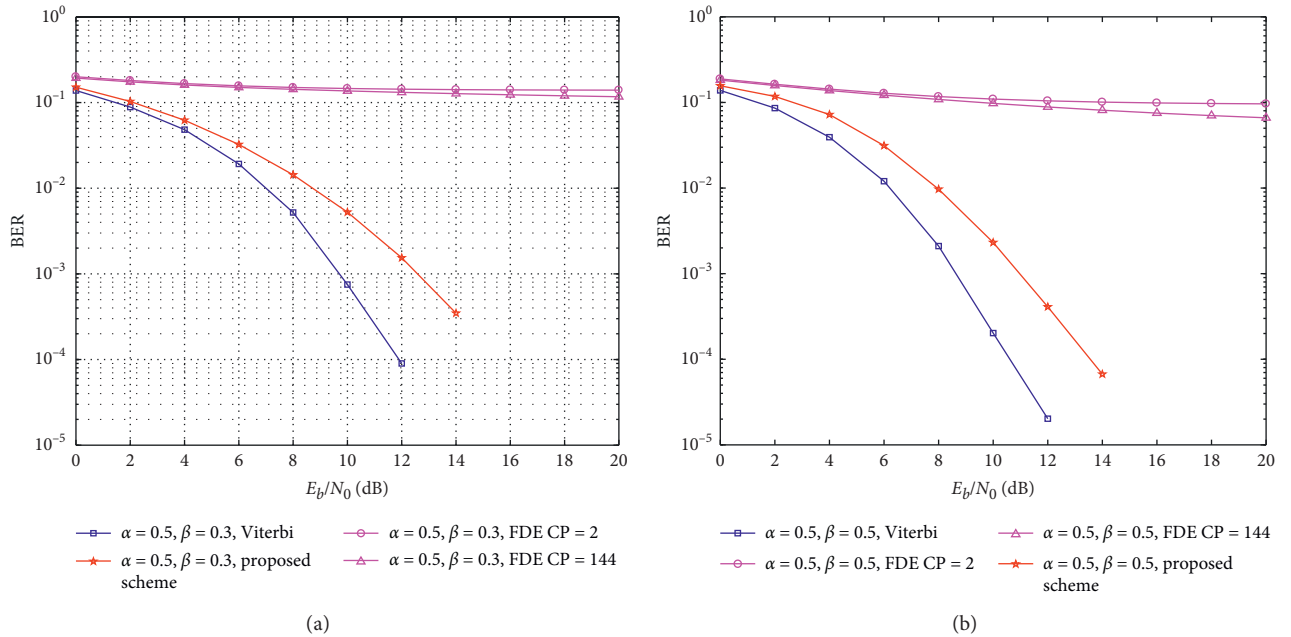


FIGURE 12: (a) BER performance of the proposed scheme, Viterbi, and FDE algorithms in single-user case at $\alpha = 0.5$ and $\beta = 0.3$. (b) BER performance of the proposed scheme, Viterbi, and FDE algorithms in single-user case at $\alpha = 0.5$ and $\beta = 0.5$.

worse than single-user case. For two-user FTN-NOMA case, the SIC receiver has double computational complexity than single-user scenario with the Viterbi detection method. As seen from Figure 12, the sliding window detection scheme has high detection accuracy which approaches the SIC receiver with Viterbi algorithm (ML detection method). In our proposed scheme, the descent of computational complexity is much more remarkable in NOMA case.

We assume that the training network consists of input layer, three hidden layers, and output layer with a , b , c , d , and e neurons (e depends on the number of users). So, the sizes of four trained weight matrixes are $a \times b$, $b \times c$, $c \times d$, and $d \times e$ respectively. Our proposed scheme can directly transform the received symbols into transmitted bits with matrix multiplication operation and the computational complexity only depends on operation times of the matrix

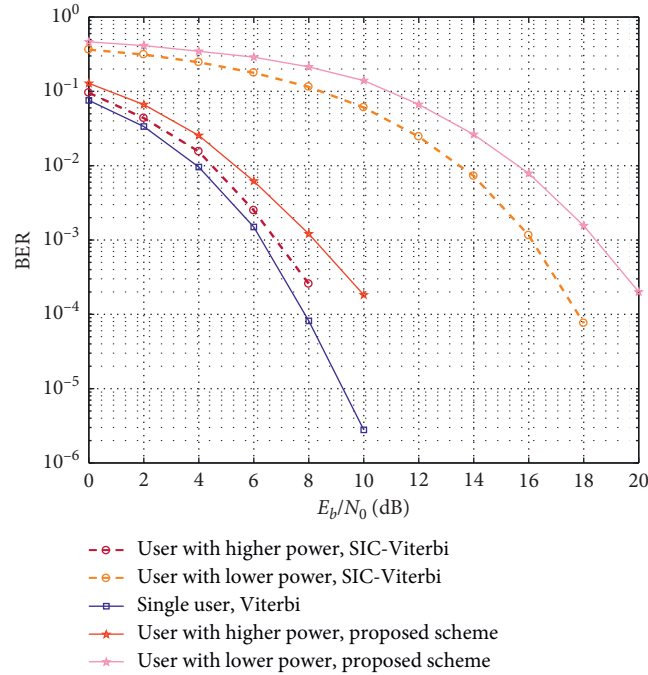


FIGURE 13: BER performance of the proposed scheme, Viterbi, and FDE algorithms in two-user case at $\alpha=0.8$ and $\beta=0.5$.

TABLE 1: Total multiplication operation times of different receiver schemes.

	Viterbi	MMSE-FDE	Proposed scheme
Single user	$N \times 2^N$	$2N^2 + 6N$	$N \times (a \times b + b \times c + c \times d + d)$
M -user	$M \times N \times 2^N$	$M \times (2N^2 + 6N)$	$N \times (a \times b + b \times c + c \times d + d \times M)$

elements. As mentioned above, the total complexity of the Viterbi receiver exponentially increases with the number of detected symbols. Considering M users, N transmitted bits of each user, and BPSK modulation, Table 1 provides the total times of multiplication operation with different receiver algorithms, which represents the computational complexity of these schemes. As can be seen, our proposed scheme has a remarkable effect on reducing the computational complexity of detection compared with the Viterbi receiver.

5. Conclusions

In this paper, we have extended the conventional FTN-based transmissions in NOMA and proposed a DL-aided receiver method of FTN-NOMA based on sliding window detection, which greatly reduces the computational complexity while maintaining high detection accuracy. Based on the data-driven optimization method, a neural network model of the FTN-NOMA receiver is constructed, which directly transforms the received signals into transmitted bits with simple matrix multiplication and addition. The proposed scheme avoids the complicated iteration and convolution operation of the conventional receiver like the BCJR or Viterbi method, which makes it suitable for low-latency scenes in smart city and IoT. Simulation results reveal that BER performance of our proposed scheme outperforms MMSE-

FDE algorithm and can approach the ML method which has the ideal performance in both OMA and NOMA cases. Moreover, the computational complexity can be remarkably reduced compared with the Viterbi or ML method due to the sliding window detection and DNN structure.

In the future, we will consider the DL-aided efficient detection methods of FTN-NOMA in multicarrier and multipath cases. In addition, we will study the robustness of DL methods considering high-order modulations in FTN-NOMA.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Authors' Contributions

J. Pan and N. Ye contributed equally to this study.

Acknowledgments

This work was supported by the Advance Research Projects of 13th Five-Year Plan of Civil Aerospace Technology

(B0105) and National Natural Science Foundation of China (61771051).

References

- [1] S. Han, S. Xu, W. Meng, and L. Cheng, "An agile confidential transmission strategy combining big data driven cluster and OBF," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 11, pp. 10259–10270, 2017.
- [2] V. Gazis, M. Goertz, M. Huber et al., "IoT: challenges, projects, architectures," in *Proceedings of the 2015 18th International Conference on Intelligence on Next Generation Networks*, pp. 145–147, Paris, France, February 2015.
- [3] S. Han, Yi Zhang, W. Meng, and H.-H. Chen, "Self-interference-cancellation-based SLNR precoding design for full-duplex relay-assisted system," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 9, pp. 8249–8262, 2018.
- [4] N. Ye, X. Li, H. Yu, A. Wang, W. Liu, and X. Hou, "Deep learning aided grant-free NOMA toward reliable low-latency access in tactile internet of Things," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 5, pp. 2995–3005, 2019.
- [5] S. Han, S. Xu, W. Meng, and Li Cheng, "Dense-device-enabled cooperative networks for efficient and secure transmission," *IEEE Network*, vol. 32, no. 2, pp. 100–106, 2018.
- [6] J. E. Mazo, "Faster-than-Nyquist signaling," *The Bell System Technical Journal*, vol. 54, no. 8, pp. 1451–1462, 1975.
- [7] C.-K. Wang and L.-S. Lee, "Practically realizable digital transmission significantly below the Nyquist bandwidth," *IEEE Transactions on Communications*, vol. 43, no. 2/3/4, pp. 166–169, 1995.
- [8] A. D. Liveris and C. N. Georghiadis, "Exploiting faster-than-Nyquist signaling," *IEEE Transactions on Communications*, vol. 51, no. 9, pp. 1502–1511, 2003.
- [9] F. Rusek and J. B. Anderson, "Constrained capacities for faster-than-nyquist signaling," *IEEE Transactions on Information Theory*, vol. 55, no. 2, pp. 764–775, 2009.
- [10] J. B. Anderson, F. Rusek, and V. Owall, "Faster-than-nyquist signaling," *Proceedings of the IEEE*, vol. 101, no. 8, pp. 1817–1830, 2013.
- [11] I. Kanaras, A. Chorti, M. R. D. Rodrigues, and I. Darwazeh, "Spectrally efficient FDM signals: bandwidth gain at the expense of receiver complexity," in *Proceedings of the 2009 IEEE International Conference on Communications*, pp. 1–6, Dresden, Germany, June 2009.
- [12] S. Li, B. Bai, J. Zhou, Q. He, and Q. Li, "Superposition coded modulation based faster-than-nyquist signaling," *Wireless Communications and Mobile Computing*, vol. 2018, Article ID 4181626, 10 pages, 2018.
- [13] P. Şen, T. Aktaş, and A. Ö. Yilmaz, "A low-complexity graph-based LMMSE receiver designed for colored noise induced by FTN-signaling," in *Proceedings of the 2014 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 642–647, Istanbul, Turkey, April 2014.
- [14] H. Yu, Z. Fei, N. Yang, and N. Ye, "Optimal design of resource element mapping for sparse spreading non-orthogonal multiple access," *IEEE Wireless Communications Letters*, vol. 7, no. 5, pp. 744–747, 2018.
- [15] N. Ye, H. Han, L. Zhao, and A. Wang, "Uplink nonorthogonal multiple access technologies toward 5G: a survey," *Wireless Communications and Mobile Computing*, vol. 2018, Article ID 6187580, 26 pages, 2018.
- [16] S. Han, Y. Huang, W. Meng, Li Cheng, N. Xu, and D. Chen, "Optimal power allocation for SCMA downlink systems based on maximum capacity," *IEEE Transactions on Communications*, vol. 67, no. 2, pp. 1480–1489, 2019.
- [17] N. Ye, A. Wang, X. Li, W. Liu, X. Hou, and H. Yu, "On constellation rotation of NOMA with SIC receiver," *IEEE Communications Letters*, vol. 22, no. 3, pp. 514–517, 2018.
- [18] N. Ye, A. Wang, X. Li, W. Liu, X. Hou, and H. Yu, "Rate-adaptive multiple access for uplink grant-free transmission," *Wireless Communications and Mobile Computing*, vol. 2018, Article ID 8978207, 21 pages, 2018.
- [19] K. Yang, N. Yang, N. Ye, M. Jia, Z. Gao, and R. Fan, "Non-orthogonal multiple access: achieving sustainable future radio access," *IEEE Communications Magazine*, vol. 57, no. 2, pp. 116–121, 2019.
- [20] N. Ye, A. Wang, X. Li, H. Yu, A. Li, and H. Jiang, "A random non-orthogonal multiple access scheme for mMTC," in *Proceedings of the 2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, pp. 1–6, Sydney, Australia, June 2017.
- [21] A. T. Abebe and C. G. Kang, "Iterative decoders for FTN-based NOMA scheme to multiplex sporadic and broadband transmission," in *Proceedings of the 2018 International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 813–817, Jeju, Republic of Korea, October 2018.
- [22] A. T. Abebe and C. G. Kang, "FTN-based MIMO transmission as a NOMA scheme for efficient coexistence of broadband and sporadic traffics," in *Proceedings of the 2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*, pp. 1–5, Porto, Portugal, June 2018.
- [23] S. Sugiura, "Frequency-domain equalization of faster-than-nyquist signaling," *IEEE Wireless Communications Letters*, vol. 2, no. 5, pp. 555–558, 2013.
- [24] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate (Corresp.)," *IEEE Transactions on Information Theory*, vol. 20, no. 2, pp. 284–287, 1974.
- [25] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.
- [26] H. Li, "A neural network blind multi-user detection algorithm," in *Proceedings of the 2009 International Conference on Measuring Technology and Mechatronics Automation*, pp. 467–470, Zhangjiajie, China, April 2009.
- [27] U. Mitra and H. V. Poor, "Neural network techniques for multi-user demodulation," in *Proceedings of the IEEE International Conference on Neural Networks*, vol. 3, pp. 1538–1543, San Francisco, CA, USA, March 1993.
- [28] M. Ohta, "Neural networks for multiuser detector in asynchronous PC/SS communication system," in *Proceedings of the 2002 International Joint Conference on Neural Networks. IJCNN'02 (Cat. No.02CH37290)*, vol. 3, pp. 2809–2813, Honolulu, HI, USA, May 2002.
- [29] F. Ibikunle and Y. X. Zhong, "Probabilistic neural networks for multi-user detection in code divisional multiple access communication channels," in *Proceedings of the 1998 IEEE International Joint Conference On Neural Networks Proceedings. IEEE World Congress on Computational Intelligence (cat. No. 98CH36227)*, vol. 3, pp. 2557–2560, Anchorage, AK, USA, May 1998.

Research Article

Co-Channel Coexistence Analysis between 5G IoT System and Fixed-Satellite Service at 40 GHz

Xi Meng ¹, Liyuan Zhong ², Dong Zhou ² and Dacheng Yang¹

¹School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China

²Wireless Product Operation Division in ZTE Corporation, Xi'an 710114, China

Correspondence should be addressed to Xi Meng; ximeng@bupt.edu.cn

Received 3 April 2019; Accepted 10 September 2019; Published 7 October 2019

Guest Editor: Bo Rong

Copyright © 2019 Xi Meng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

It is a promising way to successfully operate the fifth generation (5G) system with Internet of Things (IoT) in potential mmWave spectrum bands. This paper investigates the intelligent co-channel coexistence between the 5G IoT system and the fixed-satellite service (FSS) system at 40 GHz. The key issue, as identified, is the accurate estimation of interference based on mmWave propagation characteristics. Our simulation results reveal that interference from the 5G IoT system into the FSS ground stations can be kept below the protection threshold by considering different deployment parameters, such as antenna patterns, height of Earth station (ES), and separation distance.

1. Introduction

In the past, millimeter wave (mmWave) technologies were widely researched and deemed suitable for applications in space services, especially for different kinds of satellite communication services such as fixed-satellites service (FSS), mobile-satellite service (MSS), and inter-satellite service (ISS) [1]. However, today, the growth demand for both a multi-gigabits-per-second (multi-Gbps) rate and the involvement of IoT terminals becomes a major challenge for next-generation International Mobile Telecommunications (IMT) networks (5G) [2, 3]. In order to satisfy these requirements, it is necessary to introduce the mmWave, such as that at 40 GHz, to the 5G IoT system, as these bands provide an ultrawide band spectrum, the potential for spatial densification, and the ability to access a massive number of IoT terminals [4, 5].

One of the main difficulties in deploying mmWave technology is the protection of incumbents. Thus, it is essential to have an intelligent coexistence solution between the 5G IoT system and the FSS system in the same geographical area at 40 GHz, based on the exploration of the interference. Our study considers only the downlink

scenario, as ITU-R allocates other bands for FSS uplink transmission [6]. In particular, the interference into an ES receiver antenna may be generated by each single wireless transmission from base stations (BSs) to their associated IoT terminals and from the IoT terminals to their BSs.

The key issue for intelligent co-channel coexistence, as identified, is the accurate estimation of interference based on mmWave propagation characteristics. Interference estimation determines whether the 5G IoT system can share the same band with a predeployed FSS system. For example, if studies have shown that the two systems will generate unavoidable interference with each other, then RF engineers must design additional functionalities to mitigate or suppress this sort of interference. In contrast, if the spectrum sharing study concludes that both 5G IoT and FSS systems can coexist without generating any interference with each other, then it will be unnecessary to introduce interference mitigating technologies with extra overhead.

Initial results can be achieved based on previous preliminary coexistence analyses between the 5G system and existing services [7–15]. This paper, however, varies from previous studies in many ways. First, the results obtained by some previous researchers [7–9] focused on the relationship

between the advanced IMT system, also called the 4G system, and FSS; the target frequency was mainly from a 3.5 GHz, non-mmWave band, which is obviously different from the current research goals. Second, unlike the study undertaken in the current work, other previous research [10–13] has primarily considered the 5G scenario to be a traditional enhanced mobile broadband (eMBB) scenario rather than a 5G IoT scenario, such as ultrareliable and low latency (URLLC) and massive machine-type communications (mMTC). Finally, the findings by some scholars [14, 15] were obtained only from theoretical and deterministic analyses, whereas this paper additionally provides a simulation analysis and related evaluation methodology.

The reminder of this paper is organized as follows. Section 2 investigates the typical coexistence scenario between the 5G IoT system and the FSS system. Section 3 presents the simulation results by comparing a variety of propagation models and related methodologies. Finally, Section 4 concludes the paper with future research directions.

2. System Model and Coexistence Scenarios

2.1. Typical Coexistence Scenario. The coexistence scenario and related network topology are introduced to provide a direct view into the coexistence issue of the 5G IoT system and the FSS system.

Figure 1 illustrates the coexistence scenario for the 5G IoT system and the FSS system. More specifically, the 5G IoT system shares frequency bands from 37.5 to 42.5 GHz with the FSS system downlink. Thus, it is very likely that the FSS ES will be interfered with by the signals of the BSs and terminals of the 5G IoT system, which is reflected by the red dashed arrows. Moreover, the red grid indicates the separation distance between the 5G IoT system and the FSS system, the definition of which will be discussed in the next section of the paper. Because of the limited transmitting power of satellites, indicated by the black line, and the extremely long transmission distance between the satellite and the ES, the interference of the receiving FSS downlink signal will be ignored by the 5G IoT system because the signal is too weak to cause interference to it. In addition, the potential interference resource from the FSS uplink signal can mostly be ignored due to the obvious frequency isolation between the FSS uplink and 5G IoT system. It is important to note that only interfered signals from BSs are evaluated in this paper. The main reason for this is that the transmitted power of IoT terminals is much lower than that of base stations.

Based on analysis of the coexistence scenario, related topology should be considered for further evaluation. A typical network topology is presented in Figure 2, and it corresponds to the downlink band of the FSS system at 40 GHz. The evaluated area is close to a circle with a surface area of approximately 1000 km², which represents the developed area of a large city. The inner part of the city is an urban region with a radius of 8 km, and the outer part of the city is a suburban region with a radius of 18 km. It is notable that at least one BS is located on the ring with a radius of

$d_{\text{separation}}$, which is indicated by a yellow arrow in Figure 2. This means that at least one BS is at a separation distance from the FSS Earth station, because this type of topology is the worst case for the coexistence analysis between the 5G IoT and FSS systems.

Moreover, the FSS ES is placed at the center of the circle. Additionally, it should be noted that 40 GHz, as a millimeter wave, will mostly be deployed at the hotspot rather than exhibiting seamless coverage throughout the 5G IoT network; therefore, the number of BSs is wholly related to the ratio of hotspot areas to the area of a city. The total BS number (N_{IMT}) in the simulation is derived from the following equation:

$$N_{\text{IMT}} = S_{\text{urb}} * (D_{s_{\text{urb}}} * R_{a_{\text{urb}}}) + S_{\text{sub}} * (D_{s_{\text{sub}}} * R_{a_{\text{sub}}}), \quad (1)$$

where N_{IMT} expresses the total number of IMT BSs in the simulation, S_{urb} and S_{sub} refer to the sizes of the urban and suburban areas in km², respectively, and $D_{s_{\text{urb}}}$ and $D_{s_{\text{sub}}}$ are the density values for the outdoor urban and suburban hotspot areas, respectively, i.e., the density of simultaneously transmitting the IoT terminal or number of BSs per km². Finally, $R_{a_{\text{urb}}}$ and $R_{a_{\text{sub}}}$ represent the ratios of the hotspot areas to the urban and suburban areas, respectively. In Table 1, one set of typical values for equation (1) is presented, and these values are also considered for coexistence analysis in next chapter.

2.2. Model of the 5G IoT System with Massive MIMO. The 5G IoT system will use frequency bands both below and above 6 GHz. However, higher frequency bands, especially that at 40 GHz, are suitable for the IoT system in 5G, as a wider bandwidth which supports access to massive number of IoT terminals. Moreover, mmWave better meets the demands of different kinds of IoT devices by beamforming, which is one of the key technologies in 5G.

The 5G IoT system with beamforming is a massive MIMO wireless system [16] that employs a combination of limited feedback as illustrated in Figure 3. N_t transmitting antennas and N_r receiving antennas are used at transmitter and receiver, respectively. Figure 3 shows the antenna configuration of the 5G IoT system, where V -Ant. denotes the antenna with vertical (V) polarization, and H -Ant. denotes the antenna with horizontal (H) polarization. The wireless channel matrix is described by an input-output relation with H to H , H to V , V to V , and V to H polarized waves, based on the flat fading assumption.

More specifically, the signal y received by the IoT terminal from the 5G BS is represented as follows:

$$y = \sqrt{\rho} \mathbf{H} f x + n, \quad (2)$$

where f is the precoding matrix at the transmitter side, n is the N_r -dimensional noise vector, which has independent and identically distributed entries with normal distribution CN(0, 1), s represents the transmitted symbol with normalized energy, and ρ is the signal-to-noise ratio (SNR). The wireless channel $\mathbf{H} = \mathbf{H}(t)$ is a matrix that describes the dual-polarized MIMO channel. Its modelling method is discussed in

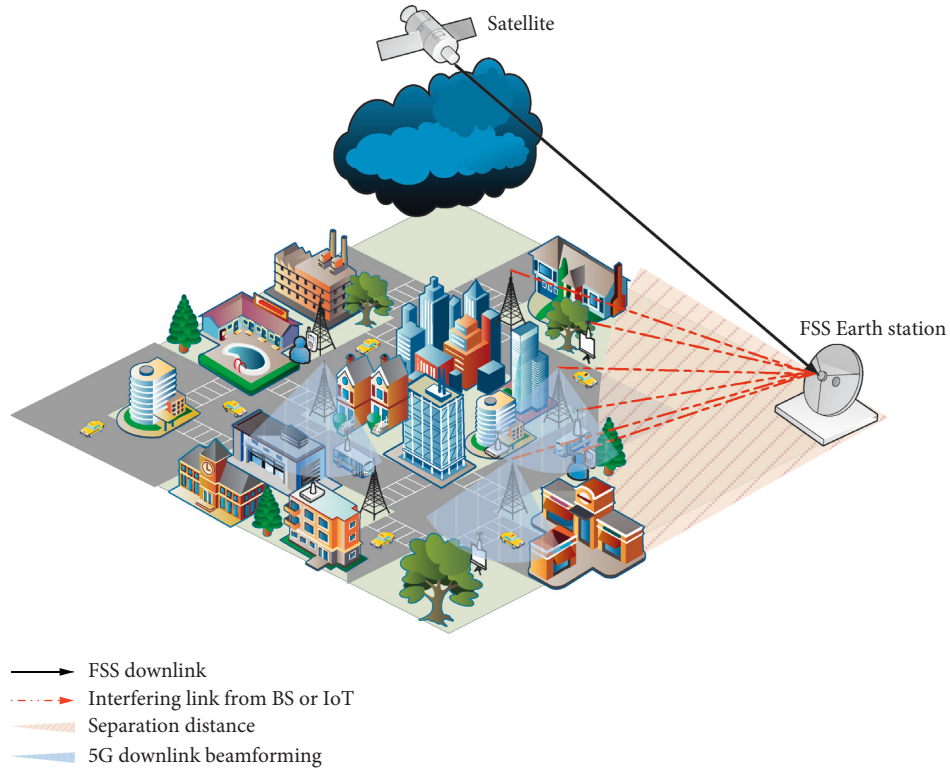


FIGURE 1: Illustration of 5G IoT sharing scenario with FSS.

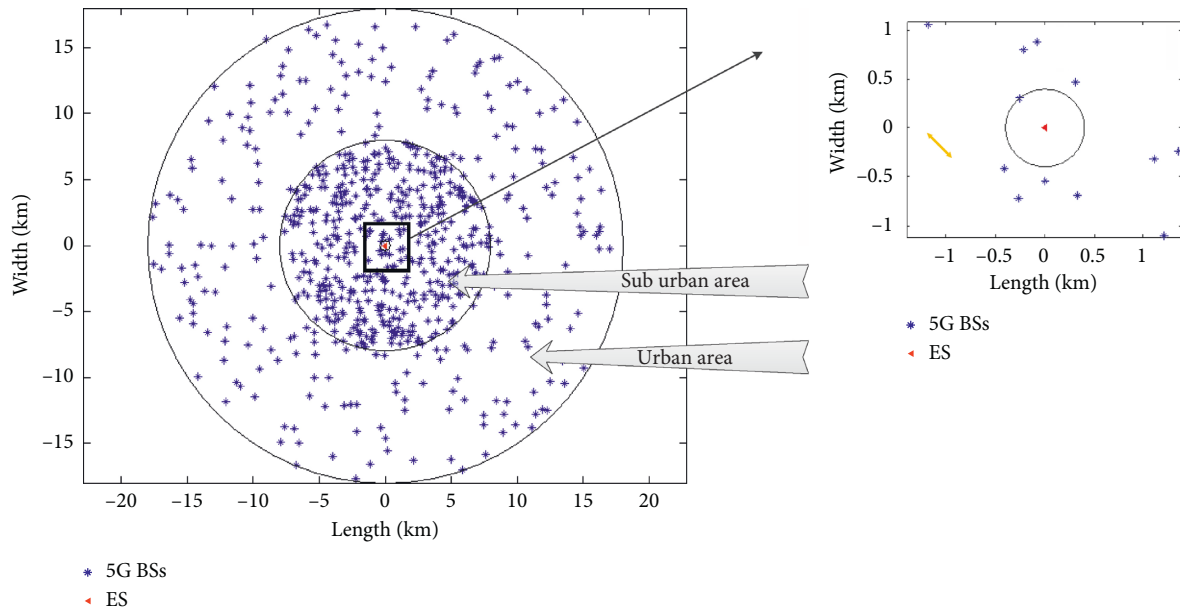


FIGURE 2: Topology of 5G IoT and FSS ES coexistence analysis.

TABLE 1: Typical value of D_s and R_a .

	Outdoor urban hotspot	Outdoor suburban hotspot
Density (devices per km^2)	$D_{s_{\text{urb}}} = 30$	$D_{s_{\text{sub}}} = 10$
Ratio (%)	$R_{a_{\text{urb}}} = 7$	$R_{a_{\text{sub}}} = 3$

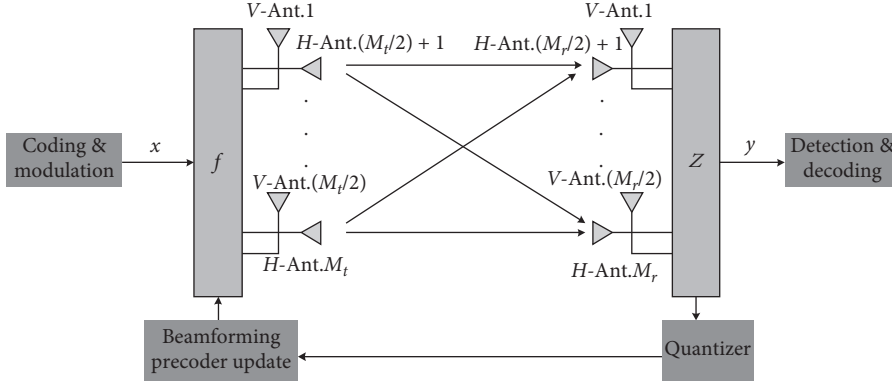


FIGURE 3: 5G IoT system model.

previous reports [17, 18], and any two receiving antenna channels (u, x) between the responses can be formulated as

$$\begin{aligned}
 H(t) = & \sqrt{P_n} \sum_{m=1}^M \begin{bmatrix} F_{rx,u,V}(\varphi_{n,m}, \gamma_{n,m}) \\ F_{rx,u,H}(\varphi_{n,m}, \gamma_{n,m}) \end{bmatrix}^T \\
 & \cdot \begin{bmatrix} \exp(j\Phi_{n,m}^{vv}) & \sqrt{\kappa} \exp(j\Phi_{n,m}^{vh}) \\ \sqrt{\kappa} \exp(j\Phi_{n,m}^{hv}) & \exp(j\Phi_{n,m}^{hh}) \end{bmatrix} \\
 & \cdot \begin{bmatrix} F_{tx,u,V}(\phi_{n,m}, \theta_{n,m}) \\ F_{tx,u,H}(\phi_{n,m}, \theta_{n,m}) \end{bmatrix} \\
 & \cdot \exp(j2\pi\lambda_0^{-1}\bar{r}_s \cdot \bar{\Phi}_{n,m}) \exp(j2\pi\lambda_0^{-1}\bar{r}_u \cdot \bar{\Psi}_{n,m}) \\
 & \cdot \exp(jk\|v\|\cos(\theta_{n,m,AoA} - \theta_v)t),
 \end{aligned} \quad (3)$$

where both n and m refer to the index of the cluster and each ray and P_n is the ray power in the n^{th} cluster. Additionally, $\varphi_{n,m}, \gamma_{n,m}$ represents the angles of arrival of the m^{th} ray, whereas $\phi_{n,m}, \theta_{n,m}$ expresses the angles of divergence of the m^{th} ray in the n^{th} cluster. The $F_{rx,u,V}$ and $F_{rx,u,H}$ functions are the multipath directions of $\varphi_{n,m}, \gamma_{n,m}$. Corresponding to the vertical and horizontal polarization antenna gain, κ represents the cross polarization power ratio in linear scale and λ_0^{-1} expresses the carrier frequency. Moreover, $\exp(j\Phi_{n,m}^{vv})$ and $\exp(j\Phi_{n,m}^{hh})$, respectively, represent the phase response of each ray in the randomly generated vertical polarization direction and horizontal plane, whereas $\exp(j\Phi_{n,m}^{vh})$ and $\exp(j\Phi_{n,m}^{hv})$ represent each horizontal to vertical, and vertical to horizontal polarization component leaked corresponding to random phase response. Then, \bar{r}_u and \bar{r}_s express the receiving end and the end of the antenna sends. $\bar{\Phi}_{n,m}$ and $\bar{\Psi}_{n,m}$ are vector expressions for the angles of arrival and departure, respectively. Finally, the IoT terminal movement speed is expressed as v .

In this paper, the antenna array configuration of the BS was assumed to be 8 and 16 elements placed along the horizontal and vertical directions, respectively, and a set of the main BS parameters is presented in Table 2. More importantly, the composite antenna gain is important for coexistence analysis [17] because it is the resulting beamforming antenna pattern from the logarithmic sum of the array gain in massive MIMO in mmWave. For example, the composite antenna gain can be formulated as

$$G_A(\theta, \varphi) = G_E(\theta, \varphi) + 10 \log \left(\left| \sum_{i=1}^{N_H} \sum_{j=1}^{N_V} w_{i,j} \cdot v_{i,j} \right|^2 \right), \quad (4)$$

where $G_E(\theta)$, represents the antenna element gain, $w_{i,j}$ is the weighting, and $v_{i,j}$ is the super position vector.

Based on equations (3) and (4), the antenna gain pattern for 126 (8×16) elements in the 5G IoT BS is shown in Figure 4. More specifically, its composite antenna gain pattern is three-dimensionally depicted, with the beam range covered from 0 to 180 degrees in vertical direction, but from -180 to 180 degrees in horizontal direction.

On the other hand, 4 antenna elements in both directions are assumed for the antenna array configuration of the IoT terminal. Additional related significant parameters of the IoT terminal can be found in Table 3, which also includes some important parameters of the BS. Furthermore, the antenna gain pattern was also created with the same method as that for the 5G BS and is demonstrated in Figure 5. The composite antenna gain pattern of the IoT terminal with 16 (4×4) antenna elements is three-dimensionally drawn too.

It is obvious from both Figures 4 and 5 that in the massive MIMO 5G IoT system, the width of the beam is narrower with the increase in the of number of antenna elements. The main reason for this is due to one of the advantage of massive MIMO in mmWave; although the coverage of the beam is limited, the transmitted power is more concentrated, which effectively improves the channel capacity. In other words, regarding sharing with the FSS ES system, the probability of interference will decrease, as it is difficult to point to the antenna of the ES from a narrower beam of the 5G IoT system.

2.3. Model of FSS System. FSS is one of the main applications in satellite communication. FSS provides the dual-transmission path between the satellite and its ES to provide telecommunication services including data, video, and voice transmission. In this paper, FSS-receiving ESs operate in the space-to-Earth direction from 40 GHz, more specifically from the 37.5 to 42.5 GHz frequency bands. The main downlink ES parameters are presented in Table 4.

TABLE 2: Key parameters for BS in 5G IoT system.

Parameter	Outdoor suburban hotspot/outdoor urban hotspot
Network topology and characteristics	Suburban: 10 BS/km ² Urban: 30 BSs/km ²
Frequency reuse	1
Antenna height (radiation center)	6 m
Sectorization	Single sector
Antenna down tilt	10 degrees
Antenna deployment	Below rooftop
Network loading factor	20%
BS TDD activity factor	80%
Element gain $G_E(\theta, \varphi)$	5 dBi
3 dB beam width of single element	65 degrees for both H/V
Front-to-back ratio	30 dB for both H/V
Antenna polarization	Linear $\pm 45^\circ$
Radiating element spacing	0.5 of wavelength for both H/V
Array ohmic loss	3 dB
Conducted power per antenna element	8 dB (m/200 MHz)
BS maximum coverage angle in the horizontal plane	120°

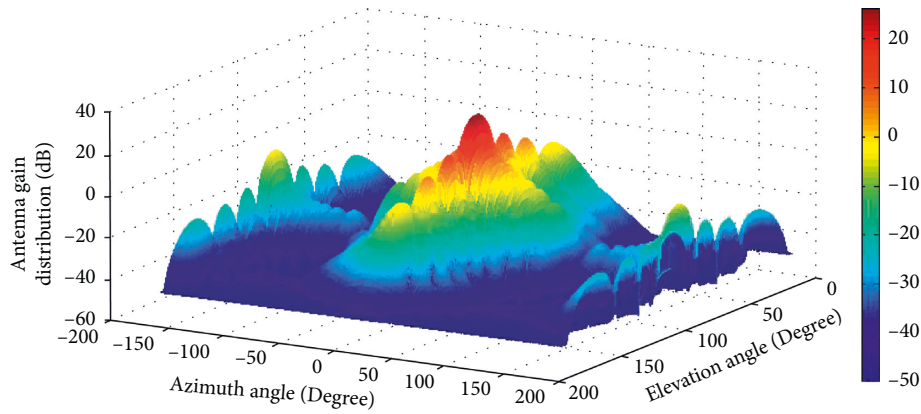
FIGURE 4: Composite antenna gain pattern for 5G BS with 8×16 antenna elements.

TABLE 3: Key parameters for terminal in the 5G IoT system.

Parameters	Outdoor suburban hotspot/outdoor urban hotspot
Network topology and characteristics	Suburban: 30 terminals/km ² Urban: 100 terminals/km ²
Body loss	4 dB
Indoor IoT terminal usage	5%
Terminals TDD activity factor	20%
3 dB beam width of single element	90° for both H/V
Front-to-back ratio	25 dB for both H/V
Antenna array configuration (row \times column)	4 \times 4 elements
Maximum IoT terminal output power	22 dBm
Transmit power target value per 180 kHz	-95 dBm
Path loss compensation factor α	1

Furthermore, in order to conduct a sharing study between the 5G IoT system and the FSS system, it is assumed that the antenna elevation angle of the ES is chosen as the minimum value, i.e., 10 degrees. Although this value is pessimistic, it means fewer dropping locations of the ES's antenna pointing towards the satellite and always indicates the worst-case assumption in further coexistence analysis.

The FSS ES antenna gain is calculated as a function of the off-boresight angles [5, 18]. Assuming ϑ is the azimuth of the 5G BS or IoT terminal without the ES main lobe, the off-boresight angles ϕ of the ES towards the BS or terminal can be computed as follows:

$$\phi = \arccos(\cos(\alpha)\cos(\varepsilon)\cos(\vartheta) + \sin(\alpha)\sin(\varepsilon)), \quad (5)$$

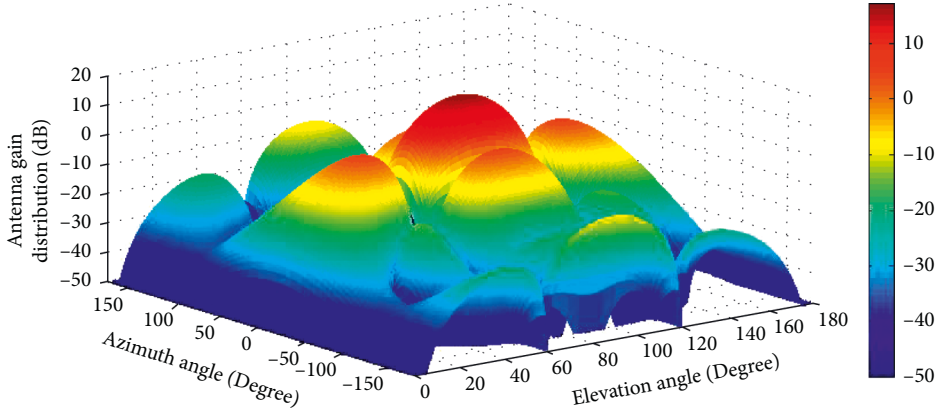


FIGURE 5: Composite antenna gain pattern for 5G IoT terminal with 4×4 antenna elements.

TABLE 4: Parameters for ES in FSS.

Parameter	Value
Frequency range (GHz)	37.5–42.5
Noise bandwidth (MHz)	100–600
ES antenna diameter (m)	6.8
Peak receive antenna gain (dBi)	68
System receive noise temperature (K)	250
Minimum ES elevation angle (degrees)	10
Interference-to-noise ratio I/N (dB)	-12.2

where α refers to the elevation angle. Additionally, ϵ is calculated as

$$\epsilon = \frac{h_{\text{ES}} - h}{d} - \frac{d}{2r}, \quad (6)$$

where h and h_{ES} are defined as the heights of the 5G BS or IoT terminals and ES in meters, respectively, and r is the effective Earth radius, approximately 8.5×10^3 km. Therefore, the FSS ES off-boresight antenna gain pattern, the unit of which is dBi, is calculated as

$$G = \begin{cases} 32 - 25 \log \varphi, & \text{for } 1^\circ \leq \varphi \leq 48^\circ, \\ -10, & \text{for } 48^\circ \leq \varphi \leq 360^\circ. \end{cases} \quad (7)$$

3. Coexistence Analysis Results for Intelligent Spectrum Sharing between 5G IoT System and FSS

3.1. Propagation Model. The propagation model refers to the propagation fading effect, which is a significant model for studies of sharing between the 5G IoT and FSS systems. In this paper, the propagation model describes a ground-to-ground wireless communications system. More importantly, the model is designed as a prediction method for the evaluation of the interference between BSs or IoT terminals and ESs on the Earth surface at frequencies from about 0.1 GHz to 50 GHz. The model should consider a set of propagation mechanisms, including attenuation by atmospheric gases, diffraction, free-space propagation, tropospheric scatter, clutter loss, and ducting reflection. However, since the propagation model assumes known information of the location of BSs, IoT terminals, and ES, its clutter loss

model is inappropriate for our analysis. Therefore, an additional method to calculate cluster loss is presented below.

The clutter loss not exceeded for $p\%$ of locations for the terrestrial to the terrestrial path [9], L_{ctt} (dB), is given by

$$\begin{aligned} L_{\text{ctt}} = & -5 \log(10^{-0.2(23.5+9.6 \log(f))}) \\ & + 10^{-0.2(32.98+23.9 \log(d)+3 \log(f))} \\ & - 6Q^{-1}\left(\frac{p}{100}\right), \end{aligned} \quad (8)$$

where d is the total path length in km, f is the frequency in GHz, and $Q^{-1}(p/100)$ is the inverse complementary normal distribution function with p .

Based on the above introduced propagation mode, Figure 6 illustrates cluster loss in different distances and frequencies, in which the green line represents a cluster loss around 40 GHz; it is also applied in our simulation.

3.2. Methodology. In this section, a novel Monte Carlo evaluation methodology is introduced to analyze the aggregate interference from the 5G IoT system, including BSs and terminals, to a FSS ES.

First, the ES is located in the center, and random distributions of both BSs and the IoT terminal are generated according to the sharing study scenario mentioned previously. In addition, their interference links from the 5G BSs or IoT terminals to the ES also are produced. It is worth noting that the BSs and terminal's beams face toward each other due to the application beamforming and the randomness of the interference resource.

Second, the related parameters and antenna pattern, such as the transmit power and network load, should be configured and are presented in Tables 2–4 for both the 5G IoT and FSS systems.

Third, the interference model from the i^{th} 5G BS or the IoT terminals linking to the FSS ES are shown by using the following equation:

$$I_{\text{IMT} \rightarrow \text{FSS}}(\theta_{\text{tx}}^i, \theta_{\text{rx}}^i) = \text{PD}_{\text{tx}} - \text{OL}_{\text{tx}} + G_{\text{tx}}(\theta_{\text{tx}}^i) - \text{PL} + G_{\text{rx}}(\theta_{\text{rx}}^i), \quad (9)$$

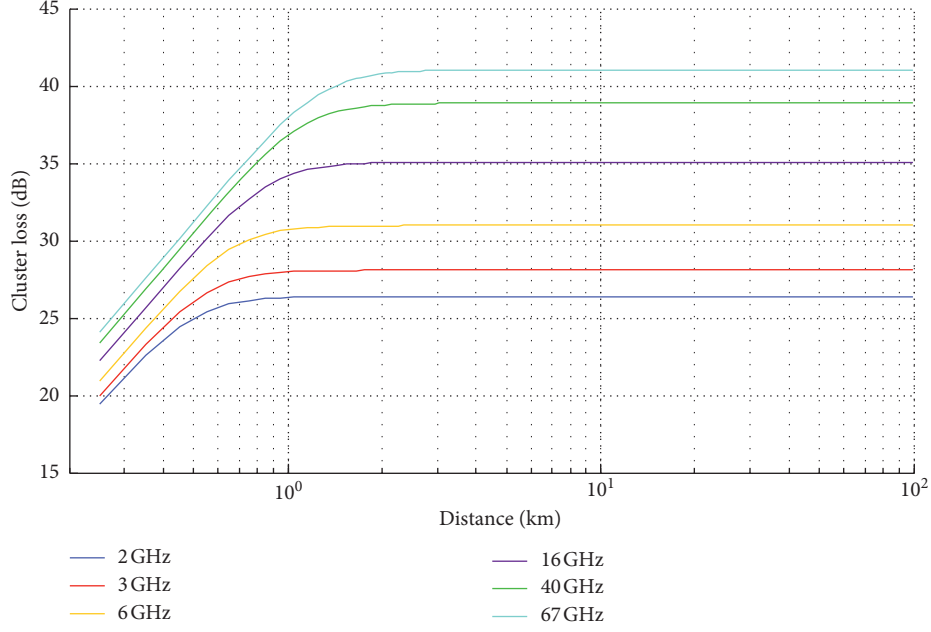


FIGURE 6: Cluster loss model with varying frequencies and distances for ES.

where PD_{tx} is the transmit station signal power density in dBm/MHz, OL_{tx} represents the transmit station array ohmic loss in dB, and $G_{tx}(\theta_{tx}^i)$ is the 5G BS or IoT terminals' antenna gain in the direction of the ES, accounting for the beamforming antenna pattern in dBi. $G_{rx}(\theta_{rx}^i)$ represents the antenna gain of the ES in the direction of the transmit station in dBi, and PL refers to propagation loss, which includes path loss, clutter loss, and polarization discrimination in dB.

Furthermore, the aggregate interference from all 5G BSs or IoT terminals at the FSS ES receiver level is calculated by the sum of the FSS ES from each interference link from equation (9).

$$I_{agg}^k = 10 \log \left(\sum_{i=1}^n 10^{I_{IMT \rightarrow FSS}(\theta_{tx}^i, \theta_{rx}^i)/10} \right). \quad (10)$$

In equation (10), I_{agg}^k represents the k^{th} running simulation to obtain the aggregate external system interference towards the interfered-with system from all n interference links in dB.

Then, when (m) numbers of simulations are run, the average aggregate interference, I_{ave}^m , can be calculated by the following equation:

$$I_{ave}^m = 10 \log \left(\frac{\sum_{k=1}^m 10^{I_{agg}^k/10}}{m} \right). \quad (11)$$

Finally, the interference margin between the 5G IoT system and the FSS system can be calculated as

$$D_{margin} = \left(\left(\frac{I}{N_{thr}} \right) + N \right) - I_{ave}^m, \quad (12)$$

where D_{margin} is the interference margin required by the 5G BS or IoT terminal, I/N_{thr} is a criterion of the interference-

noise ratio, which is presented in Table 5, and N is the thermal noise of the FSS ES. It is a common view that $((I/N_{thr}) + N)$ is also called an interference threshold, I_{thr} , in dBm/MHz.

3.3. Simulation Results. Figure 7 demonstrates the cumulative distribution functions (CDFs) of the antenna gain from the 5G IoT system towards the FSS ES, which describes the parameter $G_{tx}(\theta_{tx}^i)$ in equation (9). It is one of the most important factors that influence the final interference margin between the 5G IoT and FSS systems. More specifically, Figure 7(a) shows the CDF of the transmitted antenna gain from the BS, and Figure 7(b) displays the characteristics of the 5G IoT terminals' antenna gain in the direction of the receiver station while taking into account the beamforming antenna pattern by CDF. Moreover, from Figure 7(a), it is evident that the BS antenna gains are primarily distributed from approximately -75 to 25 dB, and a flat, increasing curve can be seen at around the 95% point of the CDF. In contrast, the distribution of the antenna gain of the IoT terminal mainly decreases from about -70 to 15 dB, but the trend at about the 95% point of the CDF increases more sharply than that of the BS. This indicates that it is possible to produce a higher aggregated gain with BSs because of the fewer number of antennas in the IoT terminals' antenna pattern.

Both Tables 6 and 7 demonstrate the results of the interference margin on links from different antenna patterns of a 5G BS to the ES. Table 8 presents the interference margin required by the IoT terminal. These results depend on different heights of the ES, isolation distances, and criteria of interference-noise ratios. The simulation details should be clarified. First, Monte Carlo simulations with 5000 iterations ($m = 5000$) can be considered due to the randomness of both the BS and IoT terminal positions in each sector. More importantly, these configurations of the BS and IoT terminals can be reused from Tables 6 and 7, respectively. In both these tables, varying

TABLE 5: Interference criteria for sharing study between 5G IoT and FSS system in 40 GHz.

Frequency ranges	Percentage of time for which the I/N_{thr} value could be exceeded (%)	I/N_{thr} criteria (dB)
37.5–51.4 GHz	20	-10
	0.001667	-1.3

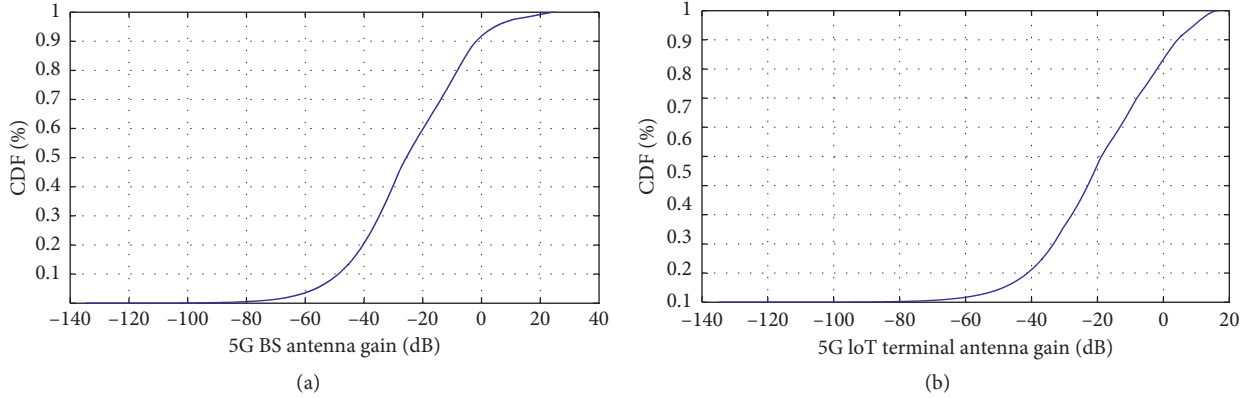


FIGURE 7: Transmitter antenna gain to the ES receiver. (a) 5G BS transmitter. (b) 5G IoT terminal transmitter.

TABLE 6: Interference margin from 5G BS (antenna pattern: 8×16) to FSS ES.

Interference criteria		Height of ES (m)	D_{margin} with isolation distance (L) km	
I/N_{thr}	I_{thr}		$L = 1$	$L = 0.4$
-10	-124.6	4.5	27.0	20.0
		10	27.1	20.8
-1.3	-115.9	4.5	32.4	27.6
		10	32.6	28.2

TABLE 7: Interference margin from 5G BS (antenna pattern: 16×16) to FSS ES.

Interference criteria		Height of ES (m)	D_{margin} with isolation distance (L) km	
I/N_{thr}	I_{thr}		$L = 1$	$L = 0.4$
-10	-124.6	4.5	26.4	19.3
		10	26.8	21.2
-1.3	-115.9	4.5	31.9	26.9
		10	32.3	28.5

TABLE 8: Interference margin from 5G IoT terminal (antenna pattern: 4×4) to FSS ES.

Interference criteria (dB)		Height of ES (m)	D_{margin} with isolation distance (L) km	
I/N_{thr}	I_{thr}		$L = 1$	$L = 0.4$
-10	-124.6	4.5	42.5	36.7
		10	42.4	36.9
-1.3	-115.9	4.5	48.0	44.0
		10	47.8	44.1

values of I/N_{thr} refer to the classifications of different cases, in which the best and normal cases are simulated.

Obvious differences in the results can be found by comparing the 8×16 and 16×16 antenna patterns in Tables 6 and 7. Although the transmission power of the 16×16 antenna pattern is approximately only 1.3 dB less than that of the 8×16 pattern, a higher transmission power of 3 dB can be found in the BS configured with the 16×16 antenna pattern because the number of antennas is two times greater than the BS configured with the 8×16 antenna pattern. The trend indicates that massive MIMO technology in mmWave can be seen as a mitigation method that sharply drops the power from the interference link. More specifically, the worse the interference margin, the greater the number of antenna elements; however, the degree of the decreasing trend in the margin is slight. The main reason for this is that, in practice, both narrower beams and more complete space diversity result in the decline of interference power.

On the other hand, it is a common view that the longer the isolation distance, the better the necessary interference margin; a gain of around 4 to 7 dB at a 1 km distance can be seen by comparing the cases of 0.4 km isolation distance in Tables 7 and 8. Moreover, the interference margin in the normal case ($I/N_{\text{thr}} = -10$ dB) is considerably lower than that in the best case ($I/N_{\text{thr}} = -1.3$ dB). However, it is a significantly low probability that the best case will appear most times. Furthermore, a higher ES also slightly promotes the value of the margin, which means that improving the height is almost meaningless. The main reason for this is that the difference of varying heights of ES is not enough to effectively block the power of each interference link from the 5G BSs and IoT terminal, although there is an increase in the ES height of more than two times. Finally, compared with the interference links from IoT terminal to ES, the interference margins of the BSs are worse; losses of around 15.2 dB can be observed.

From these three tables above, all the results of interference margin indicate that the 5G IoT system and FSS system can coexist without generating intolerance interference to the FSS system. However, more serious interference from the 5G BSs, not the IoT terminals, is also indicated in the tables due to the higher transmission power and the greater number of antennas in the sharing study between the 5G IoT system and the FSS system.

4. Conclusions and Future Research

This paper performed a detailed analysis of a coexistence scenario for the 5G IoT system in mmWave band, namely co-channel sharing with FSS downlink at 40 GHz, depending on varied separation distances, heights of ES, and interference criteria. The first part of the paper classified the interference links as BS to ES, or IoT terminal to ES, and discussed a sharing study scenario in 40 GHz. Additionally, both the 5G IoT and FSS system models and their related antenna patterns were researched. It was then demonstrated that the 5G IoT system can totally meet the interference protection criteria of the FSS system by transmitting from at least several hundred BSs and thousands of IoT terminals simultaneously. The final results also proved that the proposed massive MIMO technology can effectively suppress the interference at the ES while maintaining operable performance of the 5G IoT system in mmWave.

Notably, this type of coexistence analysis is not only essential but also mandatory to confirm whether the 5G IoT system can intelligently share spectrum with other incumbent systems, such as the FSS. This research is required for the introduction of new frequency bands during the 5G period and can steer how the 5G IoT system with mmWave can be deployed in the future.

The following points are of interest for our future research directions. First, it is a natural and significant step to conduct similar coexistence analyses for higher-potential spectrum bands, such as those above 52.6 GHz. In addition, the advanced antenna solutions and mitigation schemes are indispensable to obtain better interference margins in spectrum sharing studies. Finally, it is also important to consider a 5G IoT system coexisting with other services, such as WiFi.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] L. J. Ippolito, *Satellite Communications Systems Engineering: Atmospheric Effects, Satellite Link Design and System Performance*, Wiley Telecommunications, Hoboken, NJ, USA, 2017.
- [2] X. Meng, J. Li, D. Zhou, and D. Yang, "5G technology requirements and related test environments for evaluation," *China Communications*, vol. 13, no. 2, pp. 42–51, 2016.
- [3] S. Sun, M. Kadoch, L. Gong, and B. Rong, "Integrating network function virtualization with SDR and SDN for 4G/5G networks," *IEEE Network*, vol. 29, no. 3, pp. 54–59, 2015.
- [4] M. R. Palattella, M. Dohler, A. Grieco et al., "Internet of things in the 5G era: enablers, architecture, and business models," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 3, pp. 510–527, 2016.
- [5] Y. Wu, B. Rong, K. Salehian, and G. Gagnon, "Cloud transmission: a new spectrum-reuse friendly digital terrestrial broadcasting transmission system," *IEEE Transactions on Broadcasting*, vol. 58, no. 3, pp. 329–337, 2012.
- [6] S. Kim and S. Kota, "ITU-R standardization of fixed satellite services (FSS)," in *Proceedings of the International Workshop on Satellite & Space Communications IEEE*, pp. 247–252, Salzburg, Austria, September 2007.
- [7] S. Kim, E. Visotsky, P. Moorut, K. Bechta, A. Ghosh, and C. Dietrich, "Coexistence of 5G with the incumbents in the 28 and 70 GHz bands," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 6, pp. 1254–1268, 2017.
- [8] H. Shuai, X. Sai, M. Weixiao, and C. Li, "Dense-device-enabled cooperative networks for efficient and secure transmission," *IEEE Network*, vol. 32, no. 2, pp. 100–106.
- [9] H. Shuai, X. Sai, M. Weixiao, and C. Li, "An agile confidential transmission strategy combining big data driven cluster and OBF," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 11, pp. 10259–10270.
- [10] W. A. Hassan, H. S. Jo, and T. A. Rahman, "The feasibility of coexistence between 5G and existing services in the IMT-2020 candidate bands in Malaysia," *IEEE Access*, vol. 5, pp. 14867–14888, 2017.
- [11] F. Guidolin and M. Nekovee, "Investigating spectrum sharing between 5G millimeter wave network and fixed satellite system," in *Proceedings of the IEEE Globecom Workshops*, pp. 1–7, San Diego, CA, USA, December 2015.
- [12] H. Shuai Han, Z. Yi, M. Weixiao, and H.-H. Chen, "Self-interference-cancellation-based SLNR precoding design for full-duplex relay-assisted system," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 9, pp. 8249–8262, 2018.
- [13] J. Kim, L. Xian, and A. S. Sadri, "Numerical simulation study for frequency sharing between micro-cellular systems and fixed service systems in millimeter-wave bands," *IEEE Access*, vol. 4, no. 99, pp. 9847–9859, 2016.
- [14] B. Rong, Y. Qian, K. Lu, H.-H. Chen, and M. Guizani, "Call admission control optimization in WiMAX networks," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 4, pp. 2509–2522, 2008.
- [15] W. Weidong, Z. Fei, H. Wei, B. Wang, and Y. Zhang, "Co-existence studies between LTE system and earth station of fixed satellite service in the 3400–3600 MHz frequency bands in China," in *Proceedings of the IEEE International Conference on Broadband Network and Multimedia Technology (IC-BNMT)*, vol. 1125, no. 1130, pp. 26–28, Beijing, China, October 2010.
- [16] W. Hong, K. H. Baek, Y. Lee, Y. Kim, and S.-T. Ko, "Study and prototyping of practically large scale mmWave antenna system for 5G cellular devices," *IEEE Communication Magazine*, vol. 52, no. 9, pp. 63–69, 2014.
- [17] X. Meng, Y. Sun, D. Li et al., "Evaluated methodologies and results of IMT-2020 cellular system on millimeters waves," *Journal of Beijing University of Posts and Telecommunication*, vol. 41, no. 6, pp. 1–5, 2018.
- [18] H. Shuai, H. Yiteng, M. Weixiao, C. Li, N. Xu, and D. Chen, "Optimal power allocation for SCMA downlink systems based on maximum capacity," *IEEE Transactions on Communications*, vol. 67, no. 2, pp. 1480–1489.

Review Article

Practical Aspects for the Integration of 5G Networks and IoT Applications in Smart Cities Environments

Daniel Minoli  and **Benedict Occhiogrosso**

DVI Communications, Inc. New York, NY, USA

Correspondence should be addressed to Daniel Minoli; daniel.minoli@dvicomm.com

Received 1 April 2019; Accepted 26 May 2019; Published 5 August 2019

Guest Editor: Xi Chen

Copyright © 2019 Daniel Minoli and Benedict Occhiogrosso. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Both 5G cellular and IoT technologies are expected to see widespread deployment in the next few years. At the practical level, 5G will see initial deployments in urban areas. This is perhaps fortuitous from an IoT perspective, since many “mainstream” applications of IoT will support Smart Cities, Smart Campuses and Smart Buildings. Bandwidth demand for a number of Smart City applications is the main driver for enhanced mobile broadband (eMBB)-based 5G services in general, and new-generation 5G IoT applications, in particular. In turn, the use of the millimeter wave spectrum is required to enable 5G cellular technologies to support high data rates. Millimeter wave solutions, however, impose a requirement for small cells. Generally, an implementer tries to use one or a small handful of IoT technologies; preferably, and for managerial simplicity, the implementer would want to use a cellular/5G IoT technology for all nodes, whether indoors or outdoors, instead of a heterogeneous mix of various IoT technologies that have evolved over the years. This overview paper discusses a number of practical issues related to 5G-based IoT applications, particularly in Smart City environments, including the need for small cells, the transmission issues at millimeter wave frequencies, building penetration issues, the need for Distributed Antenna Systems, and the near term introduction of pre-5G IoT technologies such as NB-IoT and LTE-M, these being possible proxies for the commercial deployment and acceptance of 5G IoT.

1. Background

As the second decade of the 21st century comes to a close, we are witnessing an expansion of urban ecosystems, as populations continue to sustain the transition from rural and some suburban areas into large urban areas, driven by economic opportunities, demographic shifts and generational preferences. Seventy percent of the human population is expected to live in cities by the year 2050, and there already are more than 400 cities with over one million inhabitants [1]. Societal movement of people is a basic human existence dynamic and is one of the key mechanisms that drives the growth of cities. Yet, especially in the Western World, cities often have aging infrastructure, including roads, bridges, tunnels, rail yards, and power distribution plants. It follows that new technological solutions are needed to optimize the increasingly-scarce infrastructure resources, especially given the population growth and the limited financial resources available to most cities and municipalities. When cities

deploy state-of-the-art Information and Communication Technologies (ICT) on a large-scale, including Internet of Things (IoT) technologies, they are referred to as being “Smart Cities”.

Livability, infrastructure management, asset management, traffic transportation and mobility, logistics, electric power and other utilities, and physical security are the key aspects of a city’s operation. IoT technologies offer the opportunity to improve resource management of many assets related to city life and urban Quality of Life (QoL), including Intelligent Transportation Systems (vehicular automation and traffic control), energy consumption, the flow of goods, smart buildings, space/occupancy management (indoors and outdoors), pollution monitoring (for example from automobile traffic, factories, incinerators, crematoria), resource monitoring and sensing, immersive services (including wearables and crowdsensing), physical security, sustainability, and the greening of the environment. Smart City IoT applications cover indoor and outdoor applications; they also span



FIGURE 1: Illustrative example of Smart City resources that can benefit from IoT in general and 5G cellular in particular.

stationary and mobile end-nodes and sensors. There is an extensive body of literature on this topic; some references of interest include but are certainly not limited to [2–12]. Up to the present, the IoT has been mostly utilized for supporting a large population of relatively low-bandwidth sensing devices, and where the sensing devices are typically deployed in stationary locations (e.g., electric meters, building management systems, meteorological weather stations). However, video-oriented applications that require streams ranging up to Ultra-High Definition resolution are emerging (e.g., surveillance, physical security). In the evolving IoT environment, the endpoint devices include environmental and situational sensors, vehicles, wearables, drones, robots, and Virtual Reality gear. In some applications IoT actuators are also utilized to control the physical ecosystem in response to a sensed set of data or some analytical calculation – for example, changing the barriers and signs on a road to reverse traffic lanes during the day; or, changing the parameters of a pump to control water or sewer flows. Figure 1 depicts graphically some of the common Smart City applications.

While a large number of definitions and descriptions of the IoT exist, this is one definitional/descriptive quote

from the author’s previous work, which we utilize here: “The basic concept of the IoT is to enable objects of all kinds to have sensing, actuating, and communication capabilities, so that locally-intrinsic or extrinsic data can be collected, processed, transmitted, concentrated, and analyzed for either cyber-physical goals at the collection point (or perhaps along the way), or for process/environment/systems analytics (of a predictive or historical nature) at a processing center, often “in the cloud”. Applications range from infrastructure and critical-infrastructure support (for example smart grid, Smart City, smart building, and transportation), to end-user applications such as e-health, crowdsensing, and further along, to a multitude of other applications where only the imagination is the limit” [13–27]. According to the Global System for Mobile Communications Association (GSMA), between 2018 and 2025, the number of global IoT connections will triple to 25 billion, while global IoT revenue will quadruple to \$1.1 trillion [28]; others offer higher numbers (e.g., according to Statistica there will be about 80 billion IoT devices worldwide in 2025 [29]).

5G (5th Generation) is the term for the next-generation cellular/wireless service provider network that aims at

delivering higher data rates -- 100 times faster data speeds than the current 4G Long Term Evolution (LTE) technology -- lower latency, and highly-reliable connectivity. In a sense, it is an evolution of the previous generations of cellular technology.

Smart Cities do not depend on any unique or specific IoT technology per se, but include a panoply of IoT technologies, such as mission-specific sensors, appropriate networks, and function-and-use-efficient analytics, these often in the cloud. Wireless connectivity plays an important role in the utility of this technology, especially at the geographic scope of a large or even medium-size city. For practical reasons wireless is also important in Smart Campus and Smart Building applications. Table 1 identifies a number of Smart City challenges and needs, possible IoT-based solutions, wireless requirements, and the applicability of 5G solutions. 5G IoT is *licensed cellular IoT*. In this table “low bandwidth” equates to 200 kbps or less, “medium bandwidth” equates to 200 kbps to 2 Mbps, and “high bandwidth” equates to more than 2 Mbps. Some IoT applications entail periodic “batch” communication while other applications require real-time communication; in the table “low latency” means real-time and “medium latency” means 1-to-5 seconds. Table 2 provides a snapshot of key wireless technologies that are applicable to the IoT environment. A number of wireless technologies are available, each with its specific applicability and functionality. The direct use of traditional cellular services (e.g., 4G/LTE networks) is not optimal for IoT applications, both for cost and nodal power-consumption reasons. Furthermore, these services are not ideal for a number of IoT applications where a small amount of data is transmitted infrequently (e.g., electric, gas, or water meters for reading). Node density is also an issue. Cellular IoT solutions endeavor to address low-power, low data rate requirements. Several iterations and alternatives solutions have emerged in recent years (e.g., Cat1/Rel 8, Cat 0/Rel 12, Cat-M/Rel 13, EC-GSM, and NB-IoT/Rel 13). The 5G IoT system is the next evolutionary step, perhaps also affording some simplification and technology homogeneity.

Figure 2 depicts the pre-5G and the 5G IoT connectivity ecosystem, which is further elaborated in the rest of this paper. The figure illustrates a typical case of Wi-Fi (in-building) aggregation of sensor data for a handoff to the cloud over a traditional router; it illustrates the use of Low Power Wide Area Network (LPWAN) overlay technologies such as LoRa and Sigfox; it shows the use of pre-5G IoT technologies; and then illustrates the use of 5G IoT in a native mode, or in a more realistic Distributed Antenna System (DAS)-assisted mode.

This review, position and assessment paper provides an overview of salient 5G features and then discusses some practical design issues applicable to the IoT. A lot of the important 5G IoT information is included in the figures and tables. This paper is not intended to be a full 5G overview per se, nor a discussion of IoT, for both of which there are many references (e.g., [30–34] for 5G, and close to one hundred books on the IoT topic alone).

2. 5G Concepts and Technology

5G cellular networks are now starting to be deployed around the world, as the underlying standards and the system-wide technology become more mature (the term “*International Mobile Telecommunications-2020 [IMT-2020]*” is also used by the standards bodies.) Industry observers predict that societal developments will lead to changes in the way communication systems are used and that these developments will, in turn, lead to a significant increase in mobile and wireless traffic volume; such traffic volume is expected to increase a thousand-fold over the next decade. Observations such as this one are common in the literature positioning the technology: “*Unlike previous generations of mobile networks, the fifth generation (5G) technology is expected to fundamentally transform the role that telecommunications technology plays in the society*” [34].

The 5G system expands the 4G environment by adding New Radio (NR) capabilities, but doing so in such a manner that LTE and NR can evolve in complementary ways. As it might be envisioned, a 5G system entails devices connected to a 5G access network, which in turn is connected to a 5G core network. The 5G access network may include 3GPP (3rd Generation Partnership Project) radio base stations and/or a non-3GPP access network. The 5G core network offers major improvements compared with a 4G system in the area of network slicing and service-based architectures (SBAs); in particular, the core is designed to support cloud implementation and the IoT. 5G systems subsume important 4G system concepts such as the energy saving capabilities of narrowband IoT (NB-IoT) radios, secure low latency small data transmission for low-power devices -- low latency is a requirement for making autonomous vehicles safe -- and devices using energy-preserving dormant states when possible. Network slicing allows service providers to deliver “Network as a Service (NaaS)” to large/institutional users affording them the flexibility to manage their own services and devices on the 5G provider’s network.

Applications driving wireless traffic include but are not limited to on-demand mobile information and high-resolution entertainment, augmented reality, virtual reality and immersive services, e-health, and ubiquitous IoT roll-outs. While 5G technology could still take several distinct service directions, it appears at this juncture that the view favoring a super-fast mobile network, where densely-clustered small cells provide contiguous urban coverage to mobile as well as stationary users, is the approach envisioned by the standards development bodies and by the implementers. It should be noted that in the U.S., upwards of 55 percent of residential users now utilize cellular-services-only at home in place of a landline, and about 30 percent of residential users utilize both, with the trend favoring an eventual transition to the former. Therefore, the evolving 5G systems will have to properly support this growing segment of the market. A goal of 5G networks is to be five times as fast as compared to the highest current speed of existing 4G networks, with download speeds as high as 5 Gbps -- 4G offering only up to a maximum of 1 Gbps. Deployment of 5G networks started in 2018 in some advanced countries, although further developments on fundamentals will continue; naturally, the current

TABLE I: Key Urban Challenges and IoT-supported Solutions.

Smart City Issue and Requirements	IoT support/solutions	Indoors wireless needed	Outdoors wireless needed	5G applicability	Bandwidth / latency/ reliability
Infrastructure and real estate management.	Networked sensors (possibly including drones) to provide real-time and historical trending data allowing city agencies to provide enhanced visibility into the performance of resources, facilitating environmental and safety sensing, smart parking and smart parking meters, smart electric meters, and smart building functionality.	Y	Y	High	Low/ Low/ Medium
Requirement: monitor status and occupancy of spaces, buildings, roads, bridges, tunnels, railroad crossings, and street signals.					
Livability.	Networked sensors (possibly including drones) to facilitate smart multi-modal transportation, information-rich environments, location-based services, real-time connectivity to health-monitoring resources (e.g., air quality).	Y	Y	High	Medium/ Medium/ Medium
Requirement: Quality of Life, expeditious access to services, efficient transportation, low delays, safety.					
Logistics.	Networked sensors (possibly including drones) to enable the streamlining of warehousing, transportation, and distribution of goods. Traffic management is a facet of such logistical support.	Y	Y	High	Medium/Medium/ High
Requirement: supplying city dwellers with fresh food, supplies, goods, and other materials.					
Physical security.	Networked sensors (possibly including drones and gunshot detection systems) to support IP-based surveillance video, license plate reading, gun-shot detection, bio-hazard and radiological contamination monitoring, face recognition, and crowd monitoring and control.	Perhaps	Y	High	High/ Low/ High
Requirement: security in streets, parks, stations, tunnels, bridges, trains, buses, ferries.					

TABLE I: Continued.

Smart City Issue and Requirements	IoT support/solutions	Indoors wireless needed	Outdoors wireless needed	5G applicability	Bandwidth / latency/ reliability
Power and other city-supporting utilities.					
Requirement: reliable flow of electric energy, gas, and water; optimized waste-management and sewer; safe storage of gasoline.	Smart Grid solutions and sensor-rich utility infrastructure	N	Y	High	Low/Medium/High
Traffic, transportation and mobility.					
Requirement: optimized traffic flow, low congestion, low latency and high expediency, low noise, minimal waste of fuel and CO ₂ emissions, safety.	Networked sensors to support traffic flow, driverless vehicles including driverless bus transit, and multi-modal transportation systems. For driverless vehicles sensors will allow high-resolution mapping, telemetry data, traffic and hazard avoidance mechanisms	N	Y	High	Medium-to-High/Low/High
Electric and other utility manhole monitoring.					
Requirement: Electric power manholes require monitoring to avoid and/or prevent dangerous situations	Cost-effective and reliable sensors are needed. Technology being investigated by Con Edison in New York city	N	Y	High	Low/Medium/High
Pollution monitoring.					
Requirement: monitor emission of dioxins, vaporized mercury, nanoparticles, radiation from factories, incinerators, urban crematoria, especially if these sources are close to train tracks or other wind-turbulence elements (e.g., canyons)	Networked sensors throughout town (or within 10 km of a point source) to monitor toxic, health-impacting emission from point sources including factories, generation plants (if any) and crematoria (if any) [35–46]	N	Y	High	Medium/Medium/High

TABLE I: Continued.

Smart City Issue and Requirements	IoT support/solutions	Indoors wireless needed	Outdoors wireless needed	5G applicability	Bandwidth / latency/ reliability
Environmental Monitoring. Requirements: monitor outdoor temperature, humidity and other environmental gases	Sensors to that can be placed in easy-to-deploy locations, e.g., atop existing Smart City light poles to continuously monitor temperature, humidity and other environmental gases	N	Y	High	Low/ Medium/Medium
Flood Abatement. Requirement: Flood and storm drainage control	Distributed ruggedized sensors to monitor Flood and storm drainage to provide early warning and fault detection	N	Y	High	Low/ Medium/ High
Smart City Lighting. Requirement: Conversion to LED lighting and ensuing control via IoT for weather conditions, phases of the moon, seasons, traffic occupancy, and so on	Cities spend large amounts of money yearly for street lighting (usually 1000 street lights per 10,000 inhabitants, and \$125 per year per light for 4662 hours of usage yearly and system amortization.) LED lighting requires 1/3 rd the amount of power for the same amount of luminance. Payback for conversion is now around 5-6 years. Sensors are needed for IoT-directed light management for weather conditions, phases of the moon, seasons, traffic occupancy, and so on	N	Y	High	Medium/ Medium/ Medium

TABLE 2: Key Wireless Technologies applicable to IoT.

Technology	Indoor usability	Outdoor usability	Basic aspects
5G	Yes, perhaps with Distributed Antenna Systems (DASs)	Yes, about 10-15 miles	<ul style="list-style-type: none"> (i) Evolving, not yet widely deployed (ii) Several bands, low latency, high sensor density (iii) Cellular network architecture (iv) Licensed spectrum, 0.01 Mbps in some implementations, battery ~10 years (v) Broadband features available for surveillance/multimedia (vi) Cost-effective (vii) Expected to be available worldwide (viii) Building penetration may need Distributed Antenna Systems (DASs)
NB-IoT (Narrowband IoT)	Yes	Yes, up to about 20miles	<ul style="list-style-type: none"> (i) Several bands, licensed spectrum (ii) LTE-based (iii) 0.1-0.2 Mbps data rates, battery ~10+ years (iv) Low cost, low modem complexity, low power, energy saving mechanisms (high battery life) (v) Does not require a gateway: sensor data is sent directly to the destination server (other IoT systems typically have gateways that aggregate sensor data, which then communicate with the destination server) (vi) Reasonable building penetration (improved indoor coverage) (vii) Large number of low throughput devices (up to 150,000 devices per cell)

TABLE 2: Continued.

Technology	Indoor usability	Outdoor usability	Basic aspects
LTE-M (Long-Term Evolution Machine Type Communications) Rel 13 (Cat M/Cat M)	Yes	Yes, about 10-20 miles	<ul style="list-style-type: none"> (i) Cellular network architecture, LTE compatible, easy to deploy, new cellular antennas not required (ii) Uses 4G-LTE bands below 1GHz, licensed spectrum (iii) Considered the second generation of LTE chips aimed at IoT applications (iv) Caps maximum system bandwidth at 1.4 MHz (as opposed to Cat-0's 20 MHz), thus is cost-effective for Low Power Wide Area Network (LPWAN) applications such as smart metering, where only small amount of data transfer is required. (v) 1 Mbps upload/download, battery ~10 years (vi) Relatively low complexity and low power modem (vii) Can be used for tracking moving objects (Location services provided through cell tower mechanisms)
LoRa	Yes	Yes (6-15 miles with LOS)	<ul style="list-style-type: none"> (i) Band below 1 GHz (ii) IoT-focused from the get-go (iii) Proprietary (iv) Low power
Sigfox	Somewhat limited	Yes (30 miles in rural environments; 1-6 miles in city environments)	<ul style="list-style-type: none"> (i) Band below 1 GHz (ii) Narrowband (iii) Low power (iv) Star topology

TABLE 2: Continued.

Technology	Indoor usability	Outdoor usability	Basic aspects
Wi-Fi	Yes, 300 feet	To some degree, requires inter-spot connectivity backbone (wired or wireless) (e.g., 802.11ah: distance range up to about 1/2 mile)	<ul style="list-style-type: none"> (i) Several bands (ii) In 2018 the FCC allowed the expansion of the 6 GHz band to next-generation Wi-Fi devices with 1.2 GHz of additional spectrum spanning 5.925 to 7.125 GHz (current Wi-Fi networks operate at 2.4 GHz and 5 GHz with a few vendors offering 60 GHz “WiGig”, this having a range of 30 feet – IEEE 802.11ad and IEEE 802.11ay.) (iii) High adoption; most (but not all) indoor IoT utilize Wi-Fi; good functionality (iv) Free “air time” (v) Subject to interference: malicious or non-malicious interference (e.g., too many hot spots) could impair the sensor from sending data either on a fine-grain or coarse-grain basis
Bluetooth	Yes, 30 feet	No (or for Personal Area only)	<ul style="list-style-type: none"> (i) Low bandwidth (2 Mbps) (ii) Used in medical devices and industrial sensors Low power, good for wearables (iii) Usable for Real time location systems with medium accuracy
Zigbee	Yes (30-300 feet)	No (or for Personal Area only)	<ul style="list-style-type: none"> (i) Low data rate (ii) Industrial and some home applications (e.g., home energy monitoring, wireless light switches) (iii) Low transmit power/Low battery consumption

Note: A few other legacy IoT wireless technologies exist (e.g., Cat 1, EC-GSM, Weightless) but are not included in this table

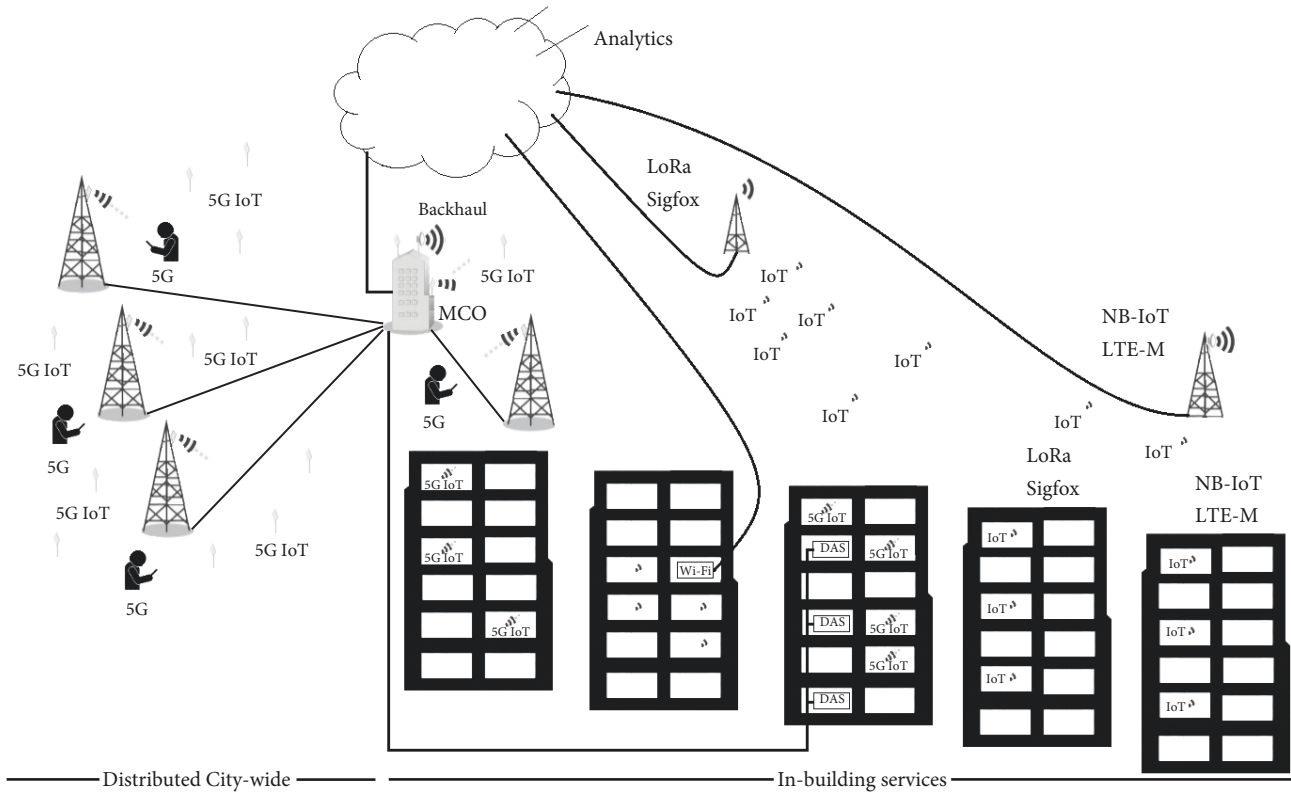


FIGURE 2: The pre-5G and the 5G IoT connectivity ecosystem.

4G/LTE and 5G are expected to coexist for many years. However, it is fair to say that like many other technologies before 5G, this technology is probably going through a “hype-cycle”, where a technology is supposed to be “all things to all people” and be the “be-all-and-end-all technology”; both claims will be abrogated in time. Proponents argue that 5G will “maximize the satisfaction of end-users by providing immersiveness, intelligence, omnipresence, and autonomy”.

2.1. 5G Standardization and Use Cases. Standardization work for 5G systems has been undertaken by several international bodies with the goal of achieving one unified global standard. Many well-known research centers, universities, standards bodies, carriers and technology providers have been involved in advancing the development of the technology for a 2020 rollout, including the Internet Engineering Task Force (IETF), the Open Network Automation Platform (ONAP), the GSMA, and the European Telecommunications Standards Institute Network Function Virtualization (ETSI NFV). In particular, work on 5G requirements, services and technical specifications has been undertaken in the past few years by three key entities: (i) International Telecommunication Union-Radio Communication Sector (ITU-R) [30], (ii) Next Generation Mobile Networks (NGMN) Alliance [31], and (iii) the 3rd Generation Partnership Project (3GPP) [32]. The ITU-R has assessed usage scenarios in three classes: ultra-reliable and low-latency communications (URLLC), massive machine-type communications (mMTC), and enhanced mobile broadband (eMBB). eMBB is probably the earliest

class of services being broadly supported and implemented. Key performance indicators are identified for each of these classes, such as spectrum efficiency, area traffic capacity, connection density, user-experienced data rate, peak data rate, and latency, among others. The ability to efficiently handle device mobility is also critical. Some examples of eMBB use cases include Non-SIM devices, smart phones, home/enterprise/venues applications, UHD (4K and 8K) broadcast, and virtual reality/augmented reality. mMTC use cases include smart buildings, logistics, tracking, fleet management and smart meters. URLLC cases include traffic safety and control, remote surgery, and industrial control. 5G systems are expected to support:

- (i) Tight latency, availability, and reliability requirements to facilitate applications related to video delivery, healthcare, surveillance and physical security, logistics, automotive locomotion, and mission-critical control, among others, particularly in an IoT context;
- (ii) A panoply of data rates, up to multiple Gbps, and tens of Mbps to facilitate existing and evolving applications, particularly in an IoT context;
- (iii) Network scalability and cost-effectiveness to support both clustered users with very high data rate requirements as well a large number of distributed devices with low complexity and limited power resources, particularly in an IoT context, where, as noted, a rapid increase in the number of connected devices is anticipated; and,

TABLE 3: Radio interface goals as defined in IMT-2020.

(i) MR for downlink peak data rate is 20 Gbps
(ii) MR for uplink peak data rate is 10 Gbps
(iii) Target downlink “user experienced data rate” is 100 Mbps
(iv) Target uplink “user experienced data rate” is 50 Mbps
(v) Downlink peak spectral efficiency is 30 bps/Hz
(vi) Uplink peak spectral efficiency is 15 bps/Hz
(vii) MR for user plane latency for eMBB is 4ms
(viii) MR for user plane latency for URLLC is 1ms
(ix) MR for control plane latency is 20ms (a lower control plane latency of around 10ms is encouraged)
(x) Minimum requirement for connection density is 1,000,000 devices per km ²
(xi) Requirement for bandwidth is at least 100 MHz
(xii) Bandwidths up to 1 GHz are required for higher frequencies (above 6 GHz)

MR = Minimal Requirement

Source: ITU-R SG05 Contribution 40, “Minimum requirements related to technical performance for IMT-2020 radio interface(s)”, Feb 2017.

- (iv) Pragmatic deployment cost metrics, along with acceptable service price points across the gamut of applications and data rates, particularly in an IoT context.

Specifically, some of the design details are a latency below 5 msec. (as low as 1 msec.), support for device densities of up to 100 devices/m², reliable coverage area, integration of telecommunications services including mobile, fixed, optical and MEO/GEO satellite, and seamless support for the IoT ecosystem. For example, the technical objective 5G as envisioned of METIS (*Mobile and Wireless Communications Enablers for the Twenty-twenty Information Society -- a European Community advocacy effort related to mobility*) are as follows [47–54]:

- (i) 1000 x higher mobile data volume per area than current systems;
- (ii) 10 to 100 x higher number of devices than current systems (i.e., dense coverage);
- (iii) 10 to 100 x higher user data rate than current systems (e.g., 1-20 Gbps);
- (iv) 10 x longer battery life for low power IoT devices than current systems (up to a 10-year battery life for machine type communications); and,
- (v) 5 x reduced end-to-end latency than current systems.

Table 3 defines the 5G radio interface goals as defined in IMT-2020. A number of these requirements are in fact being met (in various measure) by the systems now being deployed. The expectation is that to provide the full panoply of 5G services significant changes in both wireless technologies and core networks will be required.

As a point of observation, 3GPP/TR 22.891 has defined and/or described the following service groups: eMBB, Critical Communication, mMTC, Network Operations, and Enhancement of Vehicle-to-Everything (V2X). NGMN has defined and/or described the following service groups: Broadband access in dense area, Indoor ultra-high broadband access, Broadband access in a crowd, 50+ Mbps everywhere, Ultra low-cost broadband access for low ARPU areas,

Mobile broadband in vehicles, Airplanes connectivity, Massive low-cost Low long-range/low-power MTC, Broadband MTC, Ultra low latency, Resilience and traffic surge, Ultra-high reliability and Ultra low latency, Ultra-high availability and reliability, and Broadcast-like services.

Figure 3 depicts some of the key 5G services that can be utilized for the IoT, in the medium term in Smart Cities; other services shown might also be used over time. Although some have associated Smart Cities with mMTC, we are of the opinion that the early applications will be more within the eMBB domain (some others also agree [55]). Also, one would expect eMBB to be deployed more broadly, driven by the commercial “appeal” of the video services it facilitates. Augmented and/or virtual reality (AR/VR) are emerging as keys application of 5G networks, also involving some IoT aspects. To meet the requirements of lower latency and massive data transmission in AR/VR applications, software-defined networking (SDN) with a multi-path cooperative route (MCR) scheme that minimizes delay may be ideally positioned for 5G small cell networks [56]. Note parenthetically that video requirements range from about 8 Mbps for HD, 25 Mbps for UHD, 50 Mbps for 360-degree UHD video, 200 Mbps for 360-degree HDR (high dynamic range) video, and up to 1 Gbps for 6DoF/MPEG-I. The evolving MPEG-I Visual standard addresses visual technologies of immersive media; 360 video provides panoramic video texture projected onto a virtual shape surrounding the user’s head, from which the user visualizes a portion for an immersive video experience; 6DoF (6 Degrees of Freedom) supports movements along three rotation axes and three translations and presumes that full freedom of movement through the scene is possible [57]. 5G/eMBB may eventually support some (but not necessarily all) of these video applications, but these applications are well beyond the IoT applications discussed in this paper. IP-based video surveillance in Smart Cities that may be supported by IoT can operate rather well at the 0.384-2.5 Mbps bandwidth range.

Figure 4 highlights some technical features of 5G services that can be utilized for the IoT in Smart Cities in terms of data rates, latency, reliability, device density and so on. 5G IoT overcomes the well-known limitation of unlicensed LPWAN

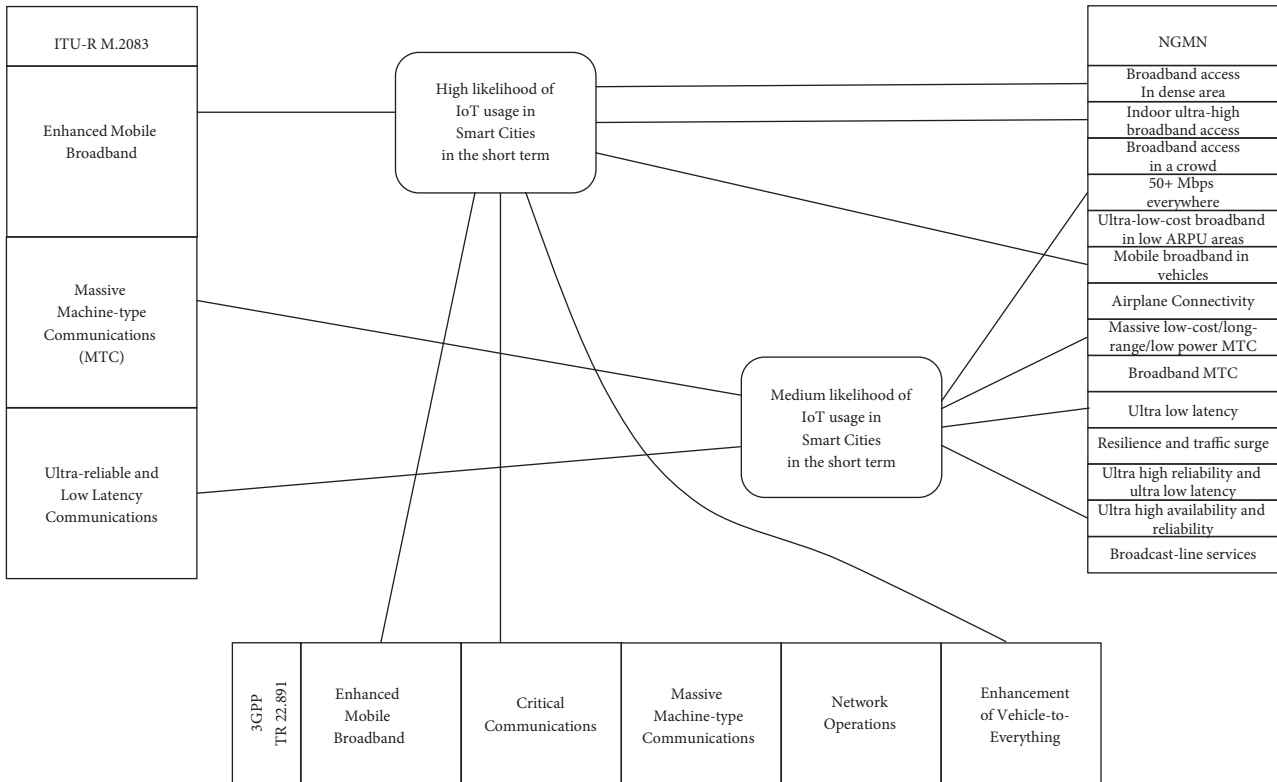


FIGURE 3: 5G services that can be utilized for the IoT in Smart Cities.

technologies that utilize crowded license-free frequency bands, especially in large cities; therefore, 5G IoT is ideal for Smart City for mission-critical and Quality of Service (QoS)-aware applications (for example, traffic management, smart grid, utility control.)

2.2. 5G Evolution. 3GPP has specified new 5G radio access technology, 5G enhancements of 4G (fourth generation) networks, and new 5G core networks. Specifically, it has defined a new 5G Core network (5GC) and a new radio access technology called 5G “New Radio” (NR). The new 5GC architecture has several new capabilities built inherently into it as native capabilities: multi-Gbps support, ultra-low latency, Network Slicing, Control and User Plane Separation (CUPS), and virtualization. To deploy the 5GC new infrastructure will be needed. There is a firm goal to support for “forward compatibility”. The 5G NR modulation technique and frame structure are designed to be compatible with LTE. The 5G NR duplex frequency configuration will allow 5G NR, NB-IoT and LTE-M subcarrier grids to be aligned. This will enable the 5G NR user equipment (UE) to coexist with NB-IoT and LTE-M signals. As might be expected, however, it is possible to integrate into 5G elements of different generations and different access technologies– two modes are allowed: the SA (standalone) configuration and the NSA (non-standalone) configuration (see Figure 5, also positioning IoT support).

(i) 5G Standalone (SA) Solution: in 5G SA an all new 5G packet core is introduced. SA scenarios utilize only one radio access technology (5G NR or the evolved

LTE radio cells); the core networks are operated independently.

(ii) 5G Non-Standalone Solution (NSA): in 5G NSA Operators can leverage their existing Evolved Packet Core (EPC)/LTE packet core to anchor the 5G NR using 3GPP Release 12 Dual Connectivity feature. This will enable operators to launch 5G more quickly and at a lower cost. This solution might suffice for some initial use cases. However, 5G NSA has a number of limitations, thus these Operators will eventually be expected to migrate to 5G Standalone solution. NSA scenario combines NR radio cells and LTE radio cells using dual-connectivity to provide radio access and the core network may be either EPC or 5GC.

Multiple evolution/deployment paths may be employed by service providers (service providers of various services, including IoT services) to reach the final target configuration; this migration could well take a decade, and may also have different timetables in various parts of a country, e.g., top urban areas, top suburban areas, secondary urban areas, secondary suburban areas, exurban areas, rural areas. Figure 6 depicts the well-known migration paths. The IoT implementer will need to be keenly aware of what 5G (5G IoT) services are available in a given area as an IoT implementation is contemplated. In Figure 6, Scenario 1 illustrates that the IoT Service provider will continue to use LTE and EPC to provide services (e.g., NB-IoT); here only legacy IoT devices can be supported. The provider only has a standalone radio

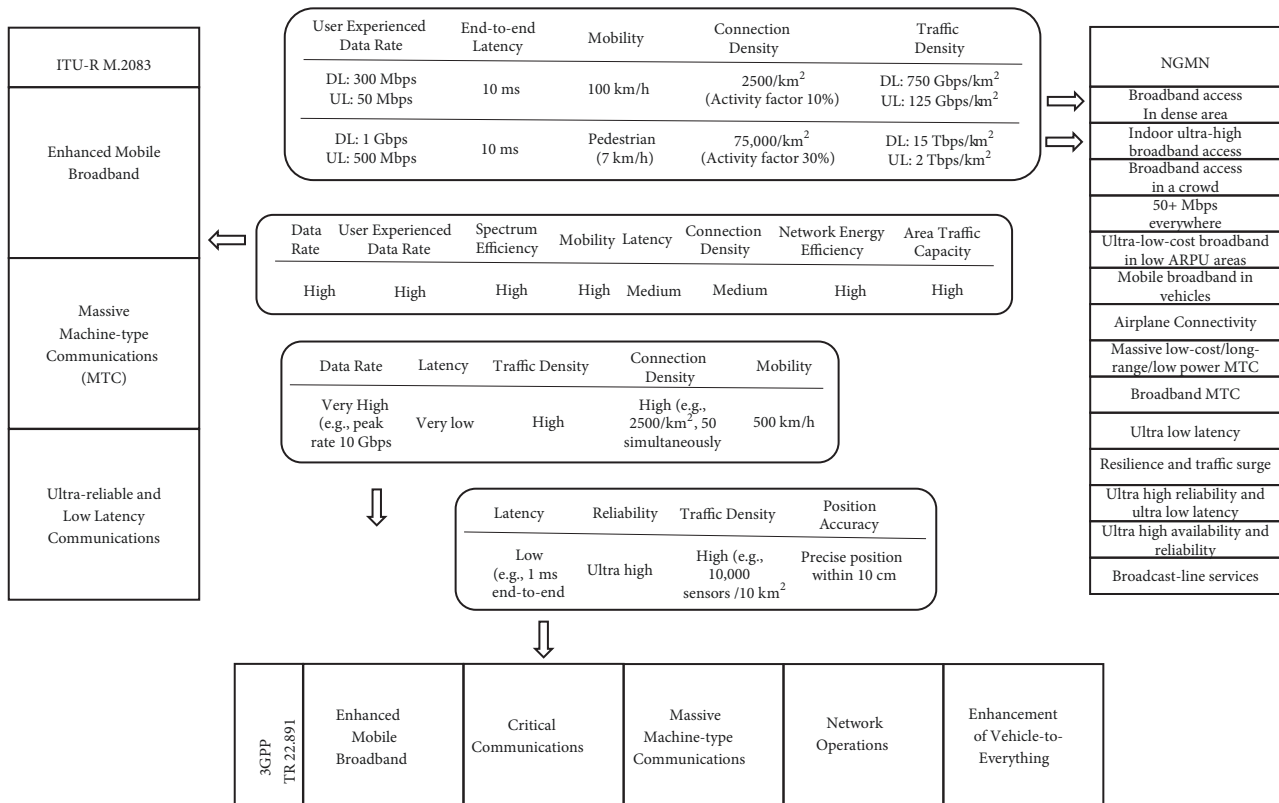


FIGURE 4: Some technical features of 5G services that can be utilized for the IoT in Smart Cities.

Core Network	5GC	New core transport
	EPC	Legacy core transport
Radio Access Network	SA	New IoT access
	NSA	Legacy IoT access

Core:

3GPP has defined a new 5G core network (5GC) and a new radio access Technology known as 5G “New Radio” (NR)

Access:

5G Standalone (SA) solution: In 5G SA an all new 5G packet core is introduced. SA scenarios utilize only one radio access technology (5G NR or the evolved LTE radio cells); the core networks are operated independently

5G Non-Standalone Solution (NSA): in 5G NSA, Operators can leverage their existing Evolved Packet Core (EPC)/LTE packet core to anchor the 5G NR using 3GPP Release 12 Dual Connectivity feature

FIGURE 5: 5G Transition Options and IoT support.

technology, in this case LTE only. Scenario 2 illustrates an IoT Service provider has migrated completely to NR (again only providing a standalone radio technology), but will retain the existing core network, the EPC. (Only) new 5G IoT devices can be used. In scenarios 5 and 6 the service providers will support both the legacy LTE and the new NR (clearly in this non-standalone arrangement, both radio technologies are deployed.) Some of these providers retain the legacy core and some will deploy the new 5GC core. Both legacy and 5G IoT devices can be supported.

3GPP approved the 5G NSA standard at the end of 2017 and the 5G SA standard in early 2018 in the context of its Release 15. Release 15 also included the support eMBB, URLLC, and mMTC in a single network to facilitate the deployment of IoT services; Release 15 also supports 28 GHz millimeter-wave (mmWave) spectrum and multi-antenna technologies for access.

2.3. 5G Frequency Bands. Focusing on the radio technology, there are number of spectrum bands that can be used in

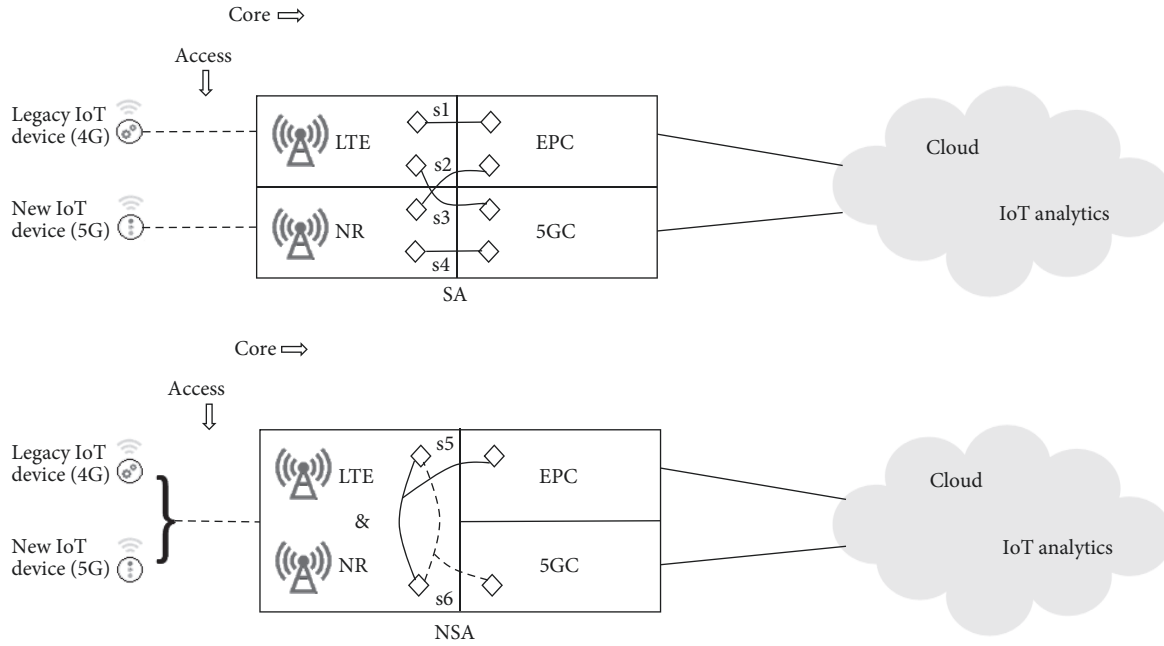


FIGURE 6: Detailed 5G Transition Options and IoT support.

5G; these bands can be grouped into three macro categories: sub-1 GHz, 1-6 GHz and above 6 GHz. The more advanced features, especially higher data rates require the use of the millimeter wave spectrum. New mobile generations are typically assigned new frequency bands and wider spectral bandwidth per frequency channel (1G up to 30 kHz, 2G up to 200 kHz, 3G up to 5 MHz, and 4G up to 20 MHz). Up to now cellular networks have used frequencies below 6 GHz. Generally, without advanced MIMO (Multiple In Multiple Out) antenna technologies one can obtain about 10 bits-per-Hertz-of-channel bandwidth. But the integration of new radio concepts such as Massive MIMO, Ultra Dense Networks, Device-to-Device, and mMTC will allow 5G to support the expected increase in the data volume in mobile environments and facilitate new IoT applications. Implementable standards for 5G are being incorporated in 3GPP Release 15 onwards. As noted, 3GPP Rel 15 defines New 5G Radio and Packet Core evolution to facilitate interoperable deployment of the technology.

The millimeter wave spectrum, also known as Extremely High Frequency (EHF), or more colloquially mmWave, is the band of electromagnetic spectrum running between 30 GHz and 300 GHz. Bands within this spectrum are being considered by the ITU and the Federal Communications Commission in the U.S. as a mechanism to facilitate 5G by supporting higher bandwidth. The use of a 3.5 GHz frequency to support 5G networks is also gaining some popularity, but the higher speeds networks will use other frequency bands, including millimeter-wave frequencies (these bands ranging from 28 GHz to 73 GHz, specifically the 28, 37, 39, 60 and 72-73 GHz bands). In the U.S., recently the FCC approved spectrum for 5G, including millimeter-wave frequencies in the 28 GHz, 37 GHz and 39 GHz bands, although these targeted cellular frequencies may nominally

overlap with other pre-existing users of the spectrum, for example point-to-point microwave paths, Direct Broadcast satellite TV, and high throughput satellite (HTS) systems (Ka-band transmissions).

Initially 5G will, in many cases, use the 28 GHz band, but higher bands will very likely be utilized later on; initial implementations, will support a maximum speed of 1 Gbps. Lower frequencies (at the so-called C band) are less subject to weather impairments, can travel longer distances, and penetrate building walls more easily. Waves at higher frequencies (Ku, Ka and E/V bands) do not naturally travel as far or penetrate walls or objects as easily. However, a lot more channel bandwidth is available in millimeter-wave bands. Furthermore, developers see the need for “an innovative utilization of spectrum”; “small cell” approaches are required to address the scarcity of the spectrum, but at the same time covering the geography. V band spectrum covers 57-71 GHz, which in many countries is an “unlicensed” band, and E band spectrum covers 71-76 GHz, 81-86 GHz and 92-95 GHz.

In the U.S., in 2018 the FCC also opened up, as an “interim” step for 5G, a “mid-band” radio spectrum at 3.5 GHz which was previously reserved for naval radar use. The 3.5 GHz band provides a combination of signal propagation distance, acceptable building penetration, and increased bandwidth. The FCC created 15 channels within the 3.550-3.700 GHz band, auctioning seven channels to “priority access licenses” and making eight channels available for general access -- the U.S. Navy still getting priority across the band when and as needed. With this approval, 5G devices can be built to support the same 3.5GHz ranges across North America, Europe, and Asia [58].

In addition to new bands, 5G technology is expected to use beam-forming and beam-tracking, where a cell’s antenna can focus its signal to reach a specific mobile device and

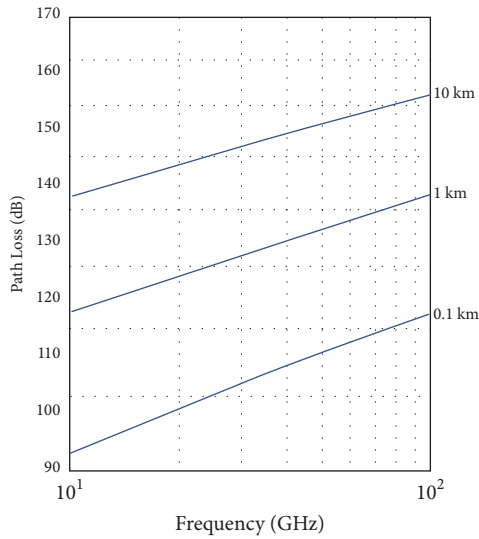


FIGURE 7: Path loss as a function of distance and frequency.

then track that device as it moves. Beamforming utilizes a large number (hundreds) of antennas at a base station to achieve highly directional antenna beams that can be “steered” in a desired direction to optimize transmission and throughput performance. Massive MIMO is a system where a transmission node (base station) is equipped with a large number (hundreds) of antennas that simultaneously serve multiple users; with this technology multiple messages for several terminals can be transmitted on the same time-frequency resource.

2.4. 5G Transmission Characteristics at Higher Frequencies. Due to RF propagation phenomena that are more pronounced at the higher frequencies, such as multipath propagation due to outdoor and indoor obstacles, free space path loss, atmospheric attenuation due to rain, fog, and air composition (e.g., oxygen), small cells will almost invariably be needed in 5G environments, especially in dense urban environments. Additionally, Line of Sight (LOS) will typically be required. ITU-R P series of recommendations has useful information on radio wave propagation including ITU-R P.838-3, 2005; ITU-R P.840-3, 2013; ITU-R P.676-10, 2013; and ITU-R P.525-2, 1994. Figures 7, 8, 9, and 10 highlight the issues at the higher frequencies, including the millimeter-wave frequencies. Figure 7 depicts the path loss as a function of distance and frequency. Figure 8 shows the attenuation as a function of precipitation and frequency. Figure 9 illustrates the attenuation as a function of fog density and frequency. Figure 10 depicts the attenuation as a function of atmospheric gases and frequency (notice high attenuation around 60 GHz).

In addition to the broad service requirements briefly highlighted in Table 3 (for example latency, user density distribution, etc.), there are specific IoT nodal considerations that have to be taken into account as one develops the next generation network. For example, IoT nodes typically are low-complexity devices and have limited on-board power. 5G systems have to take these restrictions and considerations

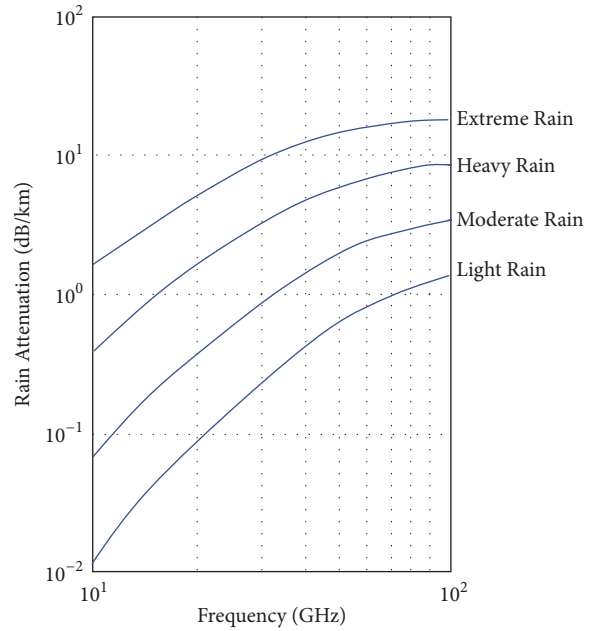


FIGURE 8: Attenuation a function of precipitation and frequency.

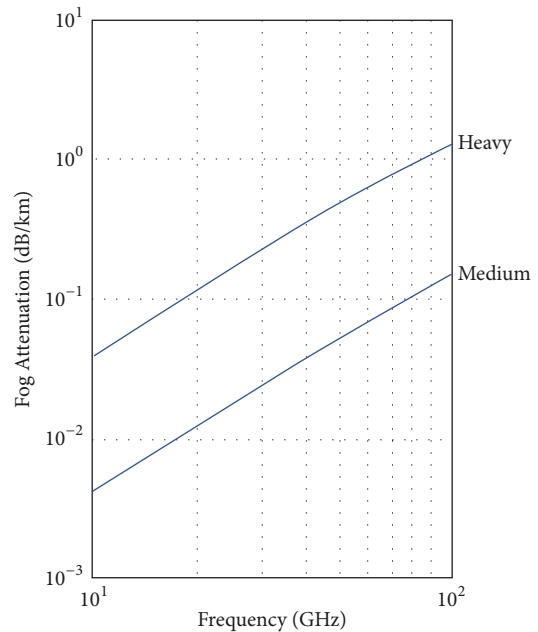


FIGURE 9: Attenuation a function of fog density and frequency.

into account. Table 4 provides a summary of some of these considerations and the 5G support.

3. Small Cell and Building Penetration Issues

As expected, communications at mmWave frequencies have attracted a lot of interest, due to the large available spectrum bandwidth that can potentially result in multiple gigabit per second transmissions per user. This follows a similar trend

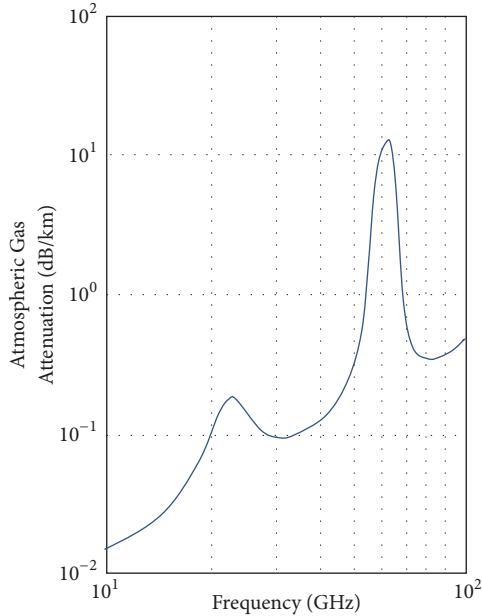


FIGURE 10: Attenuation a function of atmospheric gases and frequency. (notice high attenuation around 60 GHz).

in satellite communications with the introduction of Ka services, especially HTSs. High bandwidth will typically require a wide spectrum. Millimeter wave frequencies (signals with wavelength ranging from 1 millimeter to 10 millimeters) support a wide usable spectrum. The millimeter wave spectrum includes licensed, lightly licensed and unlicensed portions. Bandwidth demand and goals are the main driver for the need to use the millimeter wave spectrum, particularly for eMBB-based applications, allowing users to receive 100 Mbps as a bare minimum and 20 Gbps as a theoretical maximum. The use of millimeter wave frequencies, however, will imply the use of a much smaller tessellation of cells and supportive towers or rooftop transmitters due, as noted, to transmission characteristics, such as high attenuation and directionality. This is an important design consideration for 5G, especially in dense city/urban environments. The aggregation of these towers will by itself require a significant backbone network, whether a mesh based on some point-to-point microwave links, an fiber network or a set of “wireless fiber” links. Millimeter wave system utilize smaller antennas compared to systems operating at lower frequencies: the higher frequencies, in conjunction with MIMO techniques, can achieve sensible antenna size and cost. The millimeter wave technology can be utilized both for indoors and outdoors high-capacity fixed or mobile communication applications. The term “densification” is also used to describe the massive deployment of small cells in the near future.

MmWave products used for backhauling typically operate at 60 GHz (V Band) and 70/80 GHz (E Band) and offer solutions in both Point to Point and Point to Multipoint (PtMP) configurations providing end to end multi-gigabit wireless networks, for example, 1 Gbps up to 10 Gbps symmetric performance. Very small directional antennas, typically less than a half-square foot in area, are used to transmit and/or receive

signals which are highly focused beams. stationary radio systems are often installed on rooftops or towers. MmWave products are now appearing on the market targeting high capacity Smart City applications, 5G Fixed, Gigabit Wireless Access solutions, and Business Broadband. Urban canyons, however, may limit the utility of this technology to very short LOS paths. Mobile applications of mmWave technology are more challenging. On the other hand, one advantage of this technology is that short transmission paths (high propagation losses) and high directionality allow for spectrum reuse by limiting the amount of interference between transmitters and/or adjacent cells. Near LOS (NLOS) applications may be possible in some cases (especially for short distances).

Currently, mm wave frequencies are being utilized for high-bandwidth indoor applications, for example streaming (“miracasting”) of HD or UHD video and VR support (e.g., using 802.11ad Wi-Fi). Traditionally, these frequencies have not been used for outdoor broadband applications due to high propagation loss, multipath interference, and atmospheric absorption (gases, rain, fog and humidity) cited above; in addition, the practical transmission range is a few kilometers in open space [68]. Recently the FCC proposed new rules for wireless broadband in wireless frequencies above 24 GHz stating that it is “taking steps to unlock the mobile broadband and unlicensed potential of spectrum at the frontier above 24 GHz” [69]. The ITU and the 3GPP have defined two-phases of research; the first phase (expected to complete by press time) is to assess frequencies less than 40 GHz to address short-term commercial requirements; the second phase entails assessing the IMT 2020 requirements by studying frequencies up to 100 GHz. The following mmWave bands being considered, among other bands [70]:

- (i) 7 GHz of spectrum in total in the band 57 GHz to 64 GHz unlicensed.
- (ii) 3.4 GHz of spectrum in total in the 28 GHz/38 GHz licensed but underutilized region.
- (iii) 12.9 GHz of spectrum in total 71 GHz/81 GHz/92 GHz light-licensed band

Following the most recent World Radiocommunications Conference, the ITU also identified a list of proposed globally-usable frequencies between 24 GHz and 86 GHz, as follows: 24.25–27.5 GHz, 31.8–33.4 GHz, 37–40.5 GHz, 40.5–42.5 GHz, 45.5–50.2 GHz, 50.4–52.6 GHz, 66–76 GHz, and 81–86 GHz.

3.1. Cell Types. MmWave transmission will drive the requirement for small cells [71, 72]. “Small cells” refer to relatively low-powered radio communications equipment (base stations) and ancillary antennas and/or towers that provide mobile, internet, and IoT services within localized areas. Small cells typically have a range up to one-to-two kilometers, but can also be smaller -- on the other hand, a typical mobile macrocell (such as urban macro-cellular [UMa] or rural macrocell [RMa]) has a range of several kilometers up to 10-to-20 of kilometers.) The terms femtocells, picocells, micro-cells, urban microcell (UMi), and metrocells are effectively synonymous with the “small cells” concept. Small(er) cells

TABLE 4: Example of IoT nodal considerations for 5G systems.

IoT device issue	5G Support
Low complexity devices	Broad standardization leads to simplification e.g., SOC (System on a Chip) and/or ASIC (Application Specific IC) development
Limited on-board power	Technology allows a battery life ~10 years
Device mobility	Good mobility support in a cellular/5G system
Open environment	Broad standardization leads to broad acceptance of the technology
Devices universe by type and by cardinality	Standardized air interfaces can reduce certain aspects of the end-node, just like Ethernet simplified connectivity to a network, regardless of the functionality of the processor per se
Always connected/always on mode of operation	Cost-effective connectivity services allow the always on mode of operation
IoT security (IoTSec) concerns [59, 60]	Security capabilities are being added. The use of 256-bit symmetric cryptography mechanisms is expected to be fully incorporated. The encryption algorithms are based on SNOW 3G, AES-CTR, and ZUC; and integrity algorithms are based on SNOW 3G, AES-CMAC, and ZUC. The main key derivation function is based on HMAC-SHA-256. Identity management (e.g., via the 5G authentication and key agreement [5G AKA] protocol and/or the Extensible Authentication Protocol [EAP]), Privacy (conforming to the General Data Protection Regulation [GDPR]), and Security assurance (e.g., using Network Equipment Security Assurance Scheme [NESAS]) are supported. Some of these mechanisms are described [61–65]. As another example, the ETSI Technical Committee on Cybersecurity issued in 2018 two encryption specifications for access control in highly distributed systems such as G and IoT. Attribute-Based Encryption (ABE) that describes how to secure personal data.
Lack of agreed-upon end-to-end standards	Broad standardization possible with 5G if the technology is broadly deployed and is cost-effective
Lack of agreed-upon end-to-end architecture	Standardization at the lower layers (Data Link Control and Physical) can drive the development of a more inclusive multi-layer multi-application architecture

have been used for years to increase area spectral efficiency -- the reduced number of users per cell provides more usable spectrum to each user. However, the smaller cells in 5G are also dictated by the propagation characteristics. In the 5G context UMi typically have radii of 5-120 meters for LOS and 20 to 270 meters in NLOS; UMA typically have radii of 60-1000 meters for LOS and 50-1500 meters for NLOS [73]. Given their size, 5G/mmWave UMi cells will be able to support high bandwidth enabling eMBB services over small areas of high traffic demand. At the mmWave operation, user-device proximity with the antenna will enable higher signal quality, lower latency, and by definition, high data rates and throughput. Also, to be noted, mmWave frequencies make the size of multi-element antenna arrays practical, enabling large Multi-user MIMO (MU-MIMO) solutions.

Signal penetration indoors may represent a challenge, just as is the case even at present with 3G/4G LTE, even for traditional voice and internet access and data services. This has driven the need for DAS systems, especially in densely-constructed downtown districts. Free space attenuation at the higher frequency, power budgets, directionality requirements, and weather, all impact 5G and 5G IoT. Outdoor small cells and building-resident Distributed Antenna Systems (DAS) systems utilize high-speed fiber optic lines or “wireless fiber” to interconnect the sites to the backbone and the Internet cloud.

Figure 11 depicts a 5G IoT ecosystem where mmWave technology is used. Figure 12 shows typical (4G LTE) urban microcell towers. Figure 13 depicts a Smart City supported via (5G) urban microcells.

3.2. *Assessment of Transmission Issues.* Reference [74] provides a fairly comprehensive assessment of the transmission channel issues as they apply to 5G. The importance of this topic is accentuated by the large number of agencies actively researching this topic, including [55, 73–87]:

- (i) METIS
- (ii) 3GPPP
- (iii) MiWEBA (Millimetre-Wave Evolution for Backhaul and Access)
- (iv) ITU-R M
- (v) COST2100
- (vi) IEEE 802.11
- (vii) NYU WIRELESS: interdisciplinary academic research center
- (viii) IEEE 802.11 ad/ay
- (ix) QuaDRiGa (Fraunhofer HHI)
- (x) 5th Generation Channel Model (5GCM)

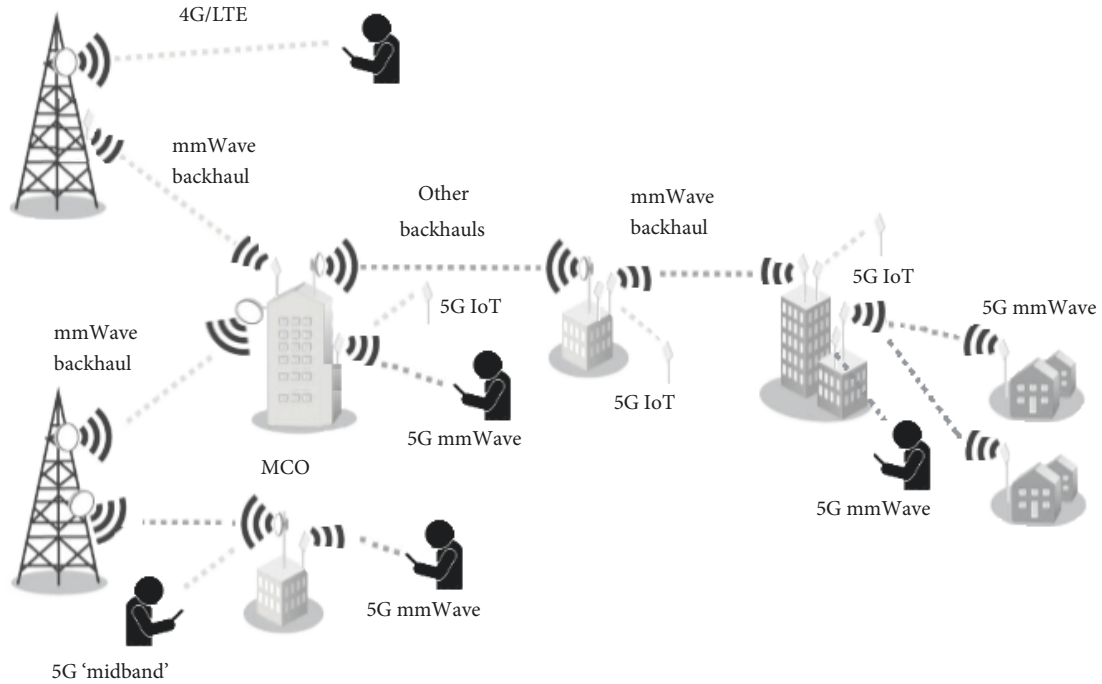


FIGURE 11: The 5G IoT ecosystems.

Microcell towers, usable in 5G and 5G IoT



FIGURE 12: Microcell towers (these for 4G, but a lot more for 5G). (non-copyrighted material from FCC-related filings [91]).

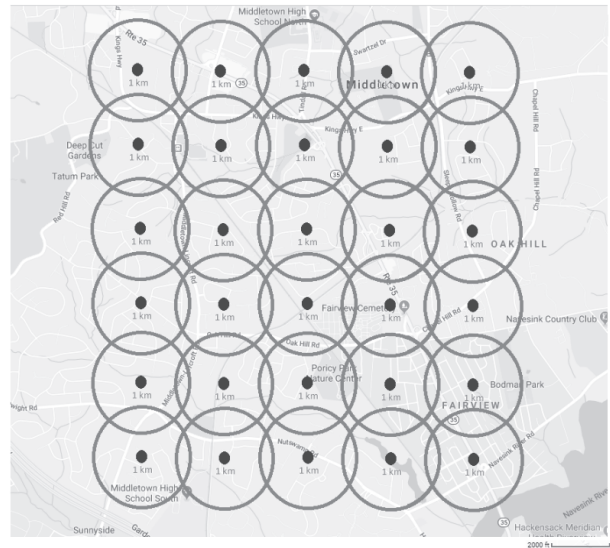


FIGURE 13: Microcells for 5G/5G IoT.

- (xi) 5G mmWave Channel Model Alliance (NIST initiated, North America based)
 - (xii) mmMAGIC (Millimetre-Wave Based Mobile Radio Access Network for Fifth Generation Integrated Communications) (Europe based)
 - (xiii) IMT-2020 5G promotion association (China based)
- (also including firms and academic centers such as, but not limited to AT&T, Nokia, Ericsson, Huawei, Intel/Fraunhofer

HHI, NTT DOCOMO, Qualcomm, CATT, ETRI, ITRI/CCU, ZTE, Aalto University, and CMCC.)
 Diffraction loss (DL) and frequency drop (FD) are just two of the path quality issues to be addressed. Although greater gain antennas will likely be used to overcome path loss, diffuse scattering from various surfaces may introduce large signal variations over travel distances of just a few centimeters, with fade depths of up to 20 dB as a receiver moved by a few centimeters. These large variations of the channel must be taken into consideration for reliable design

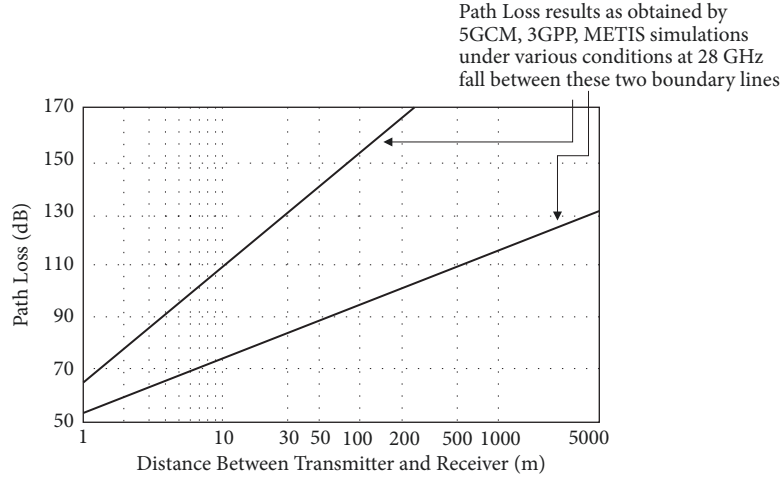


FIGURE 14: Path Loss simulations for 5G by various entities.

of channel performance, including beam-forming/tracking algorithms, link adaptation schemes, and state feedback algorithms. Furthermore, multipath interference from coincident signals can give rise to critical small-scale variations in the channel frequency response. In particular, wave reflection from rough surfaces will cause high depolarization. For LOS environment Rician fading of multipath components, exponential decaying trends and quick decorrelation in the range of 2.5 wavelengths have been demonstrated. Furthermore, received power of wideband mmWave signals has a stationary value for slight receiver movements but average power can change by 25 dB as the mobile transitions around a building corner from NLOS to LOS in an UMi setting. Additionally, human body blockage causes more than 40 dB of fading at the mmWave frequencies. Figure 14 depicts the path loss according to various simulations for 5G by various stakeholder entities.

The main parameter of the radio propagation model is the Path Loss Exponent (PLE), which is an attenuation exponent for the received signal. PLE has a significant impact on the quality of the transmission links. In the far field region of the transmitter, if $PL(d_0)$ is the path loss measured in dB at a distance d_0 from the transmitter, then the loss in signal power expected when moving from distance d_0 to d ($d > d_0$) is [88–90] is

$$PL_{d_0 \rightarrow d} (dB) = PL(d_0) + 10n \log_{10} \left(\frac{d}{d_0} \right) + \chi \quad (1)$$

$$d_f \leq d_0 \leq d$$

where

$PL(d_0)$ = Path Loss in dB at a distance d_0

n = PLE

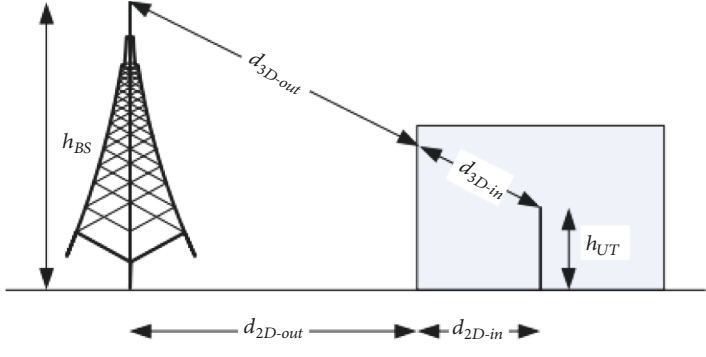
χ = A zero-mean Gaussian distributed random variable with standard deviation σ . (This is utilized only when there is a shadowing effect; if there is no shadowing effect, then this random variable is taken to be zero.)

See Figure 15. Usually PLE is considered to be known upfront but in most instances PLE needs to be assessed for the case at hand. It is advisable to estimate the PLE as accurately as possible for the given environment. PLE estimation is achieved by comparing the observed values over a sample of measurements to the theoretical values. Obstacles absorb signals, thus treating the PLE as a constant is not an accurate representation of the real environments, both indoors and outdoors (for example treating PLE as a constant which may cause serious positioning errors in complicated indoor environments [88]). Usually to model real environments the shadowing effects cannot be overlooked, by taking the PLE as a constant (a straight-line slope). To capture a shadowing effect a zero-mean Gaussian random variable with standard deviation σ is added to the equation. Here the PLE (slope) and the standard deviation of the random variable should be known precisely for a better modeling.

Table 5 provides theoretical performance equations developed by 3GPP and ETSI for outdoor channel performance [81]. As pragmatic working parameters, one has the following:

- (i) PLE values are in the 1.9 and 2.2 range for LOS and at the 28 GHz and 60 GHz bands; PLE is approximately 4.5 and 4.2 range for NLOS in the 28 GHz and 60 GHz bands.
- (ii) Rain attenuation of 2-20 dB/km can be anticipated for rain events ranging from light rain (12.5 mm/hr) to downpours (50 mm/hr) at 60 GHz (higher for tropical events). For 200-meter cells, the attenuation will be around 0.2 db for 5mm/hr rain at 28 GHz and 0.9 dB for 25mm/hr rain at 28 GHz. The attenuation will be around 0.5 db for 5mm/hr rain at 60 GHz and 2 dB for 25mm/hr rain at 60 GHz.
- (iii) Atmospheric absorption of 1-10 dB/km occurs at the mmWave frequencies. For 200-meter cells the absorption will be 0.04 dB at 28 GHz and 3.2 dB at 60 GHz.

TABLE 5: Path Loss Equations for mmWave 5G/5G IoT.



Scenario	LOS/NLOS	Pathloss [dB], (f_c is in GHz and d is in meters)	Shadow fading std [dB]	Applicability range, antenna height default values
UMi - Street Canyon	LOS	$PL_{UMi-LOS} = \begin{cases} PL_1 & 10\text{m} \leq d_{2D} \leq d'_{BP} \\ PL_2 & d'_{BP} \leq d_{2D} \leq 5\text{km}, \end{cases}$ see note 1 $PL_1 = 32.4 + 21 \log_{10}(d_{3D}) + 20 \log_{10}(f_c)$ $PL_2 = 32.4 + 40 \log_{10}(d_{3D}) + 20 \log_{10}(f_c) - 9.5 \log_{10}\left(\left(d'_{BP}\right)^2 + (h_{BS} - h_{UT})^2\right)$	$\sigma_{SF} = 4$	$1.5\text{m} \leq h_{UT} \leq 22.5\text{m}$ $h_{BS} = 10\text{m}$
	NLOS	$PL_{UMi-NLOS} = \max(PL_{UMi-LOS}, PL'_{UMi-NLOS})$ for $10\text{m} \leq d_{2D} \leq 5\text{km}$ $PL'_{UMi-NLOS} = 35.3 \log_{10}(d_{3D}) + 22.4 + 21.3 \log_{10}(f_c) - 0.3(h_{UT} - 1.5)$ Optional PL = $32.4 + 20 \log_{10}(f_c) + 31.9 \log_{10}(d_{3D})$	$\sigma_{SF} = 7.82$	$1.5\text{m} \leq h_{UT} \leq 22.5\text{m}$ $h_{BS} = 10\text{m}$
InH - Office	LOS	$PL_{InH-LOS} = 32.4 + 17.3 \log_{10}(d_{3D}) + 20 \log_{10}(f_c)$	$\sigma_{SF} = 3$	$1\text{m} \leq d_{3D} \leq 100\text{m}$
	NLOS	$PL_{InH-NLOS} = \max(PL_{InH-LOS}, PL'_{InH-NLOS})$ $PL'_{InH-NLOS} = 38.3 \log_{10}(d_{3D}) + 17.30 + 24.9 \log_{10}(f_c)$ Optional $PL'_{InH-NLOS} = 32.4 + 20 \log_{10}(f_c) + 31.9 \log_{10}(d_{3D})$	$\sigma_{SF} = 8.03$	$1\text{m} \leq d_{3D} \leq 86\text{m}$
			$\sigma_{SF} = 8.29$	$1\text{m} \leq d_{3D} \leq 86\text{m}$

Note 1: Breakpoint distance $d'_{BP} = 4h'_{BS}h'_{UT}f_c/c$, where f_c is the centre frequency in Hz, $c = 3.0 \times 10^8$ m/s is the propagation velocity in free space, and h'_{BS} and h'_{UT} are the effective antenna heights at the BS and the UT, respectively. The effective antenna heights h'_{BS} and h'_{UT} are computed as follows: $h'_{BS} = h_{BS} - h_E$, $h'_{UT} = h_{UT} - h_E$, where h_{BS} and h_{UT} are the actual antenna heights, and h_E is the effective environment height. For UMi $h_E = 1.0\text{m}$. For Uma $h_E = 1\text{m}$ with a probability equal to $1/(1 + C(d_{2D}, h_{UT}))$ and chosen from a discrete uniform distribution uniform $(12, 15, \dots, (h_{UT}-1.5))$ otherwise. With $C(d_{2D}, h_{UT})$ given by $C(d_{2D}, h_{UT}) = \{0, h_{UT} < 13\text{m}; ((h_{UT} - 13)/10)^{1.5} g(d_{2D}), 13\text{m} \leq h_{UT} \leq 23\text{m}\}$ where $g(d_{2D}) = \{0, d_{2D} \leq 18\text{m}; (5/4)(d_{2D}/100)^3 \exp(-d_{2D}/150), 18\text{m} < d_{2D}\}$
 3GPP TR 38.901 version 14.0.0 Release 14
 ETSI TR 138 901 V14.0.0 (2017-05).

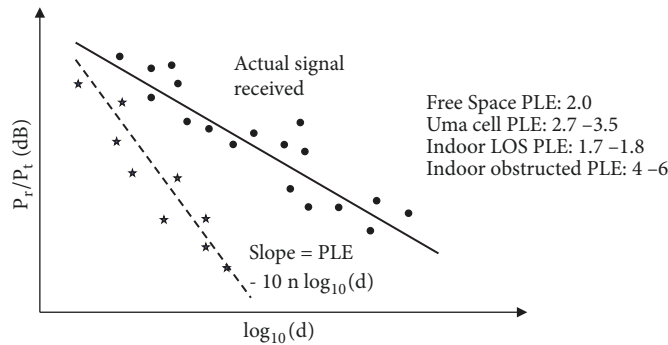


FIGURE 15: PLE.

Penetration into buildings is an issue for mmWave communication, this being a lesser concern for contemporary sub 1 GHz systems and even systems operating up to 6 GHz. O2I (Outdoor-to- Indoor) losses have to be taken into account. Actual measurements (e.g., at 38 GHz) demonstrated a penetration loss of 40 dB for brick pillars, 37 dB for a glass door, and 25 dB for a tinted glass window (indoor clear glass and drywall only had 3.6 and 6.8 dB of loss) [76]. This is why DASs are expected to be important for 5G in general and 5G IoT in particular.

3GPP and ETSI propose that the pathloss incorporating O2I building penetration loss be modelled as in the following [81]:

$$PL = PL_b + PL_{tw} + PL_{in} + N(0, \sigma_p^2) \quad (2)$$

where

PL_b is the basic outdoor path loss where d_{3D} is replaced by $d_{3D-out} + d_{3D-in}$

PL_{tw} is the building penetration loss through the external wall,

PL_{in} is the inside loss dependent on the depth into the building, and

σ_p is the standard deviation for the penetration loss

PL_{tw} is characterized as:

$$PL_{tw} = PL_{npi} - 10 \log_{10} \sum_{i=1}^N (p_i \times 10^{L_{material,i}/-10}) \quad (3)$$

where

PL_{npi} is an additional loss is added to the external wall loss to account for non-perpendicular incidence

$L_{material,j} = a_{material,j} + b_{material,j} \cdot f$, is the penetration loss of material i , example values below

Material	Penetration loss [dB]
Standard multi-pane glass	$L_{glass} = 2 + 0.2f$
IRR glass	$L_{IRRglass} = 23 + 0.3f$
Concrete	$L_{concrete} = 5 + 4f$
Wood	$L_{wood} = 4.85 + 0.12f$

Note: f is in GHz

p_i is proportion of i -th materials, where $\sum_{i=1}^N p_i = 1$; and N is the number of materials.

3GPP TR 38.901 version 14.0.0 Release 14

ETSI TR 138 901 V14.0.0 (2017-05)

In consideration of these propagation characteristics, many municipalities in the U.S. are concerned about the possible massive proliferation of small cells needed to support 5G. For example, a filing to the FCC was made in the U.S. late in 2018 by a consortium of towns known as the Communities and Special Districts Coalition in response to the Commission's September 5, 2018, Draft Declaratory Ruling and 3rd Report and Order where the FCC asserted the

claim that "small cell" deployment is a federal undertaking; furthermore the filing states that "*the massive deployment envisioned by the Commission raises substantial questions as to whether the Commission is in a position to assert that deployment is safe, given that its radio frequency emissions rules were based on technologies and deployment patterns that the Commission declares obsolete in this Order*" [74, 91]. Furthermore, it is unclear, according to the filing, what is the size of the equipment needed to support a small cell, since it could vary from a "pizza box" system to several racks that equate to 56 "pizza boxes" [91]. Although small cells will indeed need to be deployed to properly support 5G, caution is advocated. S&P Global Market Intelligence estimates that small-cell deployments reach approximately 850,000 in the U.S. by 2025 (with approximately 700,000 already deployed in 2019), with about 30% of small cell installations being outdoors; the same projection forecasts a total of 8.4 million small cells world-wide, with some regions of the world experiencing much higher deployments rates than in the U.S., e.g., doubling the 2019 numbers by the year 2025. These data show that placement within buildings is a common alternative (there will be more in-building systems than outdoor systems) [75].

4. 5G DAS for Indoor IoT Applications

The previous section discussed propagation issues at the higher frequencies. However, even the sub-6 GHz bands have issues penetrating buildings with the new building materials and infrared reflecting (IRR) glass. Indoor solutions are needed for IoT even at standard 3G/4G LTE frequencies and much more so at mmWave if cellular-based (5G) IoT transmission services for in-building applications are contemplated; outdoor 5G IoT applications do not.

Although it is in principle possible to support multiple access technologies in an IoT sensor (chipset), end-point IoT devices tend to have low complexity in order to achieve an established target price point and on-board power (battery) budget. Therefore a (large) number of applications will have devices that have a single implemented wireless uplink. It follows that -- either because of the goal of mobility support (for example, a wearable that works seamlessly indoors and in open spaces around town) or because of the designer's goal to utilize a single, consistent IoT nodal and access technology -- an all-sites wireless service for a Smart City application, is preferred. DASs may support such a goal (while city-wide Wi-Fi and/or Sigfox/LoRa could be an alternative, the ubiquity, standardization and cost-effectiveness of 5G cellular and IoT services, may well favor the latter in the future).

4.1. DAS Networks. A DAS is network of a (large) number of (small) (indoor or on-location) antennas connected to a common cellular source via fiber optic channel, providing cellular/wireless service within a given structure. DAS (sometimes also called in-building cellular) refers to the technology that enables the distribution and rebroadcasting of cellular, LTE, AWS, 5G and other RF frequencies within a building or confined/defined structural environment. While DAS is often used in large urban office buildings, DAS can also be used in

open spaces such as campuses, conference centers, stadiums, hospitals, airports, train stations, tunnels, hotels, cruise ships, and so on. DASs can and will support cellular-based IoT (e.g., LTE-M, NB-IoT, and 5G IoT.) Elements of a DAS include (see Figure 16):

- (i) (Small) Broadband antennas and amplifiers in the indoor space (typically one or more per floor) that shape the coverage. These antennas typically cover the entire spectrum of the cellular service (for/from multiple service providers);
- (ii) Coax or fiberoptic cabling to connect the structure antennas to a local Base Station;
- (iii) Remote Radio Head, a local Base Station, (“small cell”), typically in the basement; and,
- (iv) Fiberoptic connection to an aggregation point (typically in a carrier colocation space) (or the use of an outdoor donor antenna to a specific cellular provider). The former supports carrier-neutral applications, the latter typically supports only one carrier. Physical connectivity from the colocation space to each of the wireless providers is needed, typically in the form of fiber connectivity or other telecom service. Business relationships with the wireless providers are needed.

Current typical drivers include the fact that during anticipated peak times (whether in a building or in some public venue as a stadium) users will experience: coverage deficiencies, blocked connections, reduced data speeds, among other service deficiencies. Current systems support CDMA, EVDO, GSM, HSPA, UMTS, among others. Future systems will support 5G and become even more prevalent.

Given the mmWave transmission issues mentioned above (the small cells, the directionality, the free space loss and other attenuation factors) DASs will likely play a big role in 5G, both for regular voice and data services and for IoT. The large number of “small cells” cited earlier (8.4 million in 2025, with about 70% of these being considered to be indoors) supports the thesis that DASs will play a pivotal role in the future. They will be a key element of Smart City IoT support, especially for in-building sensors. As was shown in Figure 2, while a number of applications could use a Wi-Fi (or related) access technology with networked connection to the cloud, or a Sigfox/LoRa related solutions (these, however, being vendor proprietary), Smart City IoT service implementers may prefer to use a cellular service such as LTE-M or NB-IoT in the immediate future and 5G IoT as it becomes available, allowing a seamless and single-technology solution city-wide. In some cases, for example in smaller/older buildings and/or in suburbia and/or for buildings very close to a 5G cell tower, a direct 5G IoT connection may suffice. But for high-density urban and smart building applications, the use of DASs seems inevitable.

4.2. DAS Design. A single, carrier-neutral, consolidated system is often sought: a carrier-neutral system avoids multiplicity of antenna distribution, and sharing allows more coverage and higher capacity. A carrier-neutral DAS supports an end-use system, for example a smartphone, regardless of

which service provider the user is subscribed to. It would be rather expensive for a building owner to deploy a carrier-neutral DAS that supports a single building, unless it would be a very large building, campus, or installation. With carrier-neutral DAS arrangements the ownership of system is shifted from the building owner, or a specific cellular carrier to a third-party system provider, or a DAS integrator. Figure 17 depicts a typical carrier-neutral arrangement. Obtaining wireless carrier permission and coordinating between different wireless carriers is a key planning undertaking of any successful DAS rollout. Three scenarios are shown:

- (i) Scenario/Approach S1: The DAS integrator/provider wires up a remote building or space and drops a fiber link into an existing colo rack at an existing carrier-neutral provider, thus sharing all the Base Station Hotel (BSH) colo equipment and interfaces to the various wireless providers.
- (ii) Scenario/Approach S2: The DAS integrator/provider must build out the requisite base station equipment in the colo (the colo provider only provides power, rack space, HVAC, and so on). The DAS integrator/provider must also build interfaces to the wireless providers and secure business arrangements with them. The DAS integrator/provider builds out the remote buildings or venues.
- (iii) Scenario/Approach S3: The DAS integrator/provider must build out the requisite base station equipment in the colo, but the DAS integrator/provider can make use of existing interfaces and equipment to the various wireless providers. The DAS integrator/provider builds out the remote buildings or venues.

A less desirable approach is to use “donor antennas” (also shown in Figures 16 and 17). These antennas are installed on the roof of a building and are pointed at “donor” cell towers. Typically, a single cellular vendor is supported. The in-building arrangement is similar to that of a carrier-neutral arrangement, except that there typically will not be a remote base station: a combination of fiber optic cable, coaxial cable and in-building antennas is used to amplify and distribute those signals within a given space; coordination with the given carrier is still needed to make sure that the concentrated traffic is accepted by the provider.

5. 5G Deployment Snapshot

5.1. 5G Cellular Services. According to GSMA 5G is on track to account for 15% (1.4 billion) of global mobile connections by 2025. By early 2019, according to GSMA, eleven worldwide operators had announced initial 5G service launches and seven other operators had activated 5G base stations with commercial services to follow in the near future [92]. Select cities worldwide will have 5G by the end of 2019. See Table 6 for a summary of near-term 5G service-deployment activities. However, 4G services are expected to continue well past 2025: 4G will account for 59% of the connections, 3G for 20% of the connections, and 2G 5% of the connection (3G and 2G

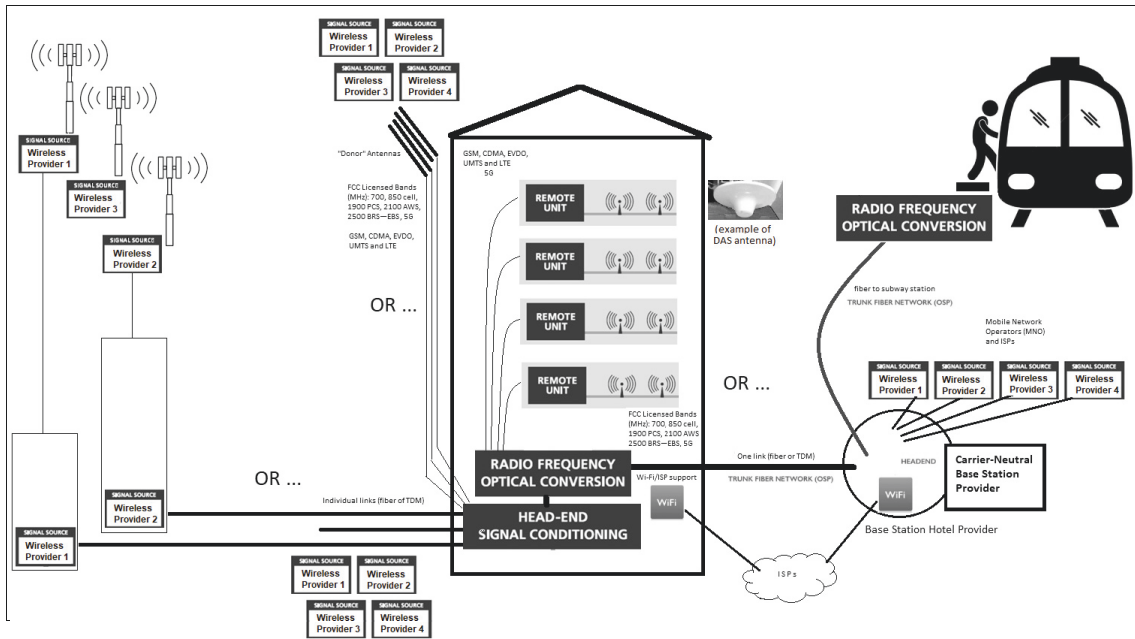


FIGURE 16: Elements of a DAS.

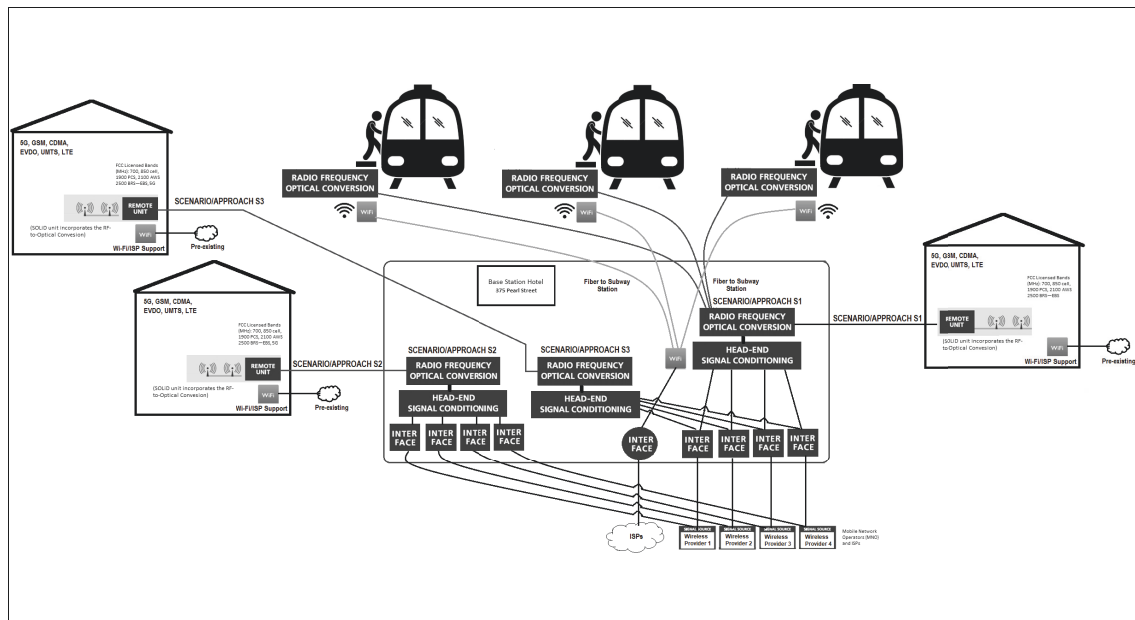


FIGURE 17: Carrier-neutral DAS.

are trending down through this period while 4G will continue to grow, but settle at around 60% by 2023 -- 5G is penetrating at approximate linear rate of CAGR around 1.5-2%). Also, according to GSMA 2019 will see 5G launches accelerate and devices hit the market:16 major markets worldwide will start to offer commercial 5G networks in 2019, following on from the first 5G launches in South Korea and the U.S. in 2018, as follows [28]:

- (i) Q4 2018: South Korea, U.S.

- (ii) Q1 2019: Bahrain, Czech Republic, Estonia, Finland, Saudi Arabia, Switzerland
- (iii) Q2 2019: Australia, Qatar
- (iv) Q3 2019: Austria, China, Hong Kong, Kuwait, Spain, UAE
- (v) Q4 2019: Portugal, UK

As of Q2 2019, there were 303 rollouts of 5G mobile networks across 294 locations worldwide, operated by 20 mobile

TABLE 6: Summary of near-term 5G service-deployment activities (2019 view).

Country or Region	Near-term 5G Activities
South Korea	Korea Telecom rolled out a trial 5G network in support of the 2018 Winter Olympics in Seoul, South Korea covering events in several cities. It has also launched a VR/AR games platform supported from the cloud over 5G. SK Telecom has acquired spectrum in the 3.5 GHz and 28 GHz frequencies in preparation of deploying 5G.
China	China plans early implementations of 5G. The GSMA estimates that by 2025 China will represent 40 percent of global 5G connections. According to the GSMA, with 460 million projected users China is expected to become the world's largest 5G market by 2025, higher than Europe's 205 million and the United States' 187 million combined. China's three major mobile operators - China Mobile, China Unicom and China Telecom - are rolling out trial operations of 5G systems in several cities, and all three aim to fully commercialize the technology by 2020 [66].
India	5G trials are contemplated by late 2019 and early deployments may happen late in 2020.
Japan	NTT DoCoMo demonstrated an advanced security service based on 5G network technology for use in the 2020 Olympics.
United States	Migration from the 4G networks should be relatively simple. The FCC has been making several new bands available as noted elsewhere. Verizon has been aggressive in its advertisement campaigns about its introduction of 5G-related services.
Europe	T-Mobile is preparing for the rollout of 5G in 2020, starting in the Netherlands. Some fear that Europe risks falling behind other regions because of restrictive regulation and weak investments; less than half of the countries in Europe have actually allocated spectrum for 5G [67].

carriers [93]. In the U.S., 21 deployments were documented, of which five were in Texas, four in California, two each in North Carolina and Florida, and one each in Oklahoma, Minnesota, Illinois, Indiana, Kentucky, Tennessee, Georgia, and Louisiana (some of these, such as the AT&T 5G network in Louisville, KY, had "Limited Availability" at that time.)

5.2. MmWave Spectrum. Regarding frequency bands 3GPP is initially focusing on 24 GHz to 43 GHz mmWave spectrum (Release 15.) Other ongoing 5G work relates to NSA and SA configurations Massive MIMO, beamforming, and LTE interoperability. 3GPP Release 16 (2019) aims at full compliance to IMT-2020 (e.g., supporting 1 GHz channels) and other spectrum capabilities (e.g., spectrum sharing, additional bands, and URLCC).

In the U.S., among other possible candidates, the FCC is making available new frequency bands for 5G use under its rubric of "Spectrum Frontiers proceeding", of which three have been instituted in the recent past. With the "July 2016 Order", the FCC designated the 27.5-28.35 GHz (known as the "28 GHz band"), 37-38.6 GHz (known as the "37 GHz band"), and 38.6-40 GHz (known as the "39 GHz band") bands for flexible mobile and fixed commercial use, and designated the 64-71 GHz band for unlicensed use (to supplement 57-64 GHz which had been made available for unlicensed use at an earlier time.) While the FCC has yet to auction any of the new Upper Microwave Flexible Use Service or (UMFUS) spectrum, in 2017, with a *Second Report and Order*, a *Memorandum Opinion and Order* it designated an additional 1700 megahertz of mmWave spectrum for licensed flexible commercial wireless fixed and mobile use. The 1700 MHz spectrum covered the 24.25-24.45, 24.75-25.25, and 47.2-48.2 GHz bands (the first two known collectively as the "24 GHz" bands, and the third, known as the "47 GHz" band). Therefore, the spectrum at 24.25-24.45 GHz is now allocated for non-Federal fixed and mobile services on a

co-primary basis, and the spectrum at 24.75-25.25 GHz for non-Federal fixed, mobile, and fixed-satellite (FSS) services on a co-primary basis [94].

In terms of rollouts, in the U.S., the spectra at 27.5 – 28.35 GHz and 37 – 40 GHz may see preliminary commercial deployments in 2019; in Korea, the spectrum at 26.5 – 29.5 GHz is similarly expected to see commercial deployments in 2019; and the EU expects commercial deployments for the 24.25 – 27.5 GHz spectrum starting around 2020.

In addition to the radio access for the end-user device, there is also interest in Backhaul and now also in Fronthaul. Backhaul mechanisms are mechanisms to connect the wireless network to the wired network by backhauling traffic from dispersed cell sites to Mobile Switching Offices (MSOs). These links typically are either traditional transmission systems (such as SONET or point-to-point microwave at various operating bands), or they are Ethernet-over-Fiber links (e.g., 1 GbE or 10 GbE). A Uma site has Baseband Unit (BBU) that processes user and control data, which is in turn connected to a Radio Unit (RU) to generate radio signals transmitted over the air via the tower-mounted antennas.

Fronthaul is related to a new type of Radio Access Network (RAN) architecture that is comprised of centralized baseband controllers and standalone radio heads installed at remote UMa or UMi sites possibly many miles away. In the fronthaul model the BBU and RU equipment is located further away from each other than is the case in the backhaul model. The RU equipment (now referred to as a Remote Radio Head [RRH]) is still located at the cell site, but the BBU is relocated to centralized location where it supports multiple RRHs. See Figure 18. The optical links that interconnect the newly centralized BBU and the multiple RRHs is referred to as fronthaul. The use of fronthaul-based C-RAN (Cloud-RAN) architectures typically improves the cell edge performance. Backhaul and fronthaul are key use cases for mmWave spectrum and will play a role in 5G and 5G IoT.

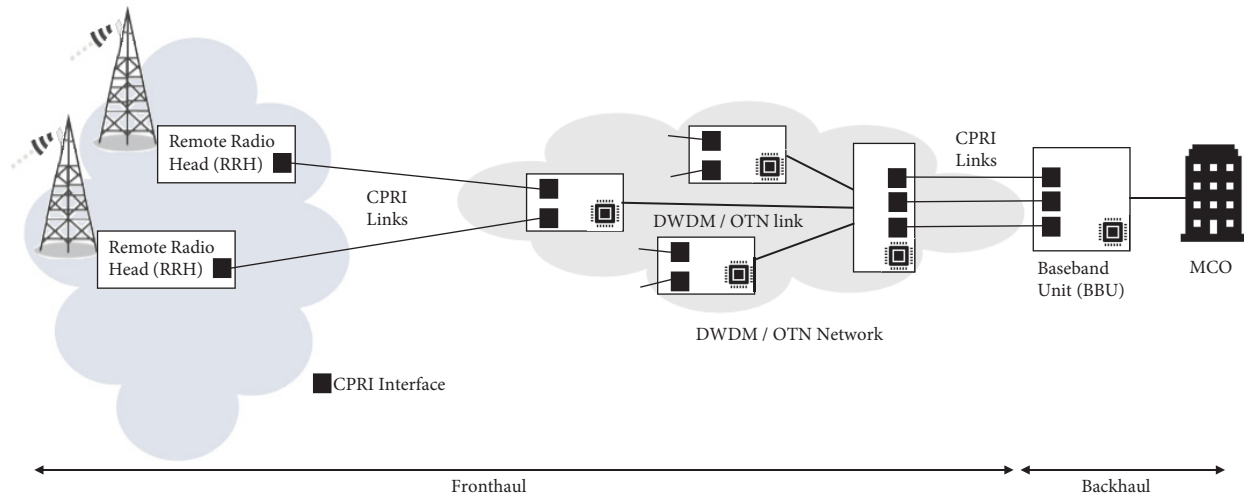


FIGURE 18: Fronthaul and Backhaul.

A consortium of wireless equipment vendors standardized the Common Public Radio Interface (CPRI) protocol that runs over these fronthaul links a few years ago; more recently, a newer eCPRI 1.0 interface has been defined; additionally, work is underway to defining a more detailed interface. The tight performance requirements of CPRI/eCPRI -- capacity, distance, and latency -- drive towards fiber connectivity such as DWDM (or more specifically OTN [Optical Transport Network]) systems between centralized BBUs and the RRHs. Ethernet-based solutions have existed for a number of years using mmWave spectrum. Work is underway in 3GPP to define backhauling solutions using the same spectrum as access. Work is also underway to define new fronthaul interfaces also utilizing mmWave spectrum.

5.3. 5G IoT Services. Global IoT revenue are expected to increase at an annual rate of 23% to 2025 to reach \$1.1 trillion (up from 267 B in 2018). As discussed in the next section, near term “5G IoT” really equate to NB-IoT and LTE-M capabilities. At the end of 2018, there were 83 commercial deployments of LTE-M and NB-IoT worldwide. However, pure connectivity will become increasingly commoditized, making it difficult for operators to compete on the data transmission alone, declining from 9% of total IoT revenue in 2018 to 5% in 2025. Service providers must develop new strategies and business models beyond connectivity services. Applications, platforms and services (e.g., cloud data analytics and IoT security) are the major growth areas of IoT; this segment will be approximately 70% of the market in 2025. Professional services (e.g., consulting, systems integration, also including managed services) will increase in share and will be approximately 25% of the market in 2025 [28].

6. Current Alternatives and Convergence to 5G

5G IoT will need to compete with other technologies, both of the cellular type (e.g., NB-IoT and LTE-M) as well as the non-cellular type (although NB-IoT and LTE-M are now considered “part of the 5G world”). The economics and

availability of these “legacy” networks in various parts of the world may be such that a level of inertia, frustrating a full migration to truly-novel 5G IoT services, will take hold. Clearly, in principle, 5G is better positioned for city/region-wide applications as contrasted with building or campus applications.

From an end-user perspective, design and implementation questions center around the following issues, which 5G IoT technology must be able to address successfully:

- (i) Availability of equipment;
- (ii) Availability of service (geographic coverage in the area of interest);
- (iii) Support of required technical details (latency, bandwidth, packet loss, and so on);
- (iv) Support of mobility (where needed, e.g., wearables, crowdsensing, Vehicle to Vehicle and Vehicle to Infrastructure applications, to name a few);
- (v) Adequate reliability (where needed, e.g., physical security, process control, Vehicle to Vehicle and Vehicle to Infrastructure applications, to name a few);
- (vi) Scalability support (functional and geographic/numerical expansion of the application)
- (vii) Initial and recurring cost of the equipment; and
- (viii) Initial and recurring cost of the service.

Recent acceptability and economics of NB-IoT and LTE-M can serve as a proxy for the near-term commercial success of 5G IoT in particular and truly-novel 5G IoT services in general. Some developers have looked at cellular services for city-wide or region-wide IoT coverage; in some instances, for example, for national truck transportation a combination of Low Earth Orbit (LEO) satellite service and cellular services have and are being used. A current drawback is the cost of the requisite (miniaturized) modems and the cost of the cellular service. New services such as NB-IoT and LTE Cat-M1 (an LTE-based 3GPP-sponsored alternative to NB-IoT,

also known as LTE-M) are short term attempts to address the cost and resource issues. In particular, NB-IoT is seen as providing a pathway to 5G IoT. 5G and truly-novel 5G IoT are the target solutions.

6.1. NB-IoT. As noted earlier, NB-IoT is a licensed low power LPWAN technology designed to coexist with existing LTE specifications and providing cellular-level QoS connectivity for IoT devices. NB-IoT was standardized by 3GPP in LTE Release 13, *but it does not operate in the LTE context per se* [95–97]. NB-IoT has attracted support from Qualcomm, Ericsson and Huawei, among many other vendors and service providers. NB-IoT (also known as LTE Cat-NB1) is based on a Direct Sequence Spread Spectrum (DSSS) modulation in a 200 kHz channel. There are several underutilized 200-kHz GSM spectrum channels, as well as other possible bands such as guard bands. NB-IoT is intended as an alternative to LoRa and Sigfox. This technology can optimize sunken financial investments by service providers and can shorten the service deployment rollout timetable for IoT services, since NB-IoT uses existing cellular infrastructure. NB-IoT service goals include: (i) low complexity end-nodes, (ii) device cost less than \$5, (iii) a device battery life expected to last for 10 years if it transmits 200 bytes of data per day, and (iv) uplink latency less than 10s (thus not a true real-time service). NB-IoT operates on 900-1800 MHz frequency bands with coverage of up to about 20 miles; it supports data rates of up to 250 Kbps for uplink and 230 Kbps for downlink communications [98–101]. NB-IoT can be implemented in a number of ways: (i) in standalone non-cellular licensed bands; (ii) in unused 200 kHz bands in the context of GSM or CDMA; and (iii) in LTE environments where base stations can allocate a resource block to NB-IoT transmissions. Since NB-IoT offers low cost for the device and for the service, it is a good choice for large-scale distributed deployment in Smart Cities and smart grid applications.

As illustrative commercial examples, in 2018 T-Mobile announced a North American NB-IoT plan that costs just \$6 a year – one tenth of Verizon’s Cat-M plans– for up to 12 MB per connected device, and several NB-IoT modules based on Qualcomm® MDM9206 LTE IoT modem that are certified for use on T-Mobile’s network. T-Mobile, in conjunction with Qualcomm and Ericsson conducted the first trial NB-IoT in the U.S. in 2017 across multiple sites; T-Mobile and the City of Las Vegas also announced a partnership to deploy IoT technology throughout the city. For applications that require more bandwidth and voice, T-Mobile offers Cat-1 IoT Access Packs [102, 103]. NB-IoT consumes minimal power: while most IoT end-nodes save power when they are quiescent, when the node and the modem are running and handling all the signal processing, the systems with simpler waveform (such NB-IoT) consume less overall power. Additionally, chipsets that support a single protocol (such as NB-IoT) are cheaper compared to a chipset that supports multiple protocols. Furthermore, *prima facie*, NB-IoT may provide deeper building penetration than LTE-M.

6.2. LTE-M. LTE-M is a power-efficient system, where two innovations support battery efficiency: LTE eDRX (Extended

Discontinuous Reception) and LTE PSM (Power Saving Mode). LTE-M allows the upload of 10 bytes of data a day (LTE-M messages are fairly short compared to NB-IoT messages), but also allows access to Mbps rates. Therefore, LTE-M can support several use cases. In the U.S., major carriers such as Verizon and AT&T offer LTE-M services (as noted, Verizon has announced support for NB-IoT -- T-Mobile and Sprint appears to lean in the NB-IoT direction) [104]. Worldwide geographies with GSM deployments will likely offer NB-IoT in the short term. Figure 19 depicts some of the IoT compatibility mechanisms to be incorporated into 5G in terms of band and bandwidth; however, the transmission frequencies will be wildly different.

In summary, LTE-M supports low nodal complexity, high nodal density, low nodal power consumption, low latency and extended geographic coverage, while allowing service operators the reuse of the LTE installed base. NB-IoT aims at improved indoor coverage, high nodal density for low throughput devices, low delay sensitivity, low node cost, low nodal power consumption and simplified network architecture. NB-IoT and LTE-M are currently providing mobile IoT solutions for smart cities, smart logistics, and smart metering, but only in small deployments to date (as of early 2018, there were 43 commercial NB-IoT and LTE-M networks worldwide [105]). As noted, the commercial success of NB-IoT and LTE-M can serve as a proxy for the eventual success of 5G IoT in a smart city context (compared with non-cellular LPWAN solutions).

NB-IoT, LTE-M and LTE are 4G standards, but advocates claim that they remain integral parts of early releases of 5G. Proponents make the case that “*enterprises deploying either NB-IoT or LTE-M are futureproofing their IoT projects because when 5G rollouts become commonplace these two Mobile IoT standards will continue into foreseeable 5G releases (from 3GPP Release 15 on)*” [102]. In the context of 3GPP Rel 15, it appears, in fact, that NB-IoT and LTE-M will be included as 5G mobile standards. In 2018 the GSMA asserted that “*NB-IoT and LTE-M, as deployed today, are part of the 5G family; with the dawn of the 5G era [...] both NB-IoT and LTE-M technologies are an integral part of 5G, and that 5G from the LPWA perspective, is already here today*” [105]. Including these technologies as initial 5G IoT standards will motivate service providers and vendors to support these implementations for IoT deployments as an evolutionary strategy to 5G. 3GPP Release 16 (targeted for the end of 2019) is considered to be the “second 5G standard” and thereafter transmitted to the ITU for consideration as a global standard. Among other functionality and capabilities, Release 16 is expected to add standards for connected cars and smart factories (notably, automobile companies have formed the 5G Automotive Association to assist 3GPP to set autonomous vehicle standards, such as 5G cellular vehicle-to-everything [C-V2X]).

7. Conclusion

This paper discussed a number of issues related to 5G-based IoT applications, particularly in Smart Cities environments,

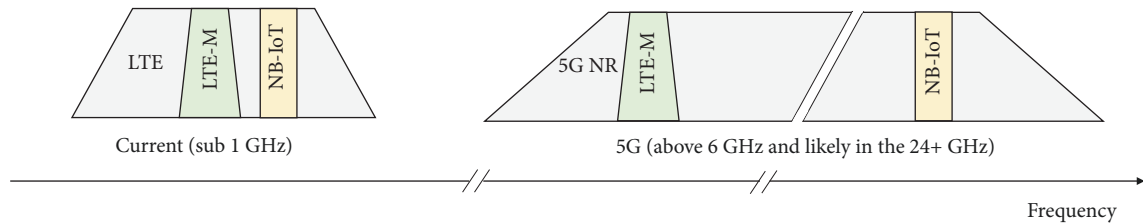


FIGURE 19: Support of LTE-M and NB-IoT under 5G.

including the need for small cells, the transmission issues at the millimeter wave frequencies, building penetration issues, the need for DAS, and the near term introduction of pre-5G IoT technologies such as NB-IoT and LTE-M, these being possible proxies for 5G IoT deployment.

A firm definition of 5G IoT has still to emerge, although a large number of use cases have been described by various industry entities. Both 3GPP NB-IoT and LTE-M technologies are seen at this juncture as integral to 5G services: these 4G technologies are expected to continue under full support in 5G networks for the immediate future. However, IoT/Smart City applications that require high bandwidth will need implementations of eMBB and mmWave frequencies.

Some controversy existed at press time about the development of 5G equipment, in the context of origin-of-manufacturing and the possible intrinsic risk related to cybersecurity [106]. If these issues are not satisfactorily resolved, some delay in the broad early deployment of 5G may result. However, the expectation is that these issues will work themselves out over time.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] National League of Cities and Center for City Solutions and Applied Research, "City of the future – technology and mobility," White Paper 1301, Pennsylvania Avenue, Suite 550, Washington, DC, USA, 2015.
- [2] A. Ramaswami, A. G. Russell, P. J. Culligan, K. Rahul Sharma, and E. Kumar, "Meta-principles for developing smart, sustainable, and healthy cities," *Science*, vol. 352, no. 6288, pp. 940–943, 2016.
- [3] D. R. Martínez, T. J. Gracia, E. M. Muñoz, and A. C. García, "Smart cities' challenge: how to improve coordination in the supply chain," in *Sustainable Smart Cities*, Innovation, Technology, and Knowledge Management, pp. 129–142, Springer International Publishing, 2016.
- [4] N. Mali, "A review on smart city through internet of things (IoT)," *International Journal of Advanced Research in Science Management and Technology*, vol. 2, no. 6, 2016.
- [5] A. Caragliu, C. D. Bo, and P. Nijkamp, "Smart cities in europe," *Journal of Urban Technology*, vol. 18, no. 2, pp. 65–82, 2011 (Chinese).
- [6] D. Minoli and B. Occhiogrosso, "Internet of things applications for smart cities," in *Internet of Things A to Z: Technologies and Applications*, Q. Hassan, Ed., Chapter 12, IEEE Press/Wiley, 2018.
- [7] A. Zanella, N. Bui, A. P. Castellani, L. Vangelista, and M. Zorzi, "Internet of things for smart cities," *IEEE Internet of Things Journal*, vol. 1, no. 1, pp. 22–32, 2014.
- [8] D. Minoli and B. Occhiogrosso, "Mobile IPv6 protocols and high efficiency video coding for smart city IoT applications," in *Proceedings of the 13th International Conference and Expo on Emerging Technologies for a Smarter World (CEWIT)*, pp. 1–6, Stony Brook, New York, NY, USA, 2017.
- [9] B. J. Wewalaarachchi, H. Shivanan, and H. Gunasingham, "Integration platform to enable operational intelligence and user journeys for smart cities and the internet of things," in *Proceedings of the Patent US20160239767 A1*, 2016.
- [10] S. Srivastava and N. Pal, "Smart cities: the support for internet of things (IoT)," *International Journal of Computer Applications in Engineering Sciences*, pp. 5–7, 2016.
- [11] O. Bates and A. Friday, "Beyond data in the smart city: repurposing existing campus IoT," *IEEE Pervasive Computing*, vol. 16, no. 2, pp. 54–60, 2017.
- [12] D. Kyriazis, T. Varvarigou, D. White et al., "Sustainable smart city IoT applications: heat and electricity management & eco-conscious cruise control for public transportation," in *Proceedings of the IEEE 14th International Symposium on "A World of Wireless, Mobile and Multimedia Networks" (WoWMoM)*, IEEE, Madrid, Spain, 2013.
- [13] D. Minoli and B. Occhiogrosso, "IoT applications to smart campuses and a case study," *European Union Digital Library*, vol. 5, article e4, pp. 2518–3893, 2017.
- [14] A. Al-Fuqaha, M. Guizani, M. Mohammadi et al., "Internet of things: a survey on enabling technologies, protocols, and applications," *IEEE Communication Surveys & Tutorials*, vol. 17, no. 4, pp. 2347–2376, 2015.
- [15] R. Gomes, H. Pombeiro, C. Silva et al., "Towards a smart campus: building-user learning interaction for energy efficiency, the lisbon case study," in *Handbook of Theory and Practice of Sustainable Development in Higher Education*, World Sustainability Series, pp. 381–398, Springer, 2016.
- [16] Z. Yu, Y. Liang, B. Xu et al., "Towards a smart campus with mobile social networking," in *Proceedings of the 4th IEEE Int'l Conference on Cyber, Physical and Social Computing (CPSCom)*, pp. 162–169, IEEE, Dalian, China, 2011.
- [17] A. Roy, J. Siddiquee, A. Datta et al., "Smart traffic & parking management using IoT," in *Proceedings of the IEEE 7th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, IEEE, Vancouver, BC, Canada, 2016.
- [18] R. Grodi, D. B. Rawat, and F. Rios-Gutierrez, "Smart parking: Parking occupancy monitoring and visualization system for

- smart cities,” in *Proceedings of the SoutheastCon*, IEEE, Norfolk, VA, USA, 2016.
- [19] D. Minoli, K. Sohraby, and B. Occhiogrosso, “IoT considerations, requirements, and architectures for smart buildings – energy optimization and next generation building management systems,” *IEEE Internet of Things Journal*, vol. 4, no. 1, pp. 269–283, 2017.
- [20] L. Kang, S. Poslad, W. Wang et al., “A public transport bus as a flexible mobile smart environment sensing platform for IoT,” in *Proceedings of the 12th International Conference on Intelligent Environments (IE)*, IEEE, London, UK, 2016.
- [21] M. Alam, J. Ferreira, and J. Fonseca, “Introduction to intelligent transportation systems,” in *Journal of Intelligent Transportation Systems*, vol. 52 of *Studies in Systems, Decision and Control*, pp. 1–17, Springer International Publishing, 2016.
- [22] J. Barbaresso, G. Cordahi, and et al., “The national academies of science, engineering and medicine, USDOT’s intelligent transportation systems (ITS): its strategic plan,” The National Academies of Science, Engineering and Medicine, USDOT’s Intelligent Transportation Systems (ITS): ITS Strategic Plan 2015–2019 FHWA–JPO–14–145, 2014.
- [23] S. H. Sutar, R. Koul, and R. Suryavanshi, “Integration of Smart Phone and IOT for development of smart public transportation system,” in *Proceedings of the International Conference on Internet of Things and Applications (IOTA)*, pp. 73–78, Pune, India, 2016.
- [24] M. Picone, S. Busanelli, M. Amoretti et al., *Advanced Technologies for Intelligent Transportation Systems*, Springer, 2015.
- [25] Q. Wang, Y. Zhao, W. Wang et al., “Multimedia IoT systems and applications,” in *Proceedings of the Global Internet of Things Summit (GloTS)*, IEEE, Geneva, Switzerland, 2017.
- [26] D. Minoli and B. Occhiogrosso, “Ultrawideband (UWB) technology for smart cities IoT applications,” in *Proceedings of the IEEE International Smart Cities Conference (ISC2)*, pp. 1–8, IEEE, Kansas City, Mo, USA, 2018.
- [27] R. K. Ganti, F. Ye, and H. Lei, “Mobile crowdsensing: current state and future challenges,” *IEEE Communications Magazine*, vol. 49, no. 11, pp. 32–39, 2011.
- [28] Global System for Mobile Communications Association (GSMA), *The Mobile Economy 2019*, 2019, <https://www.gsmain-telligence.com/research/?file=b9a6e6202e1d5f787cfebb95d36-39c5&download>.
- [29] Statista, *Internet of things (IoT) connected devices installed base worldwide from 2015 to 2025 (in billions)*, 2015, <https://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide/>.
- [30] “Recommendation ITU-R M.2083-0. IMT vision—framework and overall objectives of the future development of IMT for 2020 and beyond,” <http://www.itu.int/rec/R-REC-M.2083-0-201509-1/>.
- [31] “Next generation mobile networks,” 5G White Paper, 2015, https://www.ngmn.org/fileadmin/ngmn/content/downloads/Technical/2015/NGMN_5G_White_Paper_V1.0.pdf.
- [32] “Feasibility study on new services and markets technology enablers,” 3GPP 22.891, 2019, <http://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2897>.
- [33] H. Yu, H. Lee, and H. Jeon, “What is 5G? emerging 5G mobile services and network requirements,” *Sustainability*, vol. 9, no. 10, article 1848, 2017.
- [34] GSMA, *Road to 5G: introduction and migration*, 2018, https://www.gsma.com/futurenetworks/wp-content/uploads/2018/04/Road-to-5G-Introduction-and-Migration_FINAL.pdf.
- [35] S. W. Hu and C. M. Shy, “Health effects of waste incineration: a review of epidemiologic studies,” *Journal of the Air & Waste Management Association*, vol. 51, no. 7, pp. 1100–1109, 2001.
- [36] A. Santarsiero, G. Trevisan, G. Cappiello et al., “Urban crematoria emissions as they stand with current practice,” *Microchemical Journal*, vol. 79, no. 1-2, pp. 299–306, 2005.
- [37] M. Takaoka, K. Oshita, N. Takeda, and S. Morisawa, “Mercury emission from crematories in Japan,” *Atmospheric Chemistry and Physics*, vol. 10, no. 8, pp. 3665–3671, 2010.
- [38] N. Takeda, M. Takaoka, K. Oshita, and S. Eguchi, “PCDD/DF and co-planar PCB emissions from crematories in Japan,” *Chemosphere*, vol. 98, pp. 91–98, 2014.
- [39] Y. Xue, H. Tian, J. Yan et al., “Present and future emissions of HAPs from crematories in China,” *Atmospheric Environment*, vol. 124, pp. 28–36, 2016.
- [40] D. Muenhor, J. Satayavivad, W. Limpaseni et al., “Mercury contamination and potential impacts from municipal waste incinerator on Samui Island, Thailand,” *Journal of Environmental Science and Health, Part A: Toxic/Hazardous Substances and Environmental Engineering*, vol. 44, no. 4, pp. 376–387, 2009.
- [41] S. Sakai, K. Hayakawa, H. Takatsuki, and I. Kawakami, “Dioxin-like PCBs released from waste incineration and their deposition flux,” *Environmental Science & Technology*, vol. 35, no. 18, pp. 3601–3607, 2001.
- [42] G. D. Hinshaw and A. R. Trenholm, “Hazardous waste incineration emissions in perspective,” *Waste Management*, vol. 21, no. 5, pp. 471–475, 2001.
- [43] D. C. Ashworth, G. W. Fuller, M. B. Toledano et al., “Comparative assessment of particulate air pollution exposure from municipal solid waste incinerator emissions,” *International Journal of Environmental Research and Public Health*, vol. 2013, 13 pages, 2013.
- [44] *Chapter 4 in Waste Incineration & Public Health, National Research Council (US) Committee on Health Effects of Waste Incineration. Washington (DC): National Academies Press (US), 2000, ISBN-10: 0-309-06371-X. Also at https://www.ncbi.nlm.nih.gov/books/NBK233615/*.
- [45] S. Bose-O’Reilly, K. M. McCarty, N. Steckling et al., “Mercury exposure and children’s health,” *Current Problems in Pediatric and Adolescent Health Care*, vol. 40, no. 8, pp. 186–215, 2010.
- [46] G. González-Cardoso, N. Santiago, J. M. Hernández-Contreras, and M. Gutiérrez, “PM2.5 emissions from urban crematoriums,” *Energy Procedia*, vol. 153, pp. 359–363, 2018.
- [47] METIS, mobile and wireless communications enablers for the twenty-twenty (2020) Information society, the 5G future scenarios identified by METIS –the first step toward a 5G mobile and wireless communications system, 2013.
- [48] A. Osseiran, V. Braun, T. Hidekazu et al., “The foundation of the mobile and wireless communications system for 2020 and beyond: challenges, enablers and technology solutions,” in *Proceedings of the IEEE 77th Vehicular Technology Conference (VTC Spring)*, IEEE, Dresden, Germany, 2013.
- [49] ICT-317669 METIS project, “Requirements and general design principles for new air interface,” <https://www.metis2020.com/documents/deliverables/>, 2013.
- [50] ICT-317669 METIS project, “Positioning of multi-node/multi-antenna transmission technologies,” <https://www.metis2020.com/documents/deliverables/>, 2013.

- [51] ICT-317669 METIS project, "Intermediate description of the spectrum needs and usage principles," <https://www.metis2020.com/documents/deliverables/>, 2013.
- [52] ICT-317666 METIS project, "Summary on preliminary trade-off investigations and first set of potential network-level solutions," <https://www.metis2020.com/documents/deliverables/>, 2013.
- [53] ICT-317669 METIS project, "Novel radio link concepts and state of the art analysis," <https://www.metis2020.com/documents/deliverables/>, 2013.
- [54] ICT-317669 METIS project, "Components of a new air interface - building blocks and performance," <https://www.metis2020.com/documents/deliverables/>, 2014.
- [55] "Small cell forum, mmWave 5G eMBB use cases & small cell based hyperdense networks," 2017.
- [56] X. Ge, L. Pan, Q. Li et al., "Multipath cooperative communications networks for augmented and virtual reality transmission," *IEEE Transactions on Multimedia*, vol. 19, no. 10, pp. 2345–2358, 2017.
- [57] G. Tech, Y. Chen, K. Müller et al., "Overview of the multiview and 3D extensions of high efficiency video coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 1, pp. 35–49, 2016.
- [58] J. Horwitz, "FCC expands 3.5GHz band to 5G and opens 6 GHz band to future Wi-Fi," <https://venturebeat.com/2018/10/23/fcc-expands-3-5ghz-band-to-5g-and-opens-6ghz-band-to-future-wi-fi/>, 2018.
- [59] D. Minoli, K. Sohraby, and B. Occhiogrosso, "IoT security (IoTSec) mechanisms for e-health and ambient assisted living applications," in *Proceedings of the 2nd IEEE International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*, IEEE, Philadelphia, Pa, USA, 2017.
- [60] D. Minoli, K. Sohraby, and J. Kouns, "IoT Security (IoTSec) Considerations, Requirements," in *Proceedings of the 14th IEEE Annual Consumer Communications & Networking Conference (CCNC)*, IEEE, Las Vegas, NV, USA, 2017.
- [61] "Security architecture and procedures for 5G System," 3GPP TS 33.501, 2018, <http://www.3gpp.org/DynaReport/33501.htm>.
- [62] "Study on the security aspects of the next generation system," 3GPP TR 33.899, 2017, <http://www.3gpp.org/DynaReport/33899.htm>.
- [63] "5G Enablers for Network and System Security and Resilience," <http://www.5gensure.eu/>, 2017.
- [64] GDPR, *General Data Protection Regulation*, European Union, 2016, <http://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX:32016R0679>.
- [65] ePrivacy, *Directive on privacy and electronic communications*, European Union, 2002, <https://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX:32002L0058>.
- [66] The Strait Times Staff, *China to Be Biggest 5G Market by 2025: Report*, 2019, <https://www.straitstimes.com/world/china-to-be-biggest-5g-market-by-2025-report>.
- [67] J. Stubbs and D. Busvine, *Weak investment climate main 5G risk, not security fears: Ericsson*, <https://www.reuters.com/article/us-telecoms-mobileworld-ericsson/weak-investment-climate-main-5g-risk-not-security-fears-ericsson-idUSKCN1QE0ST>.
- [68] FCC, *Millimeter Wave Propagation: Spectrum Management Implications*, vol. 70, Federal Communications Commission Office of Engineering and Technology New Technology Development Division, Washington, DC, USA, 1997.
- [69] P. Tracy, *What is Mm Wave and How Does It Fit into 5G?*, 2016, <https://www.rcrwireless.com/20160815/fundamentals/mmwave-5g-tag31-tag99>.
- [70] X. Zhang, *Millimeter Wave for 5G: Unifying Communication and Sensing*, 2015, https://www.microsoft.com/en-us/research/wp-content/uploads/2015/03/Xinyu-Zhang_5GmmWave.pdf.
- [71] X. Ge, J. Yang, H. Gharavi, and Y. Sun, "Energy efficiency challenges of 5G small cell networks," *IEEE Communications Magazine*, vol. 55, no. 5, pp. 184–191, 2017.
- [72] X. Ge, Z. Li, and S. Li, "5G software defined vehicular networks," *IEEE Communications Magazine*, vol. 55, no. 7, pp. 87–93, 2017.
- [73] S. Sun, T. S. Rappaport, S. Rangan et al., "Propagation path loss models for 5G urban micro- and macro-cellular scenarios," in *Proceedings of the 83rd IEEE Vehicular Technology Conference (VTC Spring)*, IEEE, Nanjing, China, 2016.
- [74] "Draft declaratory ruling and third report and order," WC Docket No. 17-84, WT Docket No.17-79, FCC-CIRC1809-02, 2018.
- [75] "Small cell forum," Small Cells Market Status Report, 2018, http://www.scf.io/en/documents/050_Small_cells_market_status_report_February_2018.php?utm_source=Email%campaign&utm_medium=eshots&utm_campaign=member%eshot.
- [76] T. S. Rappaport, Y. Xing, G. R. MacCartney et al., "Overview of millimeter wave communications for fifth-generation (5G) wireless networks-with a focus on propagation models," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 12, pp. 6213–6230, 2017.
- [77] T. Rappaport, S. Sun, R. Mayzus et al., "Millimeter wave mobile communications for 5G cellular," *IEEE Access*, vol. 1, pp. 335–349, 2013.
- [78] T. Bai, A. Alkhatieb, and R. W. Heath, "Coverage and capacity of millimeter-wave cellular networks," *IEEE Communications Magazine*, vol. 52, no. 9, pp. 70–77, 2014.
- [79] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: potentials and challenges," *Proceedings of the IEEE*, vol. 102, no. 3, pp. 366–385, 2014.
- [80] W. Roh, J.-Y. Seol, J. Park et al., "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: theoretical feasibility and prototype results," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 106–113, 2014.
- [81] 3GPP 38.901-e20, "3rd generation partnership project," Technical Specification Group Radio Access Network; Study on Channel Model For Frequencies From 0.5 to 100 GHz (Release 14), 2017.
- [82] F. Khan and Z. Pi, "mmWave mobile broadband (MMB): unleashing the 3–300GHz spectrum," in *Proceedings of the 34th IEEE Sarnoff Symposium*, pp. 1–6, Princeton, NJ, USA, 2011.
- [83] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 101–107, 2011.
- [84] F. Khan and Z. Pi, "Millimeter-wave mobile broadband: unleashing 3-300 GHz spectrum," in *Proceedings of the IEEE Wireless Communications and Networking Conference*, IEEE, 2011.
- [85] S. Rajagopal, S. Abu-Surra, Z. Pi, and F. Khan, "Antenna array design for multi-Gbps mmwave mobile broadband communication," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM)*, pp. 1–6, Houston, Tex, USA, 2011.
- [86] A. Ghosh, T. A. Thomas, M. C. Cudak et al., "Millimeter-wave enhanced local area systems: a high-data-rate approach for future wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1152–1163, 2014.

- [87] G. R. MacCartney and T. S. Rappaport, "Study on 3GPP rural macrocell path loss models for millimeter wave wireless communications," in *Proceedings of the ICC 2017 - 2017 IEEE International Conference on Communications*, pp. 1–7, IEEE, Paris, France, 2017.
- [88] Y.-S. Lu, C.-F. Lai, C.-C. Hu, and Y.-M. Huang, "Path loss exponent estimation for indoor wireless sensor positioning," *KSII Transactions on Internet and Information Systems*, vol. 4, no. 3, article 243, 2010.
- [89] S. Srinivasan and M. Haenggi, "Path loss exponent estimation in large wireless networks," *Information Theory and Applications Workshop*, pp. 124–129, 2009.
- [90] M. Viswanathan, *Log Distance Path Loss or Log Normal Shadowing Model*, 2013, <https://www.gaussianwaves.com/2013/09/log-distance-path-loss-or-log-normal-shadowing-model/>.
- [91] G. L. Lederer, "Smart communities and special districts coalition – ex parte submission: accelerating wireless broadband deployment by removing barriers to infrastructure investment WT docket no. 17-79 and no. 17-84," *Best, Best & Krieger, LLP*, 2018, 2000 Pennsylvania Avenue, N.W., Suite 5300, Washington, DC 20006.
- [92] GOS World Staff, *Skyworks Unveils Sky5 Ultra Platform for 5G Architecture*, 2019, <https://www.gpsworld.com/skyworks-unveils-sky5-ultra-platform-for-5g-architecture/>.
- [93] J. Sanders, *Tracking 5G: Ookla's Coverage Map Tracks Worldwide Network Rollout*, 2019, <https://www.techrepublic.com>.
- [94] C. Yorkgitis, *FCC Adopts a Second Wave of Millimeter Wave Regulations to Support Next Generation Terrestrial Systems and Services*, Common Law Monitor, 2017, <https://www.commlaw-monitor.com/2017/12/articles/wireless-2/fcc-adopts-a-second-wave-of-millimeter-wave-regulations-to-support-next-generation-terrestrial-systems-and-services/>.
- [95] Y. Saleem, N. Crespi, M. H. Rehmani, and R. Copeland, "Internet of things-aided smart grid: technologies, architectures, applications, prototypes, and future research directions," *IEEE Access*, vol. 7, pp. 62962–63003, 2019.
- [96] Y. Li, X. Cheng, Y. Cao, D. Wang, and L. Yang, "Smart choice for the smart grid: narrowband internet of things (NB-IoT)," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 1505–1515, 2018.
- [97] P. Reininger, "3GPP standards for the internet of-things," <https://www.slideshare.net/eikoseidel/3gpp-standards-for-the-internetofthings>, 11 3gpp Standards for IoT.pdf, 2016.
- [98] "cellular system support for ultra-low complexity and low throughput internet of things (CIoT)," <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2719>.
- [99] R. Ratasuk, B. Vejlgaard, N. Mangalvedhe, and A. Ghosh, "NB-IoT system for M2M communication," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–5, 2016.
- [100] Link Labs Staff, *An Overview of Narrowband IoT (NB-IoT)*, 2018, <https://www.link-labs.com/blog/overview-of-narrowband-iot>.
- [101] Y. E. Wang, X. Lin, A. Adhikary et al., "A primer on 3GPP narrowband internet of things," *IEEE Communications Magazine*, vol. 55, no. 3, pp. 117–123, 2017.
- [102] "T-Mobile NB-IoT plan," <https://iot.t-mobile.com/pricing/>.
- [103] "Verizon wireless Cat-M Plan," <https://www.verizonwireless.com/biz/plans/m2m-business-plans/>.
- [104] GSMA, *Mobile IoT in the 5G Future- NB-IoT and LTE-M in the context of 5G*, 2018, https://www.gsma.com/iot/wp-content/uploads/2018/05/GSMAIoT_MobileIoT_5G_Future_May2018.pdf.
- [105] M. Contento, *5G and IoT – Emerging Tech with Endless Use Cases*, 2019, <https://www.telit.com/blog/state-of-5g-and-iot-current-future-applications/>.
- [106] *The Guardian View on Google Versus Huawei: No Winners*, The Guardian, 2019, <https://www.theguardian.com/commentisfree/2019/may/20/the-guardian-view-on-google-versus-huawei-no-winners>.

Research Article

Radar-Assisted UAV Detection and Identification Based on 5G in the Internet of Things

Jingcheng Zhao , **Xinru Fu** , **Zongkai Yang**, and **Fengtong Xu**

School of Electronics and Information Engineering, Beihang University, China

Correspondence should be addressed to Xinru Fu; zy1702412@buaa.edu.cn

Received 4 April 2019; Revised 2 June 2019; Accepted 23 June 2019; Published 14 July 2019

Guest Editor: Bo Rong

Copyright © 2019 Jingcheng Zhao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Unmanned aerial vehicles (UAVs) have broad application potential for the Internet of Things (IoT) due to their small size, low cost, and flexible control. At present, the main positioning method for UAVs is the use of GPS. However, GPS positioning may be affected by stronger electromagnetic signals from spoofing attacks. In this study, a radar-assisted positioning method based on 5G millimeter waves is proposed. In 5G end-to-end network slices, the rotors of UAVs can be detected and identified by deploying 5G millimeter wave radar. High-resolution range profile (HRRP) is used to obtain the UAV location in the detection zone. Micro-Doppler characteristics are used to identify the UAVs and the cepstrum method is used to extract the number and speed information of the UAV rotor. The sinusoidal frequency modulation (SFM) parameter optimization method is used to separate multiple UAVs. The proposed method provides information on the number of UAVs, the position of the UAV, the number of rotors, and the rotation speed of each rotor. The simulation results show that the proposed radar detection method is well suited for UAV detection and identification and provides a valid GPS-independent method for UAV tracking.

1. Introduction

With the maturity of unmanned aerial vehicle (UAV) technology and the improvement of relevant laws and regulations, UAVs are increasingly being used for the development of the Internet of Things (IoT). For example, UAVs have been broadly used in the military IoT, smart agriculture, and smart cities to obtain and transmit geospatial information, sensor data information, and controlling information. In smart agriculture, UAV systems are used to gather near real-time remote sensing data for precision farming. In such applications, UAV systems need to be positioned precisely to obtain data from sensors. In smart cities, almost all aspects of the city are combined with the IoT, a task that requires large amounts of data transmission. This data volume requires a large number of base stations for IoT data transmission. Researchers have equipped UAVs with various sensors, such as high-definition cameras, as well as temperature, humidity, and air pollution sensors. In addition, today's UAVs are equipped with high quality wireless communication functions, including 5G, Wifi, Bluetooth, radio-frequency identification (RFID), and other communication means [1]. This

type of equipment means that UAVs have a very important role in the IoT and can act as a mobile air base station, providing reliable downlink and uplink communications to terrestrial users, thereby increasing the capacity of wireless networks [2]. In addition, UAVs can be used as airborne wireless relays, as well as signal boosters and mobile nodes for the IoT and sensor networks [3]. If they are used as reliable air base stations, UAVs need to be positioned precisely. Another application of UAVs is for product delivery, in which case the UAVs location has to be monitored constantly. In all the aforementioned applications, the positioning of UAVs is very important. However, due to the mobility of UAVs, in certain cases, accurate positioning remains an unsolved problem. In particular, high-precision positioning is very important for remote sensing and automatic waypoint navigation of UAVs [4]. If the UAV cannot be accurately positioned, the UAV may collide with objects or other aircraft or may be deceived, captured, or interfered with.

At present, the positioning technology of UAVs is mainly achieved by using the GPS system but this method is sensitive to certain environments. Theoretically, the GPS system only requires three satellites to obtain positioning data but in

practice, the UAV needs to communicate with at least 10 satellites to obtain a stable GPS system. When flying indoors, through tunnels, or dense construction areas, the GPS signal is not reliable and UAV positioning based on the GPS system will cause difficulties [5]. In addition, if UAVs are used for transporting goods, if criminals want to capture UAVs, it is difficult to prevent these attacks using GPS systems. Since the civilian GPS signal is not encrypted, the signal structure and data of UAV positioning system used in the civil IoT is predictable. This characteristic makes the civilian Global Navigation Satellite System (GNSS) signals a simple target for spoofing attacks. Moreover, after the UAV is captured, the criminal can take control of the captured UAV and fly away from its original flight path, while the original control source still receives data that appear correct and no alert is sent. This results in great risk in the use of UAVs. Many different technologies that assist in the positioning of UAVs have been developed, such as visual aids. However, visual aids are inevitably prone to drifting during long-range surveys without a priori feature maps and are only suitable for good weather and lighting conditions [6].

In this study, the precise positioning of UAVs is achieved through radar-assisted detection and identification of UAV technology. The method prevents many problems associated with traditional GPS positioning technology, such as GPS signal failure indoors, satellite occlusion areas, and interference with GPS signals. The radar-assisted detection and identification method has the advantage of working under all weather conditions and is capable of extracting additional features of the UAV due to the micro-Doppler features [7]. According to the current technology development plan, 5G will adopt the previous generations of wireless solutions (3G and 4G) and introduce new promising mmWave technology so that a 5G base station not only can serve as a communication site for 5G networks but can also provide many powerful radar sensors in a functional device network, allowing for the detection of drones in urban environments [8–10]. The background of the UAV flight area is usually complex with other aircraft and birds; therefore, advanced clutter suppression and target detection algorithms are required. Following target detection, it is determined whether the detected target is a UAV and its category is identified. For the purposes of classification and identification, several features unique to UAVs are extracted [11]. In Doppler radar, the Doppler effect is used to determine the velocity of distant objects and accurate measurements can be obtained of the radial component of a targets velocity relative to the radar. Stepped frequency waveforms (SFW) are used to produce a high-resolution range profile (HRRP) of the target. With a wide-band Doppler radar, the HRRP and Doppler information of the target can be obtained to detect and identify UAVs.

The rest of this paper is organized as follows. The second part introduces the system model of the radar-assisted UAV positioning method in the IoT. In the third part, the scattering model of the UAV and rotors is established, the HRRP of a rotor-type UAV is introduced, and the micro-Doppler effect of the UAV is analyzed. The cepstrum method is used to extract the speed of the UAV and the number of rotors and the method for identifying multiple UAVs is

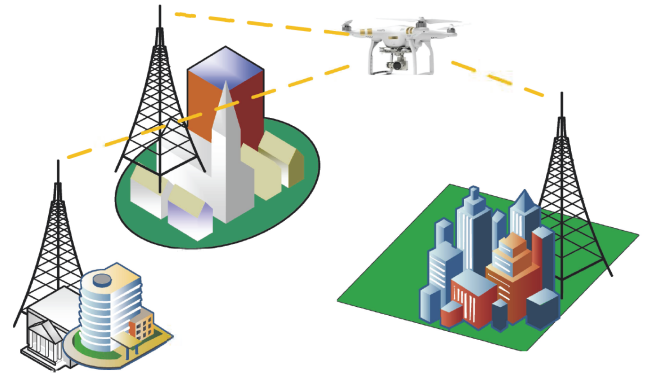


FIGURE 1: UAV detection based on 5G network.

described. In the fourth part, the simulation results and one-dimensional imaging of a multitarget model of two UAVs are provided. The identification of the UAV based on the micro-Doppler information is provided and the method for extracting the number and speed of the rotors are described. The conclusions are provided in the final section.

2. System Model

5G mmWave technology and distributed networks will facilitate radar detection and identification of small UAVs. The narrow steerable beam enables the base station to function as a millimeter wave radar system so that the base station is able to detect and identify drones in urban environments [12, 13]. In addition, through end-to-end fragmentation technology, the 5G network provides a logically isolated virtual private segmentation network for different services to implement a manual flight-specific traffic channel network [14].

The model for detecting and identifying the UAV in the 5G network is shown in Figure 1. The detection network channel of the UAV is established in the 5G network and the millimeter wave base station is established. The electromagnetic signal is transmitted by the millimeter wave base station and beamforming technology is used to reduce the interference of the clutter signal [15, 16]. The reflected echo of the target is obtained by multipath reflection, and the location of UAV is transmitted through the data processing center [17, 18]. Accurate identification and positioning are achieved in a timely manner.

2.1. Radar-Assisted Detection and Identification Method. The flowchart of the process of radar-assisted detection and identification of UAVs is shown in Figure 2. The 5G millimeter wave radar transmits a wideband signal to the target area and receives the echo reflected from the target. The HRRP provides the number of targets in the detection area. For a single target, the Doppler method is used to determine whether the target is a UAV and the cepstrum is used to extract the rotor parameters of the UAV. For multiple targets, after the Doppler and cepstrum analyses, the sinusoidal frequency modulation (SFM) parameter optimization method is used to obtain the

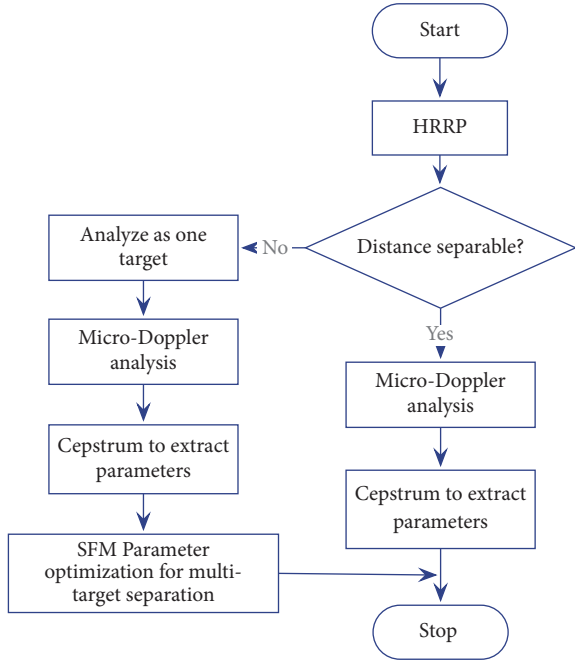


FIGURE 2: Flowchart of radar-based identification.

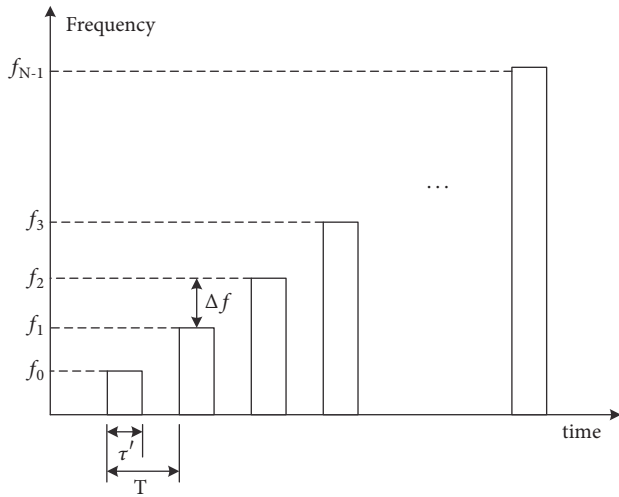


FIGURE 3: SFW burst.

position of each rotor so that multiple targets can be separated and identified.

2.2. Stepped Frequency Waveforms. High bandwidth, low latency, and good beamforming of 5G networks provide a reliable basis for resolution because future 5G mmWave systems cannot only support very high carrier frequencies but also wider bandwidths (effective operation up to 2 GHz). Wider bandwidth facilitates the detection of UAVs flying in close formation. The SFW is a type of 5G mmWave signal and produces an HRRP of the target. A typical SFW burst is shown in Figure 3; it consists of N narrow-band pulses. The frequency from pulse to pulse is stepped by a fixed frequency and each group of pulses is called a pulse train [19].

Within a pulse width τ' , the transmitted signal of the n th pulse is defined in

$$s(n, t) = \rho_n \cos(2\pi f_n t + \theta_n), \quad nT \leq t \leq nT + \tau' \quad (1)$$

where θ_n is the initial phase of the pulse and ρ_n is the amplitude constant.

The echo signal of the target is defined in

$$s_r(n, t) = \rho'_n \cos(2\pi f_n (t - \tau(t)) + \theta_n) \quad (2)$$

where ρ'_n is the amplitude of the echo signal and $\tau(t)$ is the echo delay of the signal; $\tau(t)$ is expressed as

$$\tau(t) = \frac{2R(t)}{c} \quad (3)$$

where c is the speed of light and $R(t)$ is the distance between the target and the radar at time t .

The received signal is split in the quadrature phase detector and enters the mixer of the two synchronous detectors. In the first mixer, the reference signal and the received signal are mixed and the 90° phase-shifted echo signal is mixed with the reference signal in the second mixer. After combining the outputs of the two mixers and low pass filtering, the complex signal is

$$s_r(n, t) = \rho_n \exp \left\{ -j \left[2\pi \left(f_0 + n\Delta f \right) \frac{2R(t)}{c} \right] \right\} \quad (4)$$

If we assume that the targets' moving distance is less than one range resolution cell for each group of pulse signals, then $s_r(n, t)$ is expressed as

$$s_r(n) = \rho_n \exp \left\{ -j \left[2\pi \left(f_0 + n\Delta f \right) \frac{2R}{c} \right] \right\} \quad (5)$$

The inverse discrete Fourier transform (IDFT) of $s_r(n)$ is

$$s(k) = \frac{1}{N} \sum_{n=0}^{N-1} s_r(n) \exp \left(j \frac{2\pi kn}{N} \right) \quad (6)$$

By normalizing [19], (6) yields

$$s(k) = \sum_{n=0}^{N-1} \exp \left[j \frac{2\pi n}{N} \left(k - \frac{2R\Delta f}{c} \right) \right] \quad (7)$$

which can be simplified as

$$s(k) = \frac{\sin \pi \chi}{\sin (\pi \chi / N)} \exp \left(j \frac{N-1}{2} \frac{\pi \chi}{N} \right) \quad (8)$$

where $\chi = -2NR\Delta f/c + k$. The peak of $|s(k)|$ appears at $k = 2NR\Delta f/c \pm mN, m \in \mathbb{Z}$. The distance resolution is $c/(2BW)$, where BW is the signal bandwidth.

2.3. Micro-Doppler Effect in Radar. In traditional radar applications, the antenna illuminates the target with a microwave signal and receives the echo reflected by the target. The echo signal includes the target characteristics of interest. For example, if the transmit signal hits a moving target,

the carrier frequency of the echo signal shifts, which is called the Doppler effect. The Doppler shift reflects the moving speed of the target. Mechanical vibration or rotation of a structure on the target can also cause additional frequency shifts in the returned radar signal; this is known as the micro-Doppler effect. The micro-Doppler effect allows us to determine additional properties of the target [20].

In many cases, some components of the target may have rotations or vibrations in addition to the main target translation, such as a rotor on a UAV or the flapping wing of a bird. The motion dynamics of the rotating rotor or wing will frequency-modulate the backscattered signals and cause additional Doppler shifts near the center of the translational Doppler shift, which is known as the micro-Doppler shift. Therefore, the micro-Doppler shift can be considered a unique signature of targets with vibrations, rotations, or other nonuniform motions. The micro-Doppler shift depends on the signal wavelength, the vibration or rotation speed of the target, and the incident angle of the radar line of sight (LOS) relative to the rotation or vibration plane. The reflected radar signal has the largest micro-Doppler shift when the radar LOS is parallel to the rotation or vibration plane. When the radar LOS is perpendicular to the rotation or vibration plane, the reflected signal has the smallest micro-Doppler shift. The micro-Doppler shift has been widely used to classify targets with rotations or vibrations.

Almost all UAVs have at least one or more rotating rotors. For example, a four-rotor UAV has two rotors that rotate clockwise and two that rotate counterclockwise about a vertical axis. The UAV can take off vertically, hover, and fly forward, backward, and sideways. Other multirotor UAVs, such as UAVs with six or eight rotors, have matching rotor sets that rotate in opposite directions. Therefore, it is possible to detect and identify UAVs by using the micro-Doppler effect of the rotors. In this study, we extract the characteristics of UAVs based on the micro-Doppler effect [21].

3. Detection and Identification Method of UAV

3.1. Scattering Model of UAV with Rotors. A UAV is a slow small target that can be simplified to a point target in traditional radar; however, since we plan to identify the UAV by using information on the rotors, we have to develop a model of a UAV in which the rotor is separate from the fuselage. Figure 4 shows the scattering model of a rotor-type UAV. The distance between the center of the rotor and the radar is R_0 and the elevation angle of the radar observation is β .

The fuselage scattering of the UAV can be obtained by (4). For a rotating point with a center of rotation of l and a rotational speed of Ω , the position relative to the radar at time t can be expressed as

$$R(t) = R_0 + l \cos \beta \cos (\Omega t + \varphi_0) \quad (9)$$

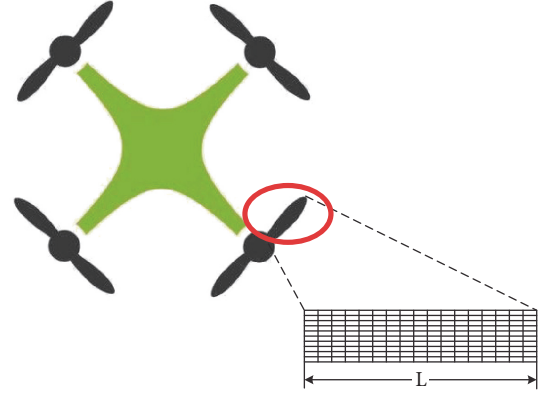


FIGURE 4: Scattering model of a rotor-type UAV.

Then the echo signal can be defined as

$$s_r(n, t) = \rho \exp \left\{ -j \left[2\pi (f_0 + n\Delta f) \times \frac{2(R_0 + l \cos \beta \cos (\Omega t + \varphi_0))}{c} \right] \right\} \quad (10)$$

We assume that the blade length is L as shown in Figure 3. Each blade can be thought of as a uniform, rigid rectangular plate that consists of arrays of rectangular facets with the scattering center of the surface of each facet assigned to its tip. The baseband signal originating from the rotors of a UAV can be expressed as

$$s_r(n, t) = \sum_{i=1}^{N_T} \sum_{m=1}^{N_M} \sum_{k=1}^{N_k} \rho_{i,m,k} \exp \left(-j \frac{4\pi (f_0 + n\Delta f)}{c} \times [R_{0i,k} + l_{i,m,k} \cos \beta_{i,k} \cos (\Omega_i t + \varphi_{0i,m})] \right) \quad (11)$$

where N_T is the total number of rotors, N_M is the number of blades per rotor, and N_k represents the number of facets.

3.2. HRRP of Rotor-Type UAV. The movement of the UAV can be divided into two parts, namely, the movement of the main body of the UAV and the rotation of the rotor. Since the time of a one-step sweep measurement is about hundreds of milliseconds and the UAV is a small slow target, its displacement during the one-step sweep measurement is negligible. Therefore, the UAV can be regarded as a stationary target during the one-step sweep measuring time. In this study, we do not consider the movement of the fuselage.

$$s_r(t) = \sum_{i=1}^{N_T} \sum_{m=1}^{N_M} \sum_{k=1}^{N_k} \rho_{i,m,k} \exp \left(-j \frac{4\pi (f_0 + \text{PRF}\Delta f t)}{c} \times [R_{0i,k} + l_{i,m,k} \cos \beta_{i,k} \cos (\Omega_i t + \varphi_{0i,k})] \right) \quad (12)$$

It is evident from (12) that the echo signal form of the rotor is an SFM signal and the equation can be expanded according to the following Bessel function:

$$e^{jA \sin \theta} = \sum_{n=-\infty}^{\infty} J_n(A) e^{jn\theta} \quad (13)$$

where $J_n(\cdot)$ is the first type of an n -order Bessel function.

The IDFT of the expanded rotor echo signal based on Eq. (13) can be obtained as [22]:

$$\begin{aligned} S_r(f) &= \sum_{i=1}^{N_T} \sum_{m=1}^{N_M} \sum_{k=1}^{N_k} \sum_{n=-N_1}^{N_1} b_{i,m,k,n} \times \delta(f - f_{R,i} - N_M n f_{\Omega,i}) \end{aligned} \quad (14)$$

where $f_{R,i}$ is the frequency offset caused by the position of the rotor center, $f_{\Omega,i}$ is the micro-Doppler frequency caused by the rotor speed, and b is the Bessel coefficient. The signal obtained by (14) is a superposition of the pulse signals of different amplitudes with a period of $N_M f_{\Omega}$, which is a comb-like function whose amplitude is modulated.

According to the characteristics of the Bessel function:

$$\sum_{m=-\infty}^{\infty} J_m^2(A) = 1 \quad (15)$$

The energy of the frequency-domain signal produces a divergence effect. The magnitude of the harmonic amplitude is modulated by the Bessel function. According to the Carson criterion [23], the number of harmonics is

$$N_1 \approx 2|A| + 3 \quad (16)$$

The width of spectral broadening caused by the rotation of the rotor is [22]

$$B \approx \frac{8\pi f_{\Omega} L \cos \beta}{N_M \lambda_0} \quad (17)$$

The rotational motion appears in the HRRP as a symmetrical range distribution centered on the position of the fuselage. When the rotor scattering is smaller than the fuselage scattering or slightly larger than the fuselage scattering, the amplitude of the rotor in the HRRP is much smaller than the amplitude of the fuselage because the rotor rotational motion is dispersed after the Bessel series expansion. For the purpose of detection, the influence of the rotor rotation on the HRRP can be neglected and the position information of the UAV can be obtained by directly reading the position of the main scattering from the fuselage. In most cases, the fuselage scattering is greater than the rotor scattering and the effects of the rotor rotation do not have to be considered for obtaining the UAV position.

When the rotor scattering is much larger than the fuselage scattering, the main body scattering is submerged in the distance harmonic caused by the rotation of the rotor. In this case, the harmonic boundary caused by the rotation can be estimated and the center position of the two boundary positions can be used to estimate the target position. Alternatively, the SFM parameter optimization method described below can be used to compensate for the effect of rotor rotation on the HRRP.

3.3. UAV Identification

3.3.1. Micro-Doppler Effect of UAV with Rotors. The Doppler shift of the received signal can be expressed as

$$\begin{aligned} f_D &= \frac{1}{2\pi} \frac{d\varphi(t)}{dt} = \frac{2}{c} \{ [f_0 + PRF \\ &\times \Delta f t] [-\Omega L \cos \beta \sin(\Omega t \\ &+ \varphi_0)] + \Delta f (L \cos \beta \cos(\Omega t + \varphi_0) + R_0) \times PRF \} \\ &= \frac{2}{c} \left\{ L \cos \beta \sqrt{(\Omega f_0 + \Omega PRF \times \Delta f t)^2 + (\Delta f \times PRF)^2} \right. \\ &\left. \cdot \cos(\Omega t + \theta) + \Delta f R_0 \times PRF \right\} \end{aligned} \quad (18)$$

The micro-Doppler frequency shift changes over time and the joint time-domain analysis method is the most intuitive means to observe the target's micro-Doppler effect. Short-time Fourier transform (STFT) [24] is a commonly used joint time domain analysis methods. The basic process of the STFT is to decompose a long time signal into shorter segments of the same length, calculate the Fourier transform on each short segment, and visualize the variation of the echo frequency over time.

3.3.2. Micro-Doppler Parameter Extraction Algorithm. The extraction of the parameters of the rotor, such as the number of rotors and the speed, is used to identify the type of UAV. For single-rotor UAVs, the STFT can be used for the identification. For multirotor UAVs, the micro-Doppler frequencies of different rotors are aliased together, which makes the target difficult to resolve and requires a new extraction method.

It is evident in (14) that the echo signals are superimposed on the pulse signals of different amplitudes in the frequency domain with $N_M f_{\Omega}$ as the period. The period of the signal is related to the rotor speed and the rotors of different speeds produce comb-like functions of different periods. The number of rotors can be determined by extracting the number of different periodic signals and the rotational speed of each rotor of the UAV is estimated by extracting the period of the signal.

Cepstrum is an effective method used in speech signal processing [25, 26] for the periodic detection of the pitch and is now also commonly used for mechanical state detection and fault diagnosis. In this study, the power cepstrum is used for parameter extraction. By multiplying the signal by its own conjugate signal, the value of the micro-Doppler can be doubled [11], which contributes to providing better periodic signal extraction. We set $g(t)$ to

$$g(t) = s(t) * s^*(-t) \quad (19)$$

A Fourier transform is performed on $g(t)$ and the result is expressed as $G(f)$. $G(f)$ can be regarded as a comb function $C(f)$ that is multiplied by some form of amplitude modulation $A(f)$. The results of $G(f)$ are shown in Figure 5.

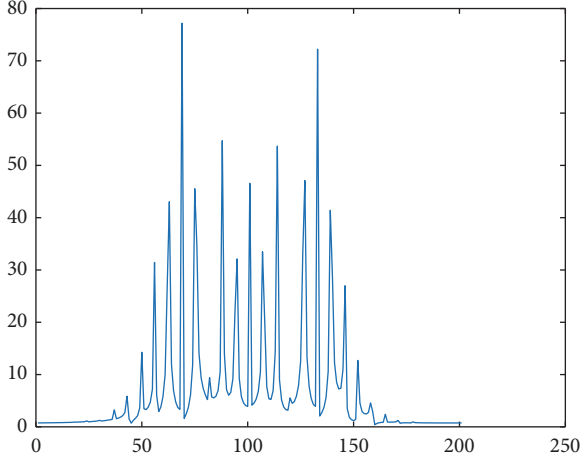
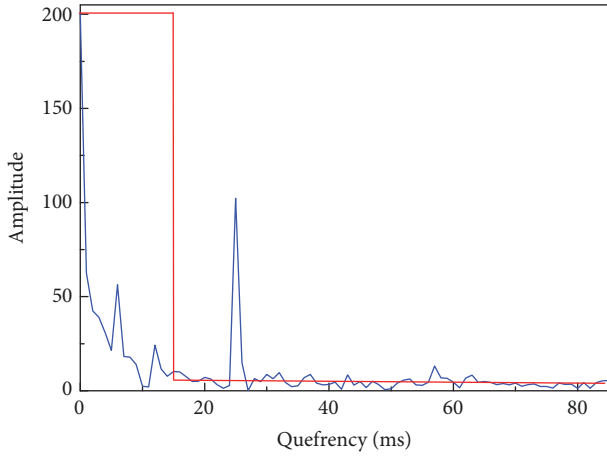
FIGURE 5: Function diagram of $G(f)$.

FIGURE 6: Cepstrum domain.

By calculating the log spectrum, the multiplicative relationship between the comb function and amplitude modulation is converted into a summation. Then, the inverse Fourier transform is performed and the square of the amplitude is used to obtain the cepstrum of the function.

The specific equation is as follows:

$$C(q) = |\text{IFFT} \{ \log |\text{FFT}(g(t))| \}|^2 \quad (20)$$

The cepstrum domain is shown in Figure 6; the modulated function $A(f)$ is a nonperiodic function whose peak appears at the zero value of the cepstrum and $G(f)$ appears as a peak at its periodic frequency. By using a high-pass filter, the influence of $G(f)$ can be removed and the rotation period of the UAV is obtained.

3.4. Estimation of Multi-UAV Positions. The rotor rotation widens the HRRP distance. If the distance between two targets is large enough so that they can be separated in the HRRP, each target can be separately processed to determine whether it is a UAV. If the distance between the targets is very small due to distance expansion by rotation, a multi-SFM

TABLE I: Variable simplification.

Initial parameter	Simplified parameters
$\Delta f \frac{2R_0}{c}$	f_c
$(f_0 + n\Delta f) \frac{2l \cos \beta \cos(\Omega n / \text{PRF} + \varphi_0)}{c}$	m_{fi}
$f_0 \frac{2R_0}{c}$	θ
$\frac{\Omega}{\text{PRF}}$	f_{mi}

signal separation method based on parameter optimization is used to compensate for the rotational motion.

3.4.1. Multi-SFM Signal Separation Based on Parameter Optimization. We discretize (12):

$$s_r(n) = \sum_{i=1}^{N_T} \sum_{m=1}^{N_M} \sum_{k=1}^{N_k} \rho_{i,m,k} \exp \left(-j \frac{4\pi (f_0 + n\Delta f)}{c} \right) \times \left[R_{0i,k} + l_{i,m,k} \cos \beta_{i,k} \cos \left(\frac{\Omega_i n}{\text{PRF}} + \varphi_{0i,k} \right) \right] \quad (21)$$

Table 1 shows the simplified variables in (21):

Equation (21) can be expressed in the form of multiple SFMs:

$$x(n) = \sum_{i=1}^M A_i \exp \left\{ -j \left[2\pi (f_{ci} n + m_{fi} \cos(2\pi f_{mi} n + \varphi_i) + \theta_i) \right] \right\} \quad (22)$$

The unknown parameters in (22) include the center frequency f_{ci} , frequency modulation factor m_{fi} , modulation period f_{mi} , initial phase of the SFM signal φ_i , initial phase of the FM signal θ_i , and the amplitude A_i . f_{ci} is related to the target position and thus we mainly focus on the estimation of f_{ci} from different signals. The parameter optimization method estimates the parameter estimation order of f_{ci} .

The modulation period of the signal can be obtained by the cepstrum method and will not be described here.

(a) Estimation of the modulation period f_{mi} .

(b) Estimation of the initial phase φ_i and the frequency modulation factor m_{fi} .

We assume $\varphi_i \in [0, 2\pi]$, $m_{fi} \in [m1, m2]$. The initial estimated value $\varphi_i = 0$, $m_{fi} = m1$. The reference signal can be expressed as

$$\tilde{x}_{m_i, \varphi_i}(n) = \exp \left\{ j \left[2\pi \hat{m}_{fi} \cos(2\pi \hat{f}_{mi} n + \hat{\varphi}_i) \right] \right\} \quad (23)$$

Equation (23) is multiplied by (22):

$$x(n) \tilde{x}_{m, \varphi}(n) = \sum_{i=1}^M A_i \cdot \exp \left\{ -j \left[2\pi (f_{ci} n + m_{fi} \cos(2\pi f_{mi} n + \varphi_i)) \right] + \theta_i \right\} + j \left[2\pi \hat{m}_{fi} \cos(2\pi \hat{f}_{mi} n + \hat{\varphi}_i) \right] \quad (24)$$

For the i th SFM component,

$$\begin{aligned}
 x_i(n) \tilde{x}_{mi,\varphi_i}(n) &= A_i \\
 &\cdot \exp \left\{ -j \left[2\pi f_{ci} n \left(+m_{fi} \cos(2\pi f_{mi} n + \varphi_i) + \theta_i \right) \right] \right. \\
 &+ j \left[2\pi \hat{m}_{fi} \cos(2\pi \hat{f}_{mi} n + \hat{\varphi}_i) \right] \left. \right\} = A_i \\
 &\cdot \exp \left\{ -j(2\pi f_{ci} n) + \theta_i \right\} \exp \left\{ -j2\pi \left[-2m_{fi} \right. \right. \\
 &\cdot \sin \left(2\pi f_{mi} n + \frac{\varphi_i + \hat{\varphi}_i}{2} \sin \left(\frac{\varphi_i - \hat{\varphi}_i}{2} \right) \right) \\
 &\left. \left. - (\hat{m}_{fi} - m_{fi}) \cos(2\pi \hat{f}_{mi} n + \hat{\varphi}_i) \right] \right\} \quad (25)
 \end{aligned}$$

(c) Estimation of \hat{f}_{ci} : when $\varphi_i = \hat{\varphi}_i$, the spectrum of the signal is concentrated in a band of width $2|\hat{m}_{fi} - m_{fi}| \hat{f}_{mi}$. When $\hat{m}_{fi} = m_{fi}$, the rotational motion is compensated. The original signal becomes

$$A_i \exp \left\{ -j(2\pi f_{ci} n) + \theta_i \right\} \quad (26)$$

Fast Fourier transform (FFT) is performed on (26); the position corresponding to the maximum value after the transformation is \hat{f}_{ci} , which can be expressed as

$$\hat{f}_{ci} = \arg \max_f \left| \text{FFT} \left(x_i(n) \tilde{x}_{mi,\varphi_i}(n) \right) \right| \quad (27)$$

3.4.2. Rotor Position Estimation. The steps for identifying close-range multi-UAVs using the SFM parameter optimization method are as follows:

(a) The cepstrum method is used to calculate the rotation period of each rotating target.

(b) Since the fuselage scattering will affect the SFM parameter optimization method, the influence of the subject scattering must be removed prior to using the SFM parameter optimization method. First, two subsequent frames of signals are measured and subtracted from each other for the cancellation of the fuselage scattering. Second, HRRP is used to determine the maximum value of the rotor scattering. Finally, the maximum value is set as a threshold. If the single frame in the HRRP is larger than the threshold, the single signal is set as the mean value of the subtracted HRRP.

(c) The SFM parameter optimization method is used for the rotation periods of the rotors obtained in step (a) to determine their position.

4. Simulation

4.1. Simulation Model. Figure 7 shows a schematic of the simulation model. There are three targets in the observation area, i.e., a quad-rotor UAV at a distance of 300 m from the radar, a non-UAV target 800 m from the radar, and a single-rotor UAV 850 m from the radar. The rotor length of the quad-rotor UAV is 0.12 m and the rotational speeds are 158 r/s, 150 r/s, 142 r/s, and 139 r/s, respectively. The rotor length of the

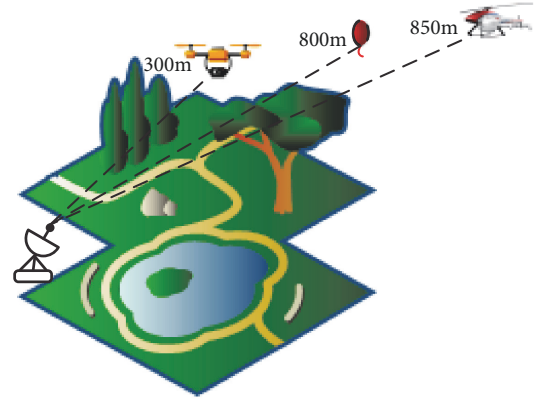


FIGURE 7: Schematic of the simulation model.

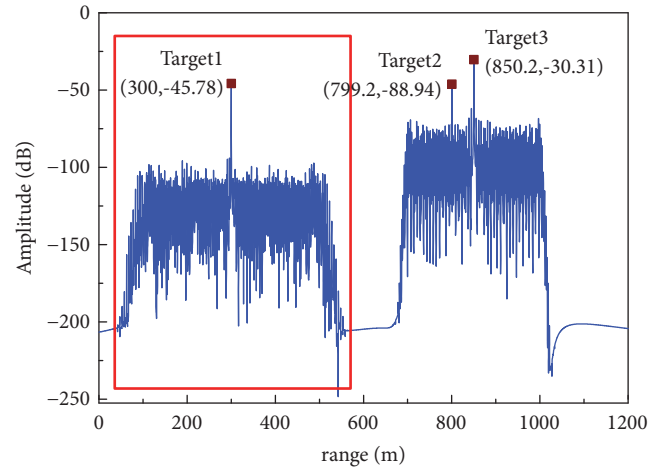


FIGURE 8: HRRP of the target.

single-rotor UAV is 0.15 m and the rotational speed is 88 r/s. The parameters of the radar are as follows: starting frequency $F_0=25$ GHz, which is in the 5G band; signal bandwidth $BW = 500$ MHz, the resolution cell is 0.3m; sampling points $N = 8001$; $PRF=150$ kHz.

4.2. Simulation Results. The HRRP of the echo signal is shown in Figure 8. Three targets are detected at distances of 300 m, 799.2 m, and 850.2 m from the radar. Target 2 and target 3 cannot be separated in the HRRP because they are in close proximity and have to be analyzed as one target. The SFM optimization method can be used to determine the UAV locations. Target 1 is far from target 2 and target 3. Therefore, the time-domain echo signal of target 1 can be separated and identified by using a distance-domain window and FFT.

4.2.1. Identification of Target 1. Figure 9 shows the result of the micro-Doppler characteristics of target 1. It can be inferred that target 1 is a rotor-type UAV. Due to the presence of the stepping frequency, the echo signal appears as a fixed frequency value that does not change over time in the time-frequency image. The micro-Doppler frequency caused by

TABLE 2: Target parameters.

	Target type	Position (m)	Rotor speed (r/s)
Target 1	Quad-rotor UAV	300	157.9, 150.0, 142.0, 138.9
Target 2	Other targets	799.2	-
Target 3	One-rotor UAV	850.2	88

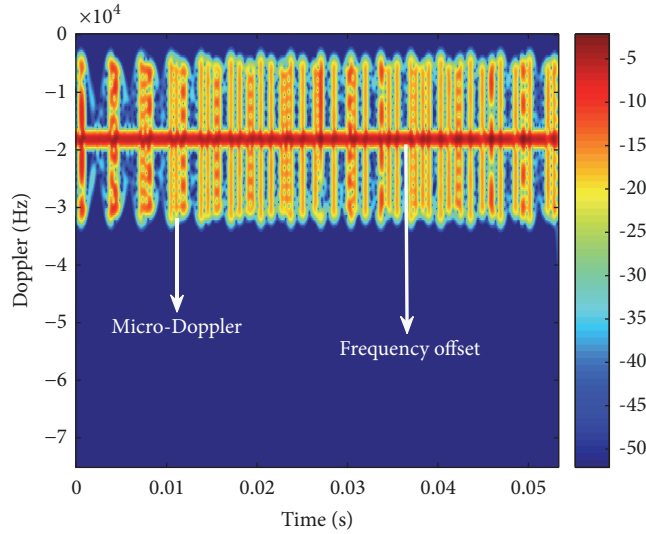


FIGURE 9: Micro-Doppler of target 1.

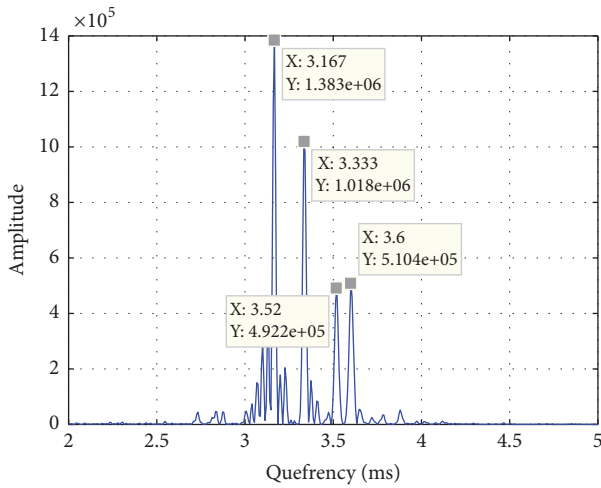


FIGURE 10: Cepstrum extraction of target 1.

the rotor changes sinusoidally over time. The difference in the amplitude of the micro-Doppler at different moments is due to the different contributions of rotor scattering to echoes at different radar observation angles.

There are multiple rotors in this target; therefore, the micro-Doppler signal is aliased and the type parameters of the UAV rotor cannot be directly obtained. The cepstrum method is used to determine the rotor type.

Figure 10 shows the cepstrum extraction results. It can be seen that target 1 is a UAV with 4 rotors. The rotational speed of the rotors is 157.9 r/s, 150.0 r/s, 142.0 r/s, and 138.9 r/s. The speed obtained by the cepstrum method is nearly the same as the predetermined speed of the rotors.

In order to verify the robustness of the cepstrum method, noise with different signal-to-noise ratios (SNRs) is added to the echo signal of target 1. Figure 11 shows the simulation results with SNR = -5, SNR = 0, and SNR = 5. It is evident from Figure 11 that after adding the noise, the peak energy of the position corresponding to the rotation period of the rotor is slightly lower. However, there is still a significant difference between the rotor values and the values of other positions and the addition of the noise has no influence on the identification of the number of rotors and the rotational speed.

4.2.2. Identifications of Target 2 and Target 3. Target 2 and target 3 are analyzed as a single target. Figure 12 shows the micro-Doppler characteristics of the two targets.

As shown in Figure 12, there is one rotor for target 2 and target 3. Figure 13 shows that the rotational speed of this rotor obtained by the cepstrum method is 88 r/s.

The position of the single rotor is estimated by the SFM parameter optimization method after the two target body scatterings are removed. The estimated results are shown in Figure 14. The parameters of the three targets in the detection area are shown in Table 2.

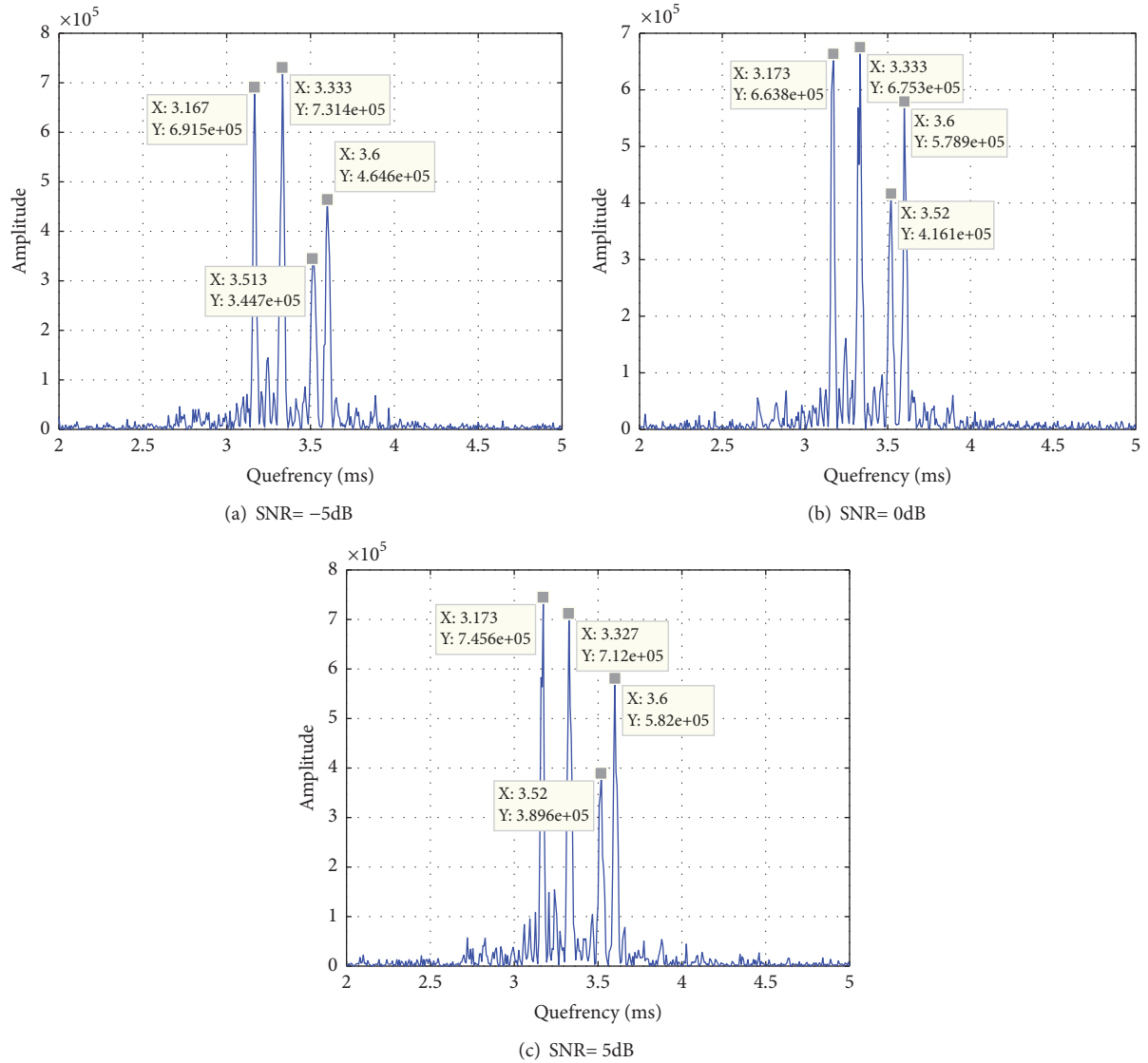


FIGURE 11: Cepstrum results with the addition of different SNRs.

Table 2 shows the good accuracy of the detection of the UAV and its position. Figure 15 shows a comparison of the UAVs predetermined route based on the GNSS signal and the actual route based on radar tracking. In the case of a GNSS spoofing attack, the captured UAV has flown away from its predetermined route but the original control source was not alerted because it received data that appeared to be correct from the attackers. In this case, the radar-assisted detection and identification method can be used for real-time monitoring of the UAV independent from the GNSS signal; the deviation between the actual route and predetermined route can be determined and an alert can be sent. It is evident that there is a considerable deviation between the GPS-based positions and the radar positions. The latter method provides accurate positioning of the UAV and allows for effective monitoring of UAVs in the IoT.

5. Conclusion

In this paper, a radar-assisted UAV detection and identification method that is independent of a GPS system was proposed. The HRRP technology of wideband radar was used for UAV detection and positioning and the micro-Doppler signal, which is capable of detecting rotating targets, was used for UAV identification. Cepstrum analysis was used to extract the number and rotation speed of the UAV rotors. The simulation results demonstrated the good robustness of this method. An SFM parameter optimization algorithm was used to compensate for the rotational motion and estimate the rotor positions; two UAVs with aliasing signals in the HRRP were separated effectively using the proposed method. The simulation results showed that the UAVs can be identified and the number of UAVs, the number of rotors on each UAV,

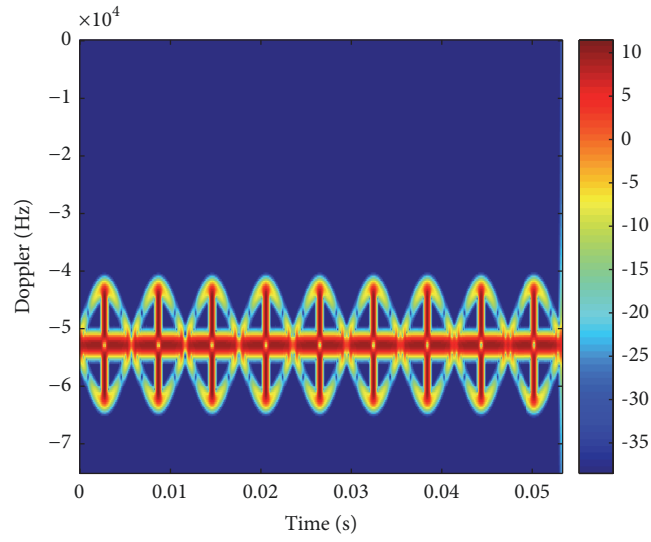


FIGURE 12: Micro-Doppler of target 2 and target 3.

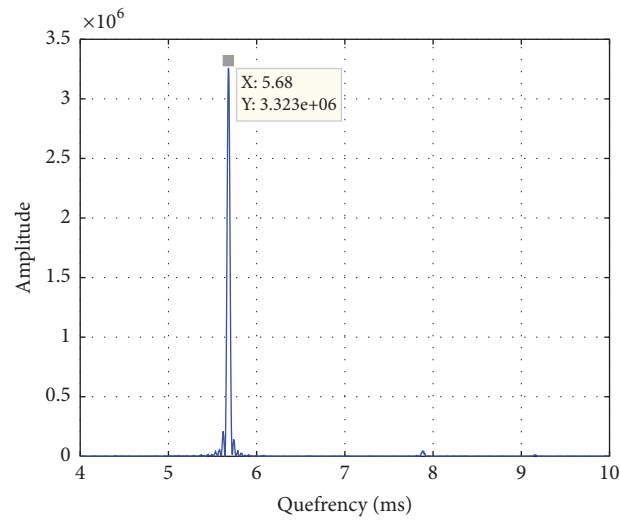


FIGURE 13: Cepstrum result of target 2 and target 3.

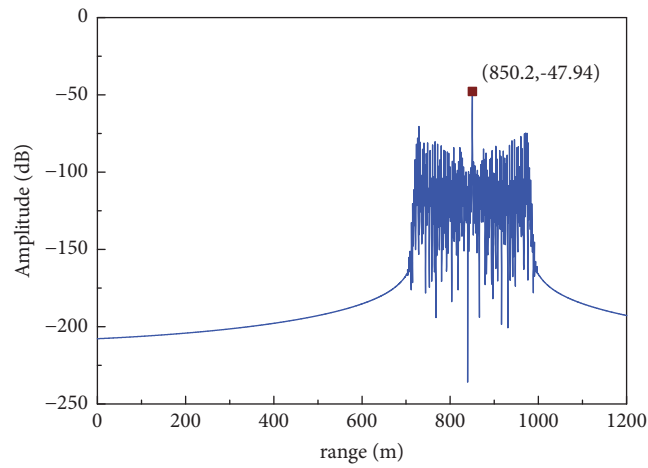


FIGURE 14: Rotor position estimation.

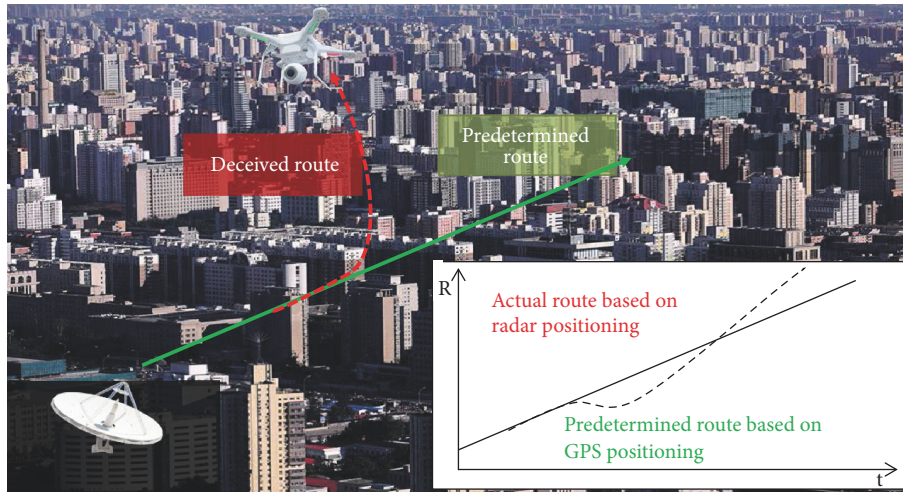


FIGURE 15: UAV flight route.

and the rotation speeds of the rotors can be determined. In addition, the proposed radar-assisted UAV detection and identification method can provide alerts if UAVs deviate from their routes during GNSS spoofing attacks.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.



References

- [1] A. Fotouhi, M. Ding, and M. Hassan, "Understanding autonomous drone maneuverability for internet of things applications," in *Proceedings of the 18th IEEE International Symposium on A World of Wireless, Mobile and Multimedia Networks, WoWMoM 2017*, IEEE, China, June 2017.
- [2] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Mobile unmanned aerial vehicles (UAVs) for energy-efficient internet of things communications," *IEEE Transactions on Wireless Communications*, vol. 16, no. 11, pp. 7574–7589, 2017.
- [3] A. Giyenko and Y. I. Cho, "Intelligent UAV in smart cities using IoT," in *Proceedings of the 16th International Conference on Control, Automation and Systems, ICCAS 2016*, pp. 207–210, IEEE, Republic of Korea, 2016.
- [4] M. Rieke, T. Foerster, J. Geipel, and T. Prinz, "High-precision positioning and real-time data processing of UAV systems," *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 38, no. 1, article no. C22, pp. 119–124, 2011.
- [5] V. M. Sinoglazov and E. V. Daskal, "IEEE 802.15.4/zigbee wireless network as a basis for uav navigation system," *Electronics and Control Systems*, vol. 2, no. 52, pp. 20–24, 2017.
- [6] A. J. Kerns, D. P. Shepard, J. A. Bhatti, and T. E. Humphreys, "Unmanned aircraft capture and control via GPS spoofing," *Journal of Field Robotics*, vol. 31, no. 4, pp. 617–636, 2014.
- [7] V. C. Chen, F. Li, S.-S. Ho, and H. Wechsler, "Micro-doppler effect in radar: Phenomenon, model, and simulation study," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 42, no. 1, pp. 2–21, 2006.
- [8] D. Solomitskii, M. Gapeyenko, V. Semkin, S. Andreev, and Y. Koucheryavy, "Technologies for efficient amateur drone detection in 5G millimeter-wave cellular infrastructure," *IEEE Communications Magazine*, vol. 56, no. 1, pp. 43–50, 2018.
- [9] S. Han, Y. Huang, W. Meng, C. Li, N. Xu, and D. Chen, "Optimal power allocation for scma downlink systems based on maximum capacity," *IEEE Transactions on Communications*, vol. 67, no. 2, pp. 1480–1489, 2019.
- [10] S. Han, Y. Zhang, W. Meng, and H. Chen, "Self-interference-cancellation-based SLNR precoding design for full-duplex relay-assisted system," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 9, pp. 8249–8262, 2018.
- [11] J. Martinez, D. Kopyto, M. Schutz, and M. Vossiek, "Convolutional neural network assisted detection and localization of UAVs with a narrowband multi-site radar," in *Proceedings of the 2018 IEEE MTT-S International Conference on Microwaves for Intelligent Mobility, ICMIM 2018*, IEEE, Germany, April 2018.
- [12] Y. Wang, T. Phelps, K. Kibaroglu, M. Sayginer, Q. Ma, and G. M. Rebeiz, "28 GHz 5G-based phased-arrays for UAV detection and automotive traffic-monitoring radars," in *Proceedings of the 2018 IEEE/MTT-S International Microwave Symposium - IMS*, pp. 895–898, IEEE, Philadelphia, PA, USA, 2018.
- [13] S. Sun, M. Kadoch, L. Gong, and B. Rong, "Integrating network function virtualization with SDR and SDN for 4G/5G networks," *IEEE Network*, vol. 29, no. 3, pp. 54–59, 2015.
- [14] T. Yang, J. Zhao, T. Hong, W. Chen, and X. Fu, "Automatic identification technology of rotor UAVs based on 5G network architecture," in *Proceedings of the 2018 IEEE International Conference on Networking, Architecture and Storage (NAS)*, pp. 1–9, IEEE, 2018.
- [15] S. Han, S. Xu, W. Meng, and C. Li, "An agile confidential transmission strategy combining big data driven cluster and OBF," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 11, pp. 10259–10270, 2017.

- [16] B. Rong, Y. Qian, K. Lu, H. Chen, and M. Guizani, "Call admission control optimization in WiMAX networks," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 4, pp. 2509–2522, 2008.
- [17] S. Han, S. Xu, W. Meng, and C. Li, "Dense-device-enabled cooperative networks for efficient and secure transmission," *IEEE Network*, vol. 32, no. 2, pp. 100–106, 2018.
- [18] Y. Wu, B. Rong, K. Salehian, and G. Gagnon, "Cloud transmission: a new spectrum-reuse friendly digital terrestrial broadcasting transmission system," *IEEE Transactions on Broadcasting*, vol. 58, no. 3, pp. 329–337, 2012.
- [19] B. R. Mahafza, *Radar Systems Analysis and Design Using MATLAB*, Chapman and Hall/CRC, 2nd edition, 2005.
- [20] V. C. Chen, F. Li, S. S. Ho, and H. Wechsler, "Analysis of micro-doppler signatures," *IEEE Proceedings-Radar, Sonar and Navigation*, vol. 150, no. 4, pp. 271–276, 2003.
- [21] M. Jian, Z. Lu, and V. C. Chen, "Experimental study on radar micro-Doppler signatures of unmanned aerial vehicles," in *Proceedings of the 2017 IEEE Radar Conference, RadarConf 2017*, pp. 0854–0857, IEEE, USA, May 2017.
- [22] G. E. Smith, K. Woodbridge, and C. J. Baker, "Radar micro-doppler signature classification using dynamic time warping," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 46, no. 3, pp. 1078–1096, 2010.
- [23] C. Weidong, X. Shanjia, W. Dongjin, and L. Falin, "Range performance analysis in linear FMCW radar," in *Proceedings of ICMMT 2000. 2000 2nd International Conference on Microwave and Millimeter Wave Technology Proceedings (Cat. No. 00EX364)*, pp. 654–657, IEEE, Beijing, China, 2000.
- [24] Y. Wang, S. Ji, and H. Xu, "Non-stationary signals processing based on STFT," in *Proceedings of the 2007 8th International Conference on Electronic Measurement and Instruments, ICEMI*, pp. 3301–3304, IEEE, China, August 2007.
- [25] H. K. Kim and R. C. Rose, "Cepstrum-domain acoustic feature compensation based on decomposition of speech and noise for asr in noisy environments," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 11, no. 5, pp. 435–446, 2003.
- [26] S. Furui, "Speaker-independent isolated word recognition using dynamic features of speech spectrum," *IEEE Transactions on Signal Processing*, vol. 34, no. 1, pp. 52–59, 1986.

Research Article

Dynamic Traffic Prediction with Adaptive Sampling for 5G HetNet IoT Applications

Shuangli Wu ^{1,2}, Wei Mao,^{1,3} Cong Liu ⁴, and Tao Tang⁴

¹Computer Network Information Center, Chinese Academy of Sciences, Beijing, China

²University of Chinese Academy of Sciences, Beijing, China

³Internet Domain Name System Beijing Engineering Research Center, Beijing, China

⁴Beihang University, Beijing, China

Correspondence should be addressed to Shuangli Wu; wushuangli@cnic.cn

Received 1 April 2019; Accepted 12 June 2019; Published 26 June 2019

Guest Editor: Bo Rong

Copyright © 2019 Shuangli Wu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Due to the proliferation of global monitoring sensors, the Internet of Things (IoT) is widely used to build smart cities and smart homes. 5G HetNets play an important role in the IoT video stream. This paper proposes an improved Call Session Control Function (CSCF) scheme. The improved CSCF server contains additional modules to facilitate IoT traffic prediction and resource reservation. We highlight traffic prediction in this work and develop a compressed sensing based linear predictor to catch the traffic patterns. Experimental results justify that our proposed scheme can forecast the traffic load with high accuracy but low sampling overhead.

1. Introduction

The Internet of Things (IoT) makes it possible to connect various physical devices. Vehicles, buildings, etc. embedded in sensors are connected through the Internet of Things, and data can be collected and exchanged via the Internet. The Internet of Things enables mutual communication between devices. In the next decade, the service targets in the Internet of Things will be improved to users in various industries, and the number of machine to machine (M2M) terminals will increase dramatically, and applications will be ubiquitous.

In order to solve the problem of explosive data volume of the IoT, a heterogeneous networks (HetNets) technology is proposed. The problem of 5G communication is solved by deploying a large number of small cells [1–3]. There are several technical challenges hindering the IoT video streaming over 5G HetNets. For example, the interactions between HetNet and 5G optical core tremendously depend on the IoT traffic from the local HetNet, which may change periodically for many reasons. Consequently, the system must have a forecast and reservation scheme to dynamically book necessary resource from optical core. In addition, a connection admission control (CAC) mechanism plays an essential role in case the actually required resource overpasses

the reserved. Traffic on future fifth-generation (5G) mobile networks is predicted to be dominated by challenging video applications such as mobile broadcasting, remote surgery, and augmented reality, demanding real-time and ultra-high quality delivery. Two of the main expectations of 5G networks are that they will be able to handle ultra-high-definition (UHD) video streaming and that they will deliver services that meet the requirements of the end user's perceived quality by adopting quality of experience (QoE) aware network management approaches.

To overcome the above challenges, this paper proposes a scheme of improved CSCF server for 5G HetNets. The improved CSCF can reside itself in a mobile edge computing (MEC) server, with three extra components added to the original, including IoT traffic prediction, bandwidth negotiation, and connection admission control [4].

The improved CSCF first forecasts the IoT traffic load from a local HetNet to the 5G optical core. After that, it employs Common Open Policy Service (COPS) protocol [5] to adaptively negotiate bandwidth resource. Finally, the connection admission control serves to keep the system in shape and guarantee communication quality.

The IoT is responsible for the communication between different devices, providing better solutions and enhanced

services for different information flows. Video signal is very important among multimedia services. At the same time, the video traffic on the Internet has increased dramatically, and a lot of research work has been done to improve the viewing experience of the audience.

Video traffic between HetNet and optical core often fluctuates because of use case's characteristic and variation of observed object. Therefore, our proposed improved CSCF server chooses a dynamic mechanism, in a bid to book bandwidth economically. Apparently the traffic prediction component of improved CSCF server has critical impact on the performance of dynamic bandwidth negotiation. In practice, sampling overhead is a real challenge, as traffic predictor must talk to two other components of the improved CSCF server during the sampling process. For instance, traffic predictor needs the calling records from connection admission control component as the input. After the prediction, it has to transfer the result to bandwidth negotiation component, which will further use COPS messages to negotiate with 5G core network. The procedure above can strongly justify the significance of sampling rate reduction.

Compressed sensing is a promising paradigm that uses signal sparsity to reduce the amount of data that needs to be measured [6, 7]. Compressed sensing (CS) theory indicates that the characteristics of sparse signals are not affected by the signal sparse basis. It can be obtained by a small amount of projection on another basis. The reconstruction of the wideband signal can be obtained by solving the L1 norm optimization. It is concluded that, in the calculation process, the sparsity of the signals in the orthogonal basis determines the minimum projection number and sampling rate required [8].

In the literature, the study of traffic prediction mainly concentrated on the least mean square (LMS) methods that take sampling and prediction as two independent aspects [9–12]. As a result, the sampling often implements a constant interval and is too prone to low efficiency in IoT environment.

Other than previous work, this paper focuses on the aggregated multiple streams in a HetNet, instead of one separated video stream.

Inspired by compressed sensing theory, we develop an adaptive sampling rate (ASR) linear predictor to overcome the shortcoming of traditional schemes. ASR linear predictor utilizes the prediction error to control sampling interval, so that the sampling and predicting overheads are significantly reduced.

The rest of the paper is organized as the remainder. Section 2 introduces 5G IoT applications in smart cities. Section 3 introduces 5G HetNets and IMS based IoT video streaming. Section 4 develops traffic prediction schemes using adaptive sampling rate and compressed sensing theory. Finally, Section 5 summarizes the advantages of our proposed improved CSCF server and ASR linear predictor.

2. 5G IoT Applications in Smart Cities

2.1. Smart Cities. The development of the IoTs and cloud computing technology has become an important symbol of

our modern information age, especially in the city, providing technical support for the commercial development of the city. The comprehensive and multilevel processing of information technology in the IoTs provides a service platform for data systems. Therefore, in the smart city information system, the digital information processing of the city can timely understand the relevant information of the city.

The use of IoT technology to improve and develop smart city information technology, shield bad information, and guide and support the exchange of useful information enables it to cover all corners of the city in a network and realizes the management and control of the city as a whole, as shown in Figure 1. At the same time, with the help of the intelligent engine and the unified information service system, the management efficiency of urbanization construction has been effectively improved.

For smart cities, their construction and development are inseparable from the IoT communication technology. It can be said that IoT communication technology is an important component of its foundation. One of the important applications is intelligent building management, which is mainly used in the following aspects.

(a) The installation of intelligent building management technology during the construction of the building will enable residents to conveniently manage and control remote items under the cooperation of the IoT.

(b) The establishment of a large public monitoring system: there are many public buildings in the city and there are many equipment and facilities. Therefore, an emergency monitoring system must be established to solve this problem quickly with the assistance of the IoT when unexpected events occur.

(c) Adopt intelligent water, coal, and electricity meters. The RFID chip is built in the table of various facilities, and the data can be collected into the data processor through the table and then transmitted to the background system through the 5G network to realize the function of automatic meter reading.

(d) Processing of abnormal equipment: remote monitoring of air conditioners, lights, etc. is carried out using wireless sensors and local area network technology during the processing of this problem.

2.2. IoT Intelligent Parking System in Smart City. With the development of economy, there are more and more vehicles in the city, and the problem of parking difficulty is followed. Smart cities can bring effective solutions to this problem. In the smart city, an intelligent parking system based on Internet of Things is proposed. The framework of the system is shown in Figure 2.

This system combines advanced IoT technology, mobile internet technology, and big data technology. As shown in the figure, the system is mainly composed of the video pile, the base station, the server, the mobile client, and the management cloud platform.

The video pile is mainly composed of a camera, a processor, and a communication device. While the system is running, the camera is used to capture images of each parking space and monitor the status of each parking space

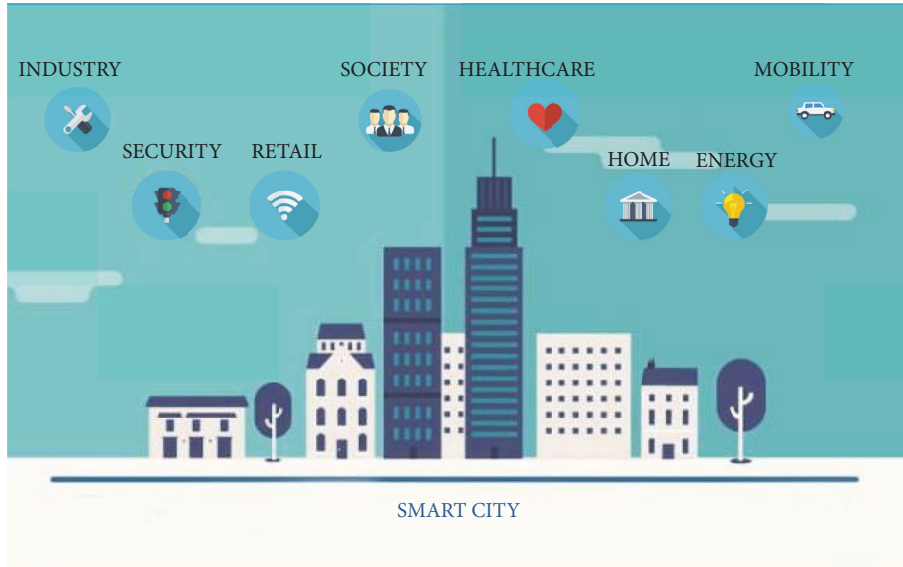


FIGURE 1: 5G IoT application in smart cities.

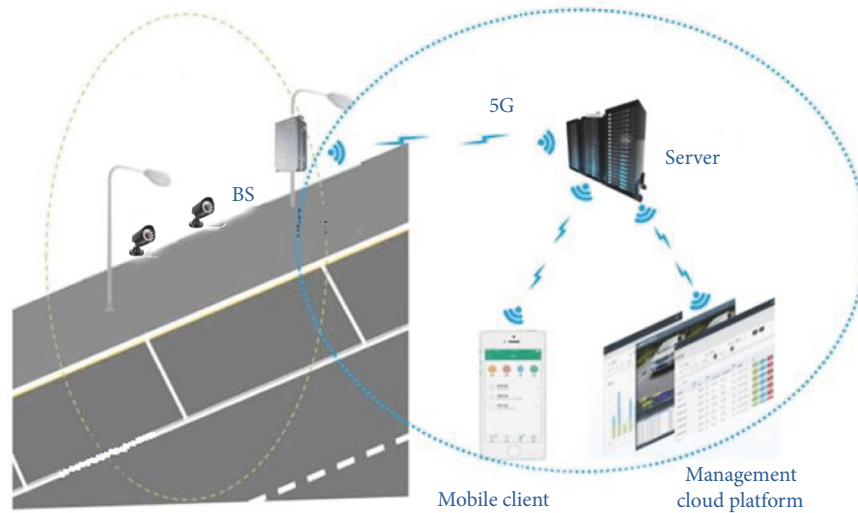


FIGURE 2: Smart parking system.

in real time through target recognition technology. Then, the video pile transmits the information to the base station via the wireless network. The base station forwards the information from different video pile to the server and the cloud management platform. The whole system adopts intelligent management, and the big data technology is adopted in the cloud management platform to carry out regional level parking space scheduling to prevent the problem of uneven resource allocation. The user can query the remaining number of parking spaces and make a parking space reservation in real time through the mobile phone client, and if the user is not familiar with the terrain, the system can recommend the best parking space for the user and provide precise navigation services. Throughout the service process,

the system performs automatic billing and supports mobile client online payment.

3. IoT Video Communication over Multilocation HetNets

3.1. 5G Core Networks Bridged by MPLS VPN. In the IoTs, HetNets networks consist of eNode BS and low-power nodes (LPNs), a.k.a. access points, a unit that makes it up. Different layers are covered by the two nodes according to the transmission power [13, 14]. This approach densifies the topology and improves space utilization and spectrum reuse. The structure of the heterogeneous network satisfies the requirements of 5G technology and greatly improves the spectral efficiency [15].

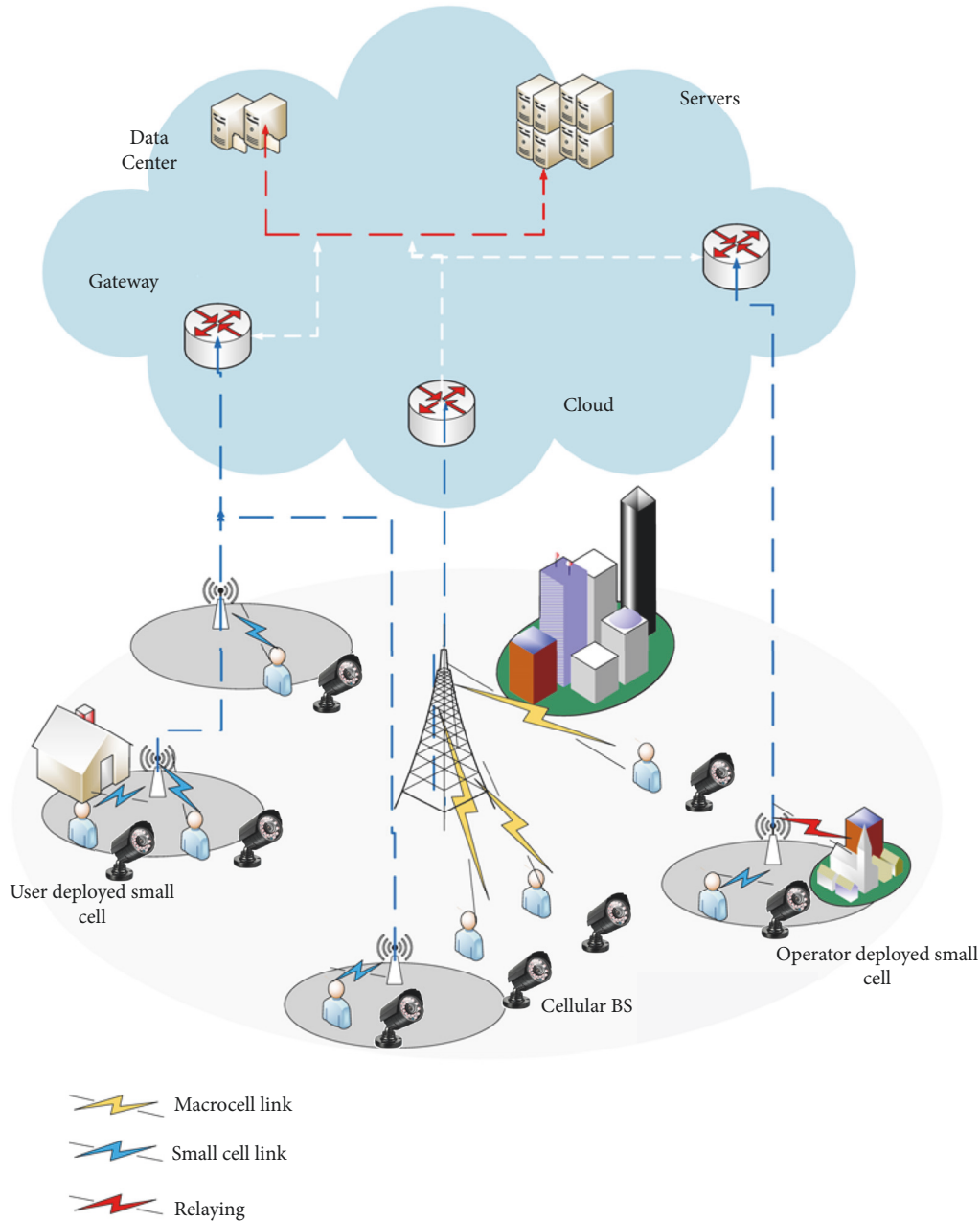


FIGURE 3: A typical 5G IoT heterogeneous network.

In order to solve this problem, 5G wireless networks have received attention and research in order to provide connectivity and meet the requirements of different quality of service (QoS). The key technologies of 5G enable the deployment of heterogeneous networks (HetNets) to support a large number of IoT data requirements by deploying a large number of small cells.

Network density is the most effective means in many ways to increase network capacity, such as spectrum expansion and spectrum efficiency enhancement [16]. By densely deploying small cells indoors and outdoors, network capacity is increased and network latency is reduced. Small cells include

microcells, picocells, and femtocells, which are closer to the user. Figure 3 is a schematic diagram of deployment of a small cell, which can be widely deployed indoors and outdoors, such as in offices, intersections, and plazas. Among them, the data traffic of indoor users can be provided by Wi-Fi. The next generation of Wi-Fi 802.11ac is expected to grow rapidly, providing multigigabit transmission rates. This led to the concept of HetNets, a multilayer network with multiple radio access technologies.

A 5G core network could be enhanced by network slicing, which allows different clients to bridge their private sites over a common provider-owned cloud. Using MPLS or another

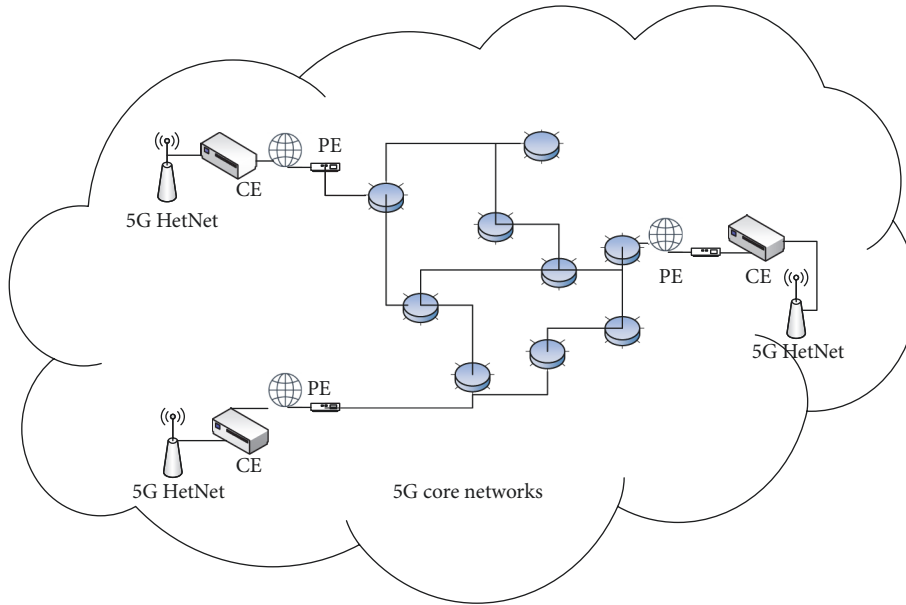


FIGURE 4: Network slicing based 5G core network connecting several local HetNets.

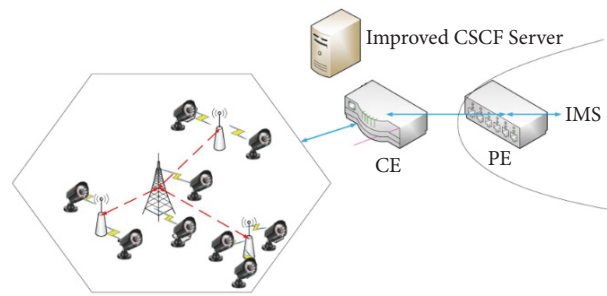


FIGURE 5: IMS based IoT video streaming over 5G HetNets.

similar technology [17], network slicing becomes a novel trend and has a wonderful capability to offer QoS guaranteed services with flexibility. Figure 4 shows an example of network slicing based 5G core that connects several local HetNets. Using MPLS technology, the core network contains the following types of equipment:

- (a) customer edge (CE) router, which serves to connect with client local network directly
- (b) provider edge (PE) router, which serves as the interface between 5G HetNet and optical core
- (c) provider (P) router, which serves to forward the user traffic arranged by CE and PE routers

3.2. IMS Based IoT Video Streaming. Internet of things (IoT) is regarded recently as the most promising form of the wireless communication environments.

In Internet devices, a large amount of data is generated and accumulated into a large amount of data streams over time. Therefore, higher requirements are imposed on the processing and decomposition of online data. In many applications, a large amount of data storage is generated, such as monitoring the generated video stream data.

5G HetNets transport the video streaming traffic by changing an IoT sensor into an IP Multimedia Subsystem (IMS) terminal. IMS realizes robust IMS service functionality by working closely with Call Session Control Function (CSCF). The CSCF node facilitates Session Initiation Protocol (SIP) to operate session setup and teardown [18].

As a popular video streaming protocol, CSCF conducts signaling and session management and enables connection information to go across network boundaries.

Figure 5 gives an example of IMS based IoT video streaming over 5G HetNets. The improved CSCF server acts as an intermediate entity that receives and forwards signaling messages. Improved server offers a bunch of crucial functions, such as authentication, authorization, and routing.

3.3. Model of Improved CSCF Server. By definition, the main module of the IMS system is call session control to allow video/voice communication over the convergence of different access networks. It comprises of all the functional modules required to handle all the signaling from end-user to services and other networks. However, conventional CSCF cannot

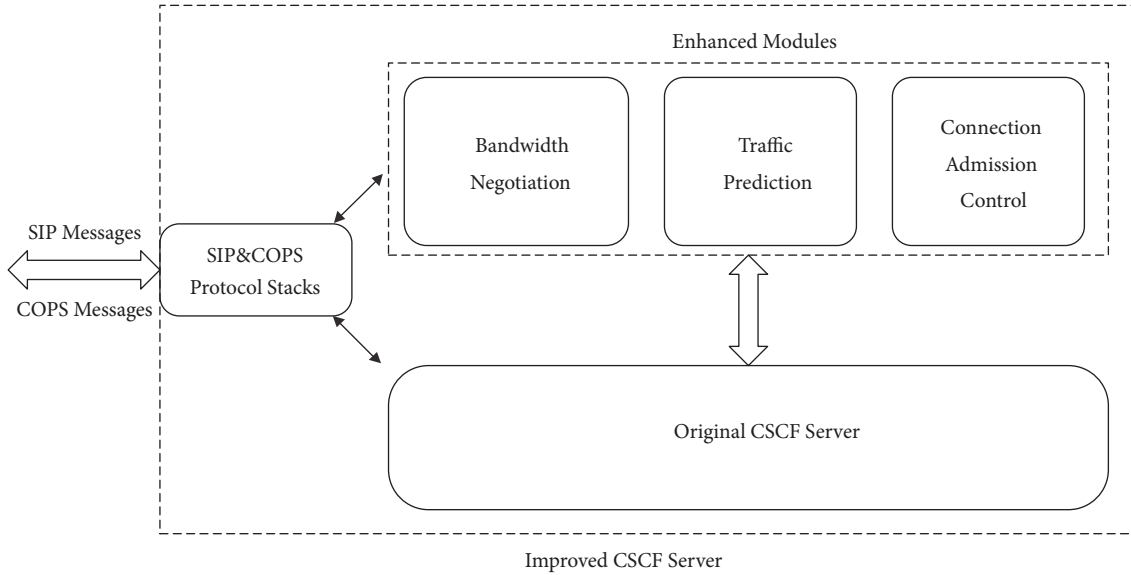


FIGURE 6: Model of improved CSCF server.

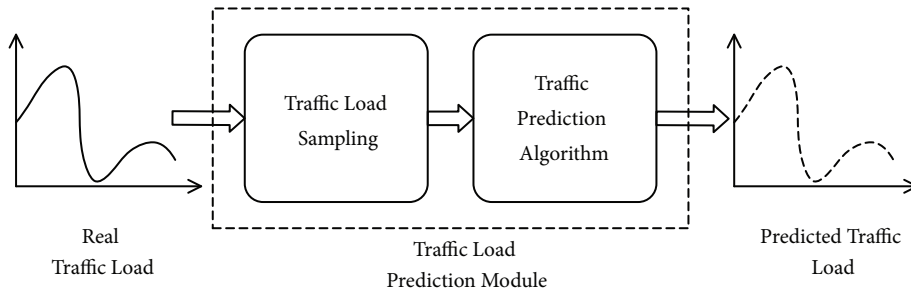


FIGURE 7: General framework of IoT traffic prediction.

100% meet the requirement of IoT video streaming and be modernized.

Figure 6 shows that an improved CSCF server contains three more modules than the original. Our design has the merit that the improved CSCF server bargains with the optical core on behalf of all IoT devices of the local HetNet in advance.

For this reason, the paths required by video streaming are preconstructed before the stream even starts, to reduce the unnecessary signaling delay.

4. Traffic Prediction in IoT Using Compressed Sensing

Traffic prediction in the IoT has been widely studied and applied in recent years. Sampling and prediction represent usually two independent parts in the past. For example, most conventional approaches keep the sampling rate unchanged, which may result in low efficiency in IoT video streaming. To overcome this problem, this paper develops a novel compressed sensing based traffic prediction, where error of forecast algorithm goes back to control the sampling interval. Our proposed scheme can significantly reduce the sampling overhead without sacrificing the accuracy.

4.1. Problem Formulation and Architecture Design. Let $B_{n-1:n}$ be the sum of IoT video streaming loads from a local HetNet at time slot $[t_{n-1} : t_n]$ ($n = 0, 1, 2, \dots$). Let improved CSCF server be exactly aware of the value of $B_{n-1:n}$ at time t_{n-1} . With this piece of information, it will bargain with 5G core for the amount of $B_{n-1:n}$ bandwidth using COPS messages. At the next time slot $[t_n : t_{n+1}]$, IoT traffic may change to a new value of $B_{n:n+1}$, and the improved CSCF server will have to rebargain with 5G core network again.

It is, nevertheless, impossible for the improved CSCF server to achieve the value of $B_{n-1:n}$ before the time of t_n . Thus, at time t_{n-1} , we can estimate the approximate value $B_{n-1:n}$, which is denoted by $B'_{n-1:n}$. If $B'_{n-1:n} < B_{n-1:n}$, the local HetNet will suffer from the shortage of bandwidth resource to accommodate all video streaming sessions. As a solution, CAC component must reject some of connection requests in order to guarantee overall quality of service.

On the other hand, if $B'_{n-1:n} > B_{n-1:n}$, part of bandwidth resource is simply overbooked and useless. In this sense, the traffic prediction is critical to ensure a good performance of our proposed scheme.

Figure 7 presents the general framework of IoT traffic prediction module with the component of traffic load sampling and the component of traffic prediction algorithm.

We will give more elaboration on these two components next.

4.2. IoT Traffic Sampling. An IoT device initiates a video stream session by sending an INVITE message to the improved CSCF server, in a bid to indicate called URL and bandwidth requirement. The improved CSCF server saves every request in the record for IoT traffic sampling usage, regardless the request is accepted by CAC or not.

In this paper, we use $B_{n-1:n}$ to represent the average IoT traffic interval $\Delta t_{n-1:n} = t_n - t_{n-1}$. In practice, the system achieves $B_{n-1:n}$ from the record in the improved CSCF server. Conventional approach conducts traffic sampling with constant rate, which rules

$$\Delta t_{0,1} = \Delta t_{1,2} = \dots = \Delta t_{n-1,n} = \dots, \quad (n = 0, 1, 2, \dots) \quad (1)$$

Let $B(t)$ be the traffic load function varying with time t , and let $B(f)$ be the frequency counterpart of $B(t)$ with the upper bound of f_{\max} . Nyquist theorem regulates the idea that sampling rate must be at least twice of f_{\max} to prevent aliasing [19]. In IoT video streaming, f_{\max} could be high in some complicated scenarios, which make constant sampling rate not applicable at all.

4.3. Compressed Sensing. Compressed sensing theory is a theory that has been widely used in recent years and is widely used in signal processing. It is also called compressed sampling or sparse sampling. CS wants to sample the signal at a sampling rate much lower than Shannon's Law and find a solution for underdetermined linear systems while maintaining the basic information of the signal. [20–24].

In order to implement the compressed sensing theory, the following solution can be used to estimate the sparse vector $x \in \mathbb{C}^N$ using the observed measurement vector $y \in \mathbb{C}^M$. A method based on measurement equations is applied. The formula is as follows:

$$y = \Phi x + w \quad (2)$$

The measurement matrix is represented by $\Phi \in \mathbb{C}^{M \times N}$. $w \in \mathbb{C}^M$ represents the unknown vector of measurement noise and modeling error. The reconstruction process can only occur under ideal conditions, i.e., only when the signal x is S sparse ($S \ll N$). That is, there can be at most S nonzero items. During the operation, the number of observations is significantly smaller than the number of variables, i.e., $\gg M$.

In practice, the signal we are dealing with is not a sparse signal. The processing method in this time is to express $\{\Psi\}_{i=1}^N$ on some basis with the corresponding sparse coefficient θ_i . Processing equation (1) yields

$$y = \Phi x + w = \Phi \Psi \theta + w = \Theta \theta + w \quad (3)$$

where $\theta \in \mathbb{C}^L$ is a L dimension sparse vector of coefficients based on the basis matrix $\Psi \in \mathbb{C}^{N \times L}$. $\Theta = \Phi \Psi$ is a $M \times L$ matrix ($M \ll L$).

However, because fewer measurements are applied than the entries in θ , the resulting solution is uncertain. In

order to recover the S sparse signal, l_0 norm of the most sparse sequence should be found from all feasible solutions. Minimization can be used for the reconstruction process. The formula is expressed as follows:

$$\begin{aligned} \hat{\theta} &= \min \quad \|\theta\|_0 \\ \text{s.t.} \quad &\|\Theta \theta - y\|_2 \leq \Xi \end{aligned} \quad (4)$$

where $\hat{\theta}$ refers to the estimated vector for θ and $\|w\| \leq \Xi$ is noise tolerance. However, for practical operations, the l_0 norm minimization has huge computational complexity, and a reconstruction algorithm with lower computational cost must be developed.

Studies have shown that the reconstruction process of signal x must satisfy a condition that matrix Θ cannot map two different S sparse signals to the same set of samples. Therefore, the matrix Θ must satisfy the restricted equidistance property (RIP) [25, 26].

4.4. Adaptive Sampling Rate. (a) The definition of traditional compressive sensing: sparse signals can be recovered from a small number of nonadaptive linear measurements. Let signal x be K sparse in basis/dictionary Ψ . For example, Ψ is DFT matrix, if the signal is frequency domain sparse. Ψ can be expressed as a matrix as shown in Figure 8(a). In Figure 9, if signal x is sparse, then we can use compressive sensing by $y = \Phi \times x$.

Φ is the measurement matrix and the theory of CS states that random Φ will work. How to understand this? For example, if the signal is sparse in frequency domain, then we make a few random samplings in time domain. In contrast, if the signal is sparse in time domain, we make a few random samplings in frequency domain.

(b) Address the situation that the appearance time of sparse signal can be predicted. In this case, Ψ is identity matrix I as shown in Figure 8(b). Then, measure matrix Φ is part of I , which involves the time points where the sparse signal appears. Φ is not random anymore. In practice, we propose variable sampling rate scheme to predict these time points (or the measurement matrix).

If a fixed sampling interval applies, as mentioned before, the sampling rate must double the highest frequency of traffic curve to achieve accurate prediction.

Accordingly, it is significant to develop an inconstant sampling rate scheme with much lower overhead.

Variable sampling rate comes from the observation of the traffic curve in Figure 10, with combination of slow and fast changing periods. The fluctuation of traffic significantly depends on the timely feature of IoT devices. For example, the cameras monitor the parking cars' moving and stopping according to people's behavior, which is obviously related to the number of car arrivals and departures. In Figure 8, fast changing zone represents the period of rapid event varying, such as morning peak hour; slow changing zone represents the period of steady event numbers, such as at night.

With the goal of reducing the sampling overhead, a straightforward adaptive rate approach is to utilize low rate in slow changing period and high rate in fast changing

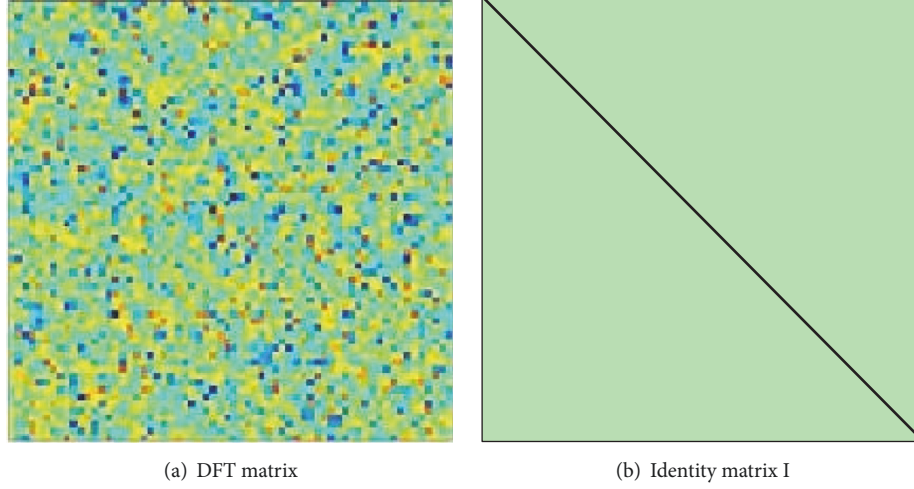


FIGURE 8: The matrix involved in the compressive sensing.

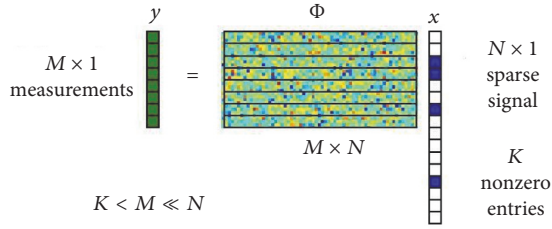


FIGURE 9: The process of compressive sensing.

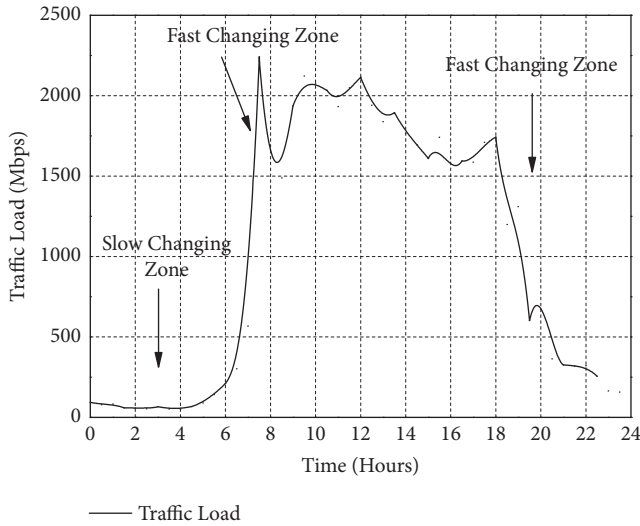


FIGURE 10: IoT traffic load curve of a parking monitor camera.

period. The adaptive sampling interval aims to sample and reconstruct the traffic curve efficiently and effectively, while keeping the same prediction accuracy as constant rate approach does.

Next, we study a typical scenario as follows. Let's divide an IoT traffic curve into a number of slow changing and fast changing periods by using f_{\max}^s and f_{\max}^f to represent the maximum frequencies, respectively. Nyquist sampling theory states that the sampling rate must be greater than $2f_{\max}^s$ in case of slow changing zone and $2f_{\max}^f$ in case of fast changing. After that, we can achieve the average rate of sampling R_{VSR}^{avg} by

$$R_{VSR}^{avg} = \frac{2f_{\max}^s T_s + 2f_{\max}^f T_f}{T_s + T_f} \quad (5)$$

where T_s and T_f are time of slow and fast changing. Since $f_{\max}^s < f_{\max}^f < f_{\max}$, we have $R_{VSR}^{avg} < 2f_{\max}$, implying that ASR approach has much less sampling overhead than its constant counterpart.

4.5. Traditional Linear Predictor. The traditional predictor used most often in practice is the least mean square (LMS) linear filter. A k -step LP makes a guess of the value of $x(n+k)$ by a linear sum of the current and past $x(n)$. Mathematically, the p th-order k -step is given by

$$\hat{x}(n+k) = \sum_{i=0}^{p-1} w_i x(n-i) = w^T x(n) \quad (6)$$

where $\hat{x}(n+k)$ is the estimated value of $x(n+k)$, $w = [w_0, w_1, \dots, w_{p-1}]^T$ is the weights, and $x(n) = [x(n), x(n-1), \dots, x(n-p+1)]^T$ is priori knowledge. We further define the prediction error as

$$e(n) = x(n+k) - \hat{x}(n+k) = x(n+k) - w^T x(n) \quad (7)$$

Then, LP updates $w(n)$ using the following recursive equation:

$$w(n+1) = w(n) + \frac{\mu e(n) x(n)}{\|x(n)\|^2} \quad (8)$$

where μ is the step size that may keep constant (constant step size (CSS)) or adaptive (adaptive step size (ASS)).

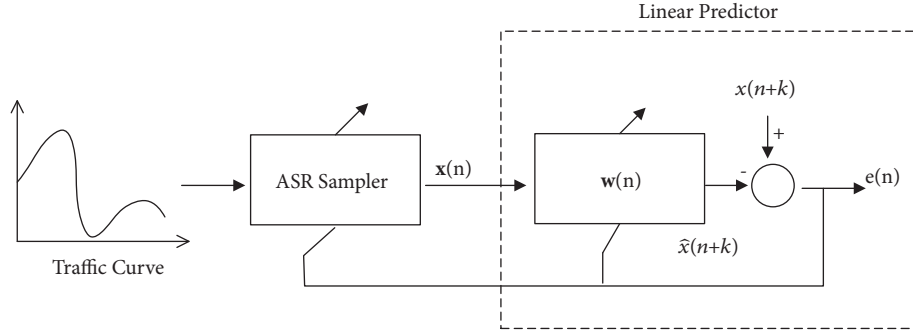


FIGURE 11: ASR-LP working with variable sampling rate.

4.6. Adaptive Sampling Rate Linear Predictor. Realistically the traffic curve is unknown at the time of prediction, and thus it is a mission impossible to chop it into slow and fast changing periods. We, therefore, develop an adaptive prediction scheme ASR-LP, as shown in Figure 11.

ASR-LP rules the sampling rate at time t_n as $R_s(n) = 1/\Delta t_n - 1$; and then update it in integration by

$$R_s'(n+1) = R_s(n) \left[q + (1-q) \frac{|e(n)|}{E_b} \right] \quad (9)$$

with $0 < q < 1$, $E_b > 0$, and

$$R_s(n+1) = \begin{cases} R_s^{\max}, & \text{if } R_s'(n+1) > R_s^{\max} \\ R_s^{\min}, & \text{if } R_s'(n+1) < R_s^{\min} \\ R_s'(n+1), & \text{otherwise} \end{cases} \quad (10)$$

where $0 < q < 1$ is the remembering factor, E_b is the targeted prediction error bound, and $R_{s_{\min}}$ and $R_{s_{\max}}$ are the upper and lower bounds of sampling rate.

The scheme of ASR-LP optimizes sampling interval by the following principal. The larger the prediction error is, the sharper the traffic curve changes. As a result, we must increase the sampling rate for satisfying prediction accuracy. In contrast, the smaller the prediction error is, the lower the sampling rate goes. Particularly, network operator can determine the value of targeted prediction error bound E_b . If $|e(n)| > E_b$, the sampling rate should be increased; if $|e(n)| < E_b$, the sampling rate should be decreased; if $|e(n)| = E_b$, the sampling rate should be kept unchanged.

In addition to E_b , there are still three other parameters in (9) and (10); i.e., $R_{s_{\max}}$ is the upper bound of sampling rate, which can lead to lower prediction error but larger computational complexity as the value rises. $R_{s_{\min}}$ is the lower bound of sampling rate, which represents the bottom line responding time of a linear predictor. Last but not least, q decides how much the current sampling rate is relying on its previous value rather than the prediction error. We have to leverage the value of remembering factor q for the optimal performance of predictor, with $q = 70\%$ as a typical case.

In our research, we have conducted simulation studies on a variety of traffic curves to evaluate the performance

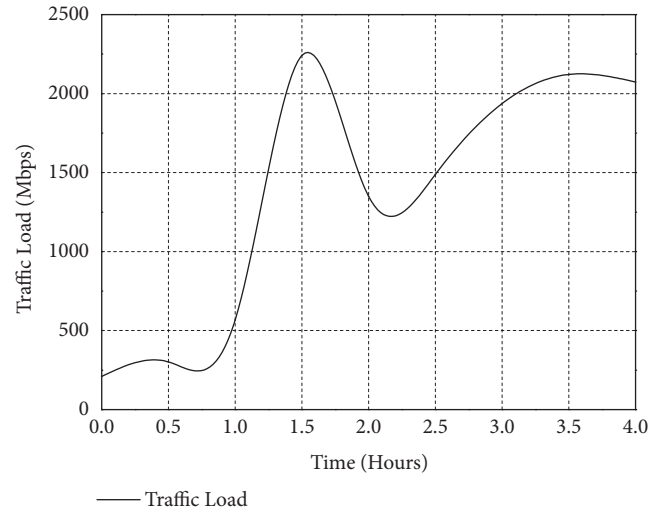


FIGURE 12: IoT Traffic load of four hours for parking monitor camera.

of our proposed ASR-LP scheme. Due to the limit of space, we demonstrate only one typical curve in Figure 10 as an example. However, the simulation results from this example are also largely true to the general cases.

Figure 10 is collected from Melbourne on street car parking sensor data [26], which reflects the typical timing feature of IoT users. It is worth noting that the curve of IoT traffic is dramatically different from previously investigated mobility models in existing literature, such as Brownian motion model.

We use Figure 12 as prototype traffic curve to compare the performance of three schemes as shown in Figure 12. In addition to its advantage in the running time, VSR-NLMS does have another important virtue, i.e., a lower average sampling rate compared to FSS-NLMS and VSS-NLMS. It shows that, with the same simulation configuration as shown in Figure 12, VSR-NLMS achieves an average sampling rate of 1/371 Hz, which is lower than 1/420 Hz for FSS-NLMS and VSS-NLMS. Clearly, VSR-NLMS has the lowest sampling rate and thus the least sampling overhead.

Figure 13 reveals that ASR-LP can basically tie ASS-LP, but far outperform CSS-LP regarding prediction accuracy.

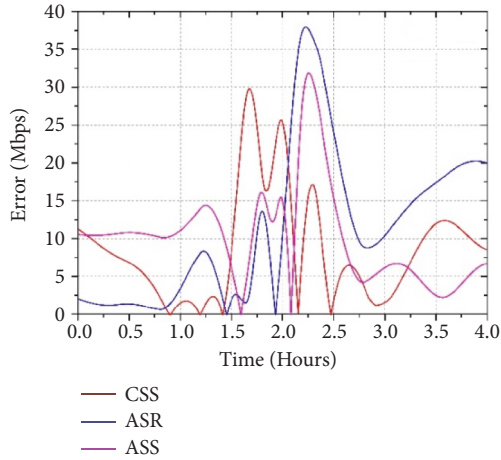


FIGURE 13: Errors of different predicting methods in four hours.

ASR-LP : $E_b = 100\text{Mbps}$, $R_{s_{\max}} = 1 = 300\text{Hz}$, $R_{s_{\min}} = 1/800\text{Hz}$, $R_s(0) = 1/420\text{Hz}$, and $q = 70\%$.

CSS-LP : $p = 4$, $\mu = 0.1$, ASS-LP : $\mu_{\max} = 0.25$, $\mu_{\min} = 0.05$.

Low overhead of average sampling is a significant advantage of ASR-LP, compared to its counterpart. Intensive simulation study reveals that ASR-LP can save at least half sampling rate over constant sampling rate linear predictors with the same accuracy.

5. Conclusions

This paper studies the deployment of IoT video streams on 5G HetNets and proposes an improved CSCF server solution to promote IoT traffic prediction and resource reservation. We further designed a linear predictor based on compressed sensing to capture traffic patterns, greatly reducing sampling overhead and computational complexity. Intensive analysis and simulation results show that the proposed scheme can effectively improve performance and save time and cost.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] M. Li, Y. Du, X. Ma, and S. Huang, "A 3-param network selection algorithm in heterogeneous wireless networks," in *Proceedings of the IEEE 9th International Conference on Communication Software and Networks (ICCSN '17)*, pp. 392–396, IEEE, Guangzhou, China, May 2017.
- [2] G. Ma, Y. Yang, X. Qiu, Z. Gao, and H. Li, "Fault-tolerant topology control for heterogeneous wireless sensor networks using Multi-Routing Tree," in *Proceedings of the 15th IFIP/IEEE International Symposium on Integrated Network and Service Management (IM '17)*, pp. 620–623, Lisbon, Portugal, May 2017.
- [3] S. Han, Y. Huang, W. Meng, C. Li, N. Xu, and D. Chen, "Optimal power allocation for SCMA downlink systems based on maximum capacity," *IEEE Transactions on Communications*, vol. 67, no. 2, pp. 1480–1489, 2019.
- [4] A. Montazerolghaem, M. H. Moghaddam, and A. Leon-Garcia, "OpenSIP: toward software-defined SIP networking," *IEEE Transactions on Network and Service Management*, vol. 15, no. 1, pp. 184–199, 2018.
- [5] A. K. Vimal, S. Pandit, A. K. Godiyal, S. Anand, S. Luthra, and D. Joshi, "An instrumented flexible insole for wireless COP monitoring," in *Proceedings of the 8th International Conference on Computing, Communications and Networking Technologies (ICCCNT '17)*, pp. 1–5, Delhi, India, July 2017.
- [6] A. Mirrashid and A. A. Beheshti, "Compressed remote sensing by using deep learning," in *Proceedings of the 9th International Symposium on Telecommunications (IST '18)*, pp. 549–552, IEEE, Tehran, Iran, December 2018.
- [7] S. Rouabah, M. Ouarzeddine, and B. Souissi, "SAR images compressed sensing based on recovery algorithms," in *Proceedings of the 2018 IEEE International Geoscience and Remote Sensing Symposium (IGARSS '18)*, pp. 8897–8900, IEEE, Valencia, Spain, July 2018.
- [8] S. Han, Y. Zhang, W. Meng, and H. Chen, "Self-interference-cancellation-based SLNR precoding design for full-duplex relay-assisted system," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 9, pp. 8249–8262, 2018.
- [9] S. Han, S. Xu, W. Meng, and C. Li, "Dense-device-enabled cooperative networks for efficient and secure transmission," *IEEE Network*, vol. 32, no. 2, pp. 100–106, 2018.
- [10] D. Ghadiyaram, J. Pan, and A. C. Bovik, "A subjective and objective study of stalling events in mobile streaming videos," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 1, pp. 183–197, 2019.
- [11] Z. Zheng, L. Pan, and K. Pholsena, "Mode decomposition based hybrid model for traffic flow prediction," in *Proceedings of the 3rd IEEE International Conference on Data Science in Cyberspace (DSC '18)*, pp. 521–526, IEEE, Guangzhou, China, June 2018.
- [12] S. Fan and H. Zhao, "Delay-based cross-layer QoS scheme for video streaming in wireless ad hoc networks," *China Communications*, vol. 15, no. 9, pp. 215–234, 2018.
- [13] T. Zhang, L. Feng, P. Yu, S. Guo, W. Li, and X. Qiu, "A handover statistics based approach for cell outage detection in self-organized heterogeneous networks," in *Proceedings of the 15th IFIP/IEEE International Symposium on Integrated Network and Service Management (IM '17)*, pp. 628–631, IEEE, Lisbon, Portugal, May 2017.
- [14] S. Sun, M. Kadoch, L. Gong, and B. Rong, "Integrating network function virtualization with SDR and SDN for 4G/5G networks," *IEEE Network*, vol. 29, no. 3, pp. 54–59, 2015.
- [15] M. Lu, Z. Qu, Z. Wang, and Z. Zhang, "Hete_MESE: multi-dimensional community detection algorithm based on multiplex network extraction and seed expansion for heterogeneous information networks," *IEEE Access*, vol. 6, pp. 73965–73983, 2018.
- [16] Z. Zhang, L. Song, Z. Han, and W. Saad, "Coalitional games with overlapping coalitions for interference management in small cell networks," *IEEE Transactions on Wireless Communications*, vol. 13, no. 5, pp. 2659–2669, 2014.

- [17] K. Zhang, S. Leng, Y. He, S. Maharjan, and Y. Zhang, "Cooperative content caching in 5G networks with mobile edge computing," *IEEE Wireless Communications Magazine*, vol. 25, no. 3, pp. 80–87, 2018.
- [18] J. Rosenberg, H. Schulzrinne, G. Camarillo et al., "SIP: session initiation protocol," RFC 3261, IETF, June 2002.
- [19] K. Zhang, S. Leng, Y. He, S. Maharjan, and Y. Zhang, "Mobile edge computing and networking for green and low-latency internet of things," *IEEE Communications Magazine*, vol. 56, no. 5, pp. 39–45, 2018.
- [20] G. Shrividya and S. H. Bharathi, "A study of optimum sampling pattern for reconstruction of MR images using compressive sensing," in *Proceedings of the Second International Conference on Advances in Electronics, Computers and Communications (ICAEECC '18)*, pp. 1–6, Bangalore, India, February 2018.
- [21] Y. Wu, B. Rong, K. Salehian, and G. Gagnon, "Cloud transmission: a new spectrum-reuse friendly digital terrestrial broadcasting transmission system," *IEEE Transactions on Broadcasting*, vol. 58, no. 3, pp. 329–337, 2012.
- [22] A. Sammoud, A. Kumar, M. Bayoumi, and T. Elarabi, "Real-time streaming challenges in Internet of Video Things (IoVT)," in *Proceedings of the 50th IEEE International Symposium on Circuits and Systems (ISCAS '17)*, pp. 1–4, Baltimore, Md, USA, May 2017.
- [23] N. Eswara, K. Manasa, A. Kommineni et al., "A continuous QoE evaluation framework for video streaming over HTTP," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 11, pp. 3236–3250, 2018.
- [24] B. Rong, Y. Qian, K. Lu, H.-H. Chen, and M. Guizani, "Call admission control optimization in WiMAX networks," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 4, pp. 2509–2522, 2008.
- [25] I. D. Irawati, A. B. Suksmono, and I. Y. M. Edward, "Local interpolated compressive sampling for internet traffic reconstruction," in *Proceedings of the International Conference on Advanced Computing and Applications (ACOMP '17)*, pp. 93–98, Ho Chi Minh City, Vietnam, December 2017.
- [26] N. Anselmi, L. Poli, G. Oliveri, and A. Massa, "Three dimensional imaging with the contrast source compressive sampling," in *Proceedings of the International Applied Computational Electromagnetics Society Symposium - China (ACES '18)*, pp. 1–2, IEEE, Beijing, China, July 2018.

Research Article

A Hybrid Predictive Strategy Carried through Simultaneously from Decision Space and Objective Space for Evolutionary Dynamic Multiobjective Optimization

Peng Xu,¹ Xiaoming Wu,^{2,3,4} Man Guo,⁵ Shuai Wang,^{2,6,7} Qingya Li ^{2,3,4} and Weiping Huang¹

¹School of Information Science and Engineering, Shandong University, Qingdao, Shandong Province 266237, China

²Qilu University of Technology (Shandong Academy of Sciences), Jinan, Shandong Province, China

³Shandong Computer Science Center (National Supercomputer Center in Jinan), Jinan, Shandong Province, China

⁴Shandong Provincial Key Laboratory of Computer Networks, Jinan, Shandong Province, China

⁵TORCH High Technology Industry Development Center, Ministry of Science and Technology 18A, Section 2 of Sanlihe Community, Xicheng District, Beijing 100045, China

⁶Heze Branch of Shandong Academy of Sciences, Heze, Shandong Province 274009, China

⁷Biological Engineering Technology Innovation Center of Shandong Province, Heze, Shandong Province 274009, China

Correspondence should be addressed to Qingya Li; liqy@sdas.org

Received 1 April 2019; Accepted 30 May 2019; Published 23 June 2019

Guest Editor: Shuai Han

Copyright © 2019 Peng Xu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

There are many issues to consider when integrating 5G networks and the Internet of things to build a future smart city, such as how to schedule resources and how to reduce costs. This has a lot to do with dynamic multiobjective optimization. In order to deal with this kind of problem, it is necessary to design a good processing strategy. Evolutionary algorithm can handle this problem well. The prediction in the dynamic environment has been the very challenging work. In the previous literature, the location and distribution of PF or PS are mostly predicted by the center point. The center point generally refers to the center point of the population in the decision space. However, the center point of the decision space cannot meet the needs of various problems. In fact, there are many points with special meanings in objective space, such as ideal point and CTI. In this paper, a hybrid prediction strategy carried through from both decision space and objective space (DOPS) is proposed to handle all kinds of optimization problems. The prediction in decision space is based on the center point. And the prediction in objective space is based on CTI. In addition, for handling the problems with periodic changes, a kind of memory method is added. Finally, to compensate for the inaccuracy of the prediction in particularly complex problems, a self-adaptive diversity maintenance method is adopted. The proposed strategy was compared with other four state-of-the-art strategies on 13 classic dynamic multiobjective optimization problems (DMOPs). The experimental results show that DOPS is effective in dynamic multiobjective optimization.

1. Introduction

In real life, there are many optimization problems which have multiple objectives but these objectives conflict with each other. These optimization problems are called multiobjective optimization problems (MOPs). However, many MOPs in the real world always contain uncertain and dynamic factors. For example, air traffic scheduling is easily affected by weather and some emergencies. So in that case the best solutions are hard to keep valid for a long time. If the objectives,

constraints, or parameters of MOPs change with time, this kind of MOPs is called dynamic MOPs (DMOPs). Due to the dynamic character, DMOPs are harder to converge than MOPs. The mathematical formula is as follows:

$$\begin{aligned} \min_{x \in \Omega} \quad & F(x, t) = (f_1(x, t), f_2(x, t), \dots, f_m(x, t))^T \\ \text{s.t.} \quad & g_i(x, t) \leq 0, \quad i = 1, 2, \dots, p; \\ & h_j(x, t) = 0, \quad j = 1, 2, \dots, q \end{aligned} \quad (1)$$

where $x = (x_1, x_2, \dots, x_n)$ is the n -dimensional decision vector whose domain of definition is Ω . t represents time variable. $F = (f_1, f_2, \dots, f_m)$ is the m -dimensional objective vector. g represents p -dimensional inequality constraints, and h is q -dimensional equality constraints. The optimal tradeoff solution set is called Pareto set (PS) in decision space and Pareto front (PF) in the objective space.

In the recent literature, whether in industrial applications or scientific research, there is a lot of contents related to DMOPs [1–3]. Industrial applications involve design [4, 5], management [6, 7], scheduling [8–10], planning [11–14], and control [15, 16]. Scientific research includes constrained optimization [17, 18], machine learning [19, 20], and bilevel optimization [21]. In some fields [22–24], we will also try to use evolutionary algorithm to solve.

The integration of 5G network and Internet of things to build a future smart city will involve a lot of sensor installations and configurations as well as the priority processing of different tasks. In addition, a good resource allocation plan will greatly promote the construction of smart cities [25, 26]. These things have a lot to do with DMOPs. In the literature [27], a strategy based on centroid-based adaptation (CBA) is put forward to solve the problem of mission plan. In this case, two objectives need to be optimized: operation execution time and operation cost. In addition, there are time-varying constraints, including changes in execution times and task-related networks. The results show that CBA is effective in dealing with mission plan.

While handling DMOPs, the traditional static algorithms [28, 29] have not been suited. Because when using these methods, the population often converges to the optimum solutions in the current environment. And then the genes of the population will become single and lose diversity. When the environment varies, it is very difficult to find the optimal solutions. So, some researchers adjusted these static algorithms to solve DMOPs [8, 30].

Afterwards, other strategies such as the random initialization methods [24, 31–33], the memory strategies [34, 35], and prediction strategies [36–40] are introduced. The memory strategy can respond to the environmental changes by recording historical optimal solutions to converge fast. However, the memory strategies have the blind character dealing with DMOPs. When dealing with periodic problems, they can achieve good results. But the effect is bad for nonperiodic problems. The prediction strategy can take advantage of historical information to predict the optimal solutions of the new environment. The accuracy of prediction has been an important aspect on research. Hatzakis and Wallace proposed the feed-forward prediction strategy (FPS) [36] in 2009. In this strategy, the historical boundary points are used to predict the new boundary points in the new environment. By the new boundary points, the PF of the new environment is located. FPS has certain effect dealing with DMOPs, but only boundary points used are hard to trace PF or PS accurately. In 2013, Wu et al. proposed a predictive multiobjective genetic algorithm (PMGA) [39]. In this method, PF is clustered into multiple center centers to guide the rapid evolution of the population toward the optimal solutions. It is good for those problems whose decision variables have linear relationship

between each other. However, for those problems whose decision variables have nonlinear relationship between each other, the effect is bad. And then, Zhou and Jin et al. proposed a population prediction strategy (PPS) [37]. PPS can not only predict the center point of the population of the next environment in the decision space, but also predict the shape distribution of the nondominated set. It has good effect for DMOPs, but it also has a drawback. In the front learning stage, the effect is bad due to lack of experience. In 2015, Wu and Jin et al. proposed a direct prediction strategy [38], which directly uses feed-forward center point to predict population optimal solutions of the next environment. It can achieve good results for DMOPs. However, it can introduce some useless individuals except nondominated individuals. And, in 2017, Jiang and Yang et al. proposed an algorithm based on steady state and generation [41], which can handle DMOPs well.

Many researchers try to find points or regions of special significance [42–44]. There are many special points that have been found, not only center points, such as boundary points and knee points [45–50]. In this paper, a hybrid prediction strategy based on center point and CTI, which can handle DMOPs from both decision space and objective space, is proposed. In the previous literature, the evolution direction of the population and the position and shape distribution of PF are judged according to the decision space in most cases. However, in this paper, we judge the evolution direction of population and distribution of optimal solutions by CTI in objective space. And then we combine it with the method based on center point to deal with DMOPs together. Then, memory strategy and adaptive diversity maintenance strategy are introduced to make up for the bad effect caused by inaccurate prediction in particularly complex problems.

The rest of this article is structured as follows. Section 2 introduces some related work. Section 3 describes DOPS in detail. Section 4 gives test problems and performance indicator. Section 5 gives experimental results and analysis. Section 6 does more discussion to further analyze the advantages and disadvantages of the prediction method based on the center point and that based on CTI. And, in the end, Section 7 gives conclusion and future work.

2. Related Work

2.1. A General Framework of Dynamic Multiobjective Genetic Algorithm. The general framework of dynamic multiobjective evolutionary algorithm is as follows.

- (1) Initialize the population and set the relevant parameters.
- (2) Detect the environmental change. When detecting the change, go to step (4); else, go to step (3).
- (3) Optimization algorithm is used to handle the problem.
- (4) Some response strategies are used, such as reinitialization, memory strategy, and prediction strategy, to respond to the environmental change.
- (5) Judge the termination condition. If not terminated, go to step (2); if terminated, then exit.

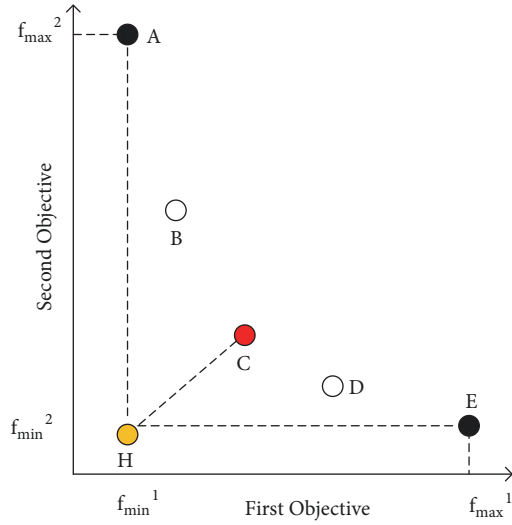


FIGURE 1: CTI schematic.

2.2. *CTI*. Many researchers have tried to find some points or regions with special meanings [36, 42–44]. Among them, ideal point and a particular point CTI [36, 45] extending from an ideal point have aroused great interest among the researchers.

The ideal point is a point composed of the smallest values of each dimension in the nondominated set. The mathematical definition is as follows:

$$H_i^* = \min \{f(\text{Non}P_i^1), f(\text{Non}P_i^2), \dots, f(\text{Non}P_i^M)\} \quad (2)$$

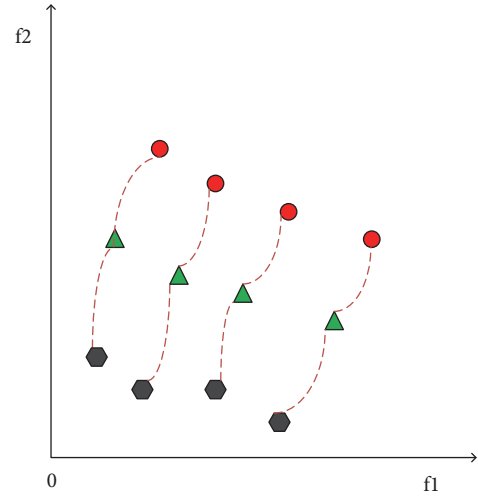
where $i = 1, 2, \dots, m$. m is the dimension of the objective space. M is the size of the nondominated set. $f(\text{Non}P_i^M)$ denotes the value of the i -th dimension of the M -th individual in the nondominated set. As shown in Figure 1, A, B, C, D, and E are five individuals in nondominated set. Among them, A and E are boundary points. Thus, H is the ideal point.

In [36, 45], CTI is as the representative point to predict the location of PF. CTI is close-to-ideal point, which means that it is the closest individual to ideal point in nondominated set. As shown in Figure 1, C is CTI. C is the point whose Euclidean distance is shortest to ideal point H.

2.3. *The Principle of the Prediction Model*. In [34], the principle of the prediction model is described in detail. In the section, the principle is described in short. Supposing that $x_1, x_2, \dots, x_t \in Q_i$, $i = 1, \dots, t$, are a series of points, these points are n -dimensional decision vectors in the decision space which are used to describe the movement of PS. The general prediction model is defined as follows:

$$x_{t+1} = F(x_t, x_{t-1}, \dots, x_{t-k+1}, t) \quad (3)$$

where K is the number of previous time steps. X_{t+1} is the individual in $t+1$ time step, which is also the individual to



Q_{t-1}
 Q_t
 Predicted points
 Prediction model

FIGURE 2: Prediction model.

predict. Figure 2 is what happens when k is equal to 2. In this paper, a simple and widely used prediction model is adopted.

$$x_{t+1} = x_t + (x_t - x_{t-1}) + \text{Gauss}(0, d) \quad (4)$$

where the direction and distance of prediction are got by $x_t - x_{t-1}$. $\text{Gauss}(0, d)$ is a Gaussian perturbation added in order to avoid falling into a local optimum. d is the variance of disturbance.

3. A Hybrid Predictive Strategy Carried through Simultaneously from Decision Space and Objective Space

Prediction strategy has been used to predict moving PS or PF. In real world, there are all kinds of DMOPs. The change locations of DMOPs may be in PS or PF or both PS and PF [37, 52, 53]. Thus, to handle different kinds of DMOPs, a hybrid predictive strategy carried through simultaneously from decision space and objective space (DOPS) is proposed. The predictive strategy in the decision space is based on center point. And the predictive strategy in the objective space is based on CTI. Two kinds of predictive strategies can both predict the nondominated set. In the paper, the two predicted nondominated sets were halved and then merged into a new nondominated set. Then, several nondominated individuals in the last environment were introduced to the predicted population as the memory set. In the end, we adopted a kind of adaptive diversity maintenance strategy, which is similar to that in [46], to keep diversity. The diversity individuals produced by this diversity strategy are called random set. As shown in Figure 3, all predicted population consists of three parts.

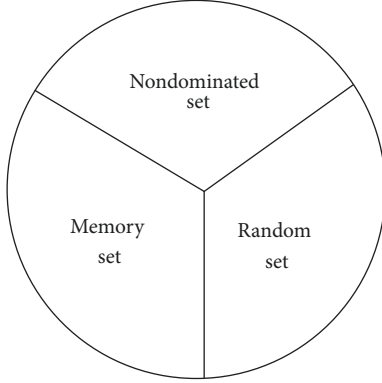


FIGURE 3: Population structure schematic.

3.1. Memory Set. In real world, there are all kinds of dynamic real problems. For some dynamic problems, the optimal solutions may appear periodically. The part of optimal solutions of the population can be recorded. If the same environment occurs in the future, then the optimal solutions recorded are also valid. In this section, a kind of simple memory strategy is selected. If the number of the optimal solutions to be memorized is $Nmem$, $Nmem$ individuals are randomly selected from the optimal individuals in the current population to be memorized, and the $Nmem$ individuals make up the memory set.

3.2. Nondominated Set. The predicted nondominated set is got by hybrid prediction strategy. It contains two prediction substrategies: one is the strategy based on the center point of decision space, and the other one is based on the CTI in objective space.

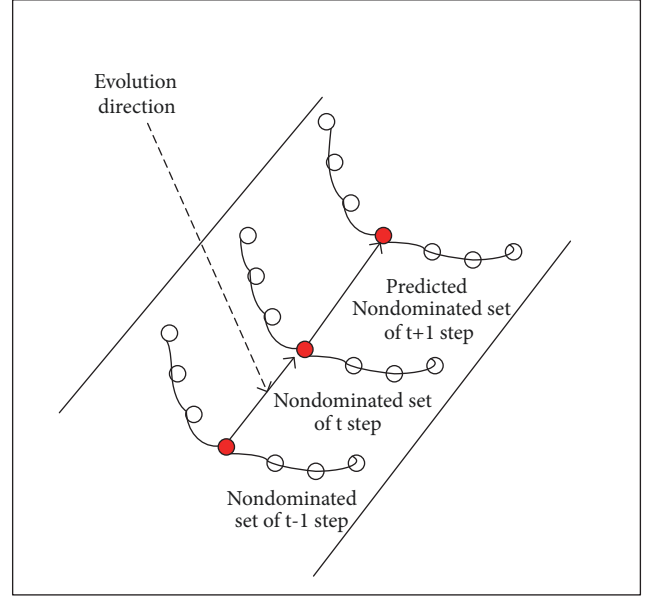
3.2.1. The Prediction Strategy Based on Center Point. The prediction strategy based on center point is a widely used strategy [37, 38, 40, 41, 45, 46, 54, 55]. The center point is defined as follows:

$$C_k^t = \frac{1}{|P_{Non-dom}^t|} \sum_{NonInd_k^t \in P_{Non-dom}^t} NonInd_k^t. \quad (5)$$

where $|P_{Non-dom}^t|$ represents the cardinality of the nondominated set. $P_{Non-dom}^t$ represents the nondominated set in t time step. $NonInd_k^t$ represents a nondominated individual in t step and in the k -th dimension. Based on the prediction model in Section 2.3, the prediction model based on center point can be expressed as

$$pop_k^{t+1} = pop_k^t + (C_k^t - C_k^{t-1}) + Gauss(0, d). \quad (6)$$

where C_k^t and C_k^{t-1} show the center points in k dimension in t and $t-1$ time step, respectively. pop_k^t denotes the individuals of the nondominated set in k dimension and t time step. And pop_k^{t+1} denotes the predicted individuals in k dimension and in $t+1$ time step. As shown in Figure 4, the nondominated set of $t+1$ step can be predicted by using the nondominated set of t plus evolution direction got by center points of t and $t-1$ time step. The detailed description is in Algorithm 1.



- Individual in nondominated set
- Center point in nondominated set

FIGURE 4: Center point prediction schematic.

3.2.2. The Prediction Strategy Based on CTI. CTI is a special point of nondominated sets in the objective space and has been widely used. However, in most cases, CTI is used to be a representative point to predict PF and to judge the position of PF [36, 45]. In this paper, we use CTI not only to judge the evolution direction of the population, but also to judge the position and shape of PF. This strategy is based on the prediction model in Section 2.3. The mathematical form is similar to formula (6), as follows:

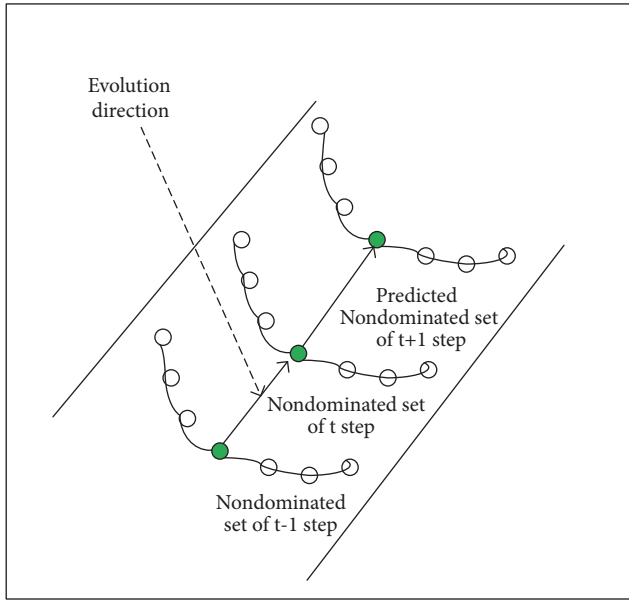
$$pop_k^{t+1} = pop_k^t + (CTI_k^t - CTI_k^{t-1}) + Gauss(0, d) \quad (7)$$

where pop_k^t and $Gauss(0, d)$ are the same as formula (6). CTI_k^t and CTI_k^{t-1} denote CTI in k dimension in t and $t-1$ time step, respectively. As shown in Figure 5, the nondominated set of $t+1$ step can be predicted by using the nondominated set of t plus evolution direction got by CTIs of t and $t-1$ time step. The detailed description is in Algorithm 2.

3.2.3. The Hybrid Prediction Strategy. The hybrid prediction strategy is a mixture of the prediction strategy based on the center point and the prediction strategy based on CTI. In theory, the final nondominated set, which consists of two nondominated sets got by two prediction strategies, respectively, is formed by cutting them in half and then merging them. But, in fact, it is related to whether the capacity of the nondominated set of the current population is odd or even, and whether $Nmem + Nnondom_t$ is greater than the population capacity $psize$. In this case, it is a little more complicated. The detailed process is shown in Algorithms 3 and 4.

Input: Pop_{t-1} and Pop_t , the population in t-1 and t time step.
Output: $Non-dom1$, the predicted non-dominated set in t+1 time step.
Step 1: Layer the populations, Pop_{t-1} and Pop_t , in t-1 and t time step, and then select the individuals in the first layer as non-dominated set, $Non-dom_{t-1}$ and $Non-dom_t$;
Step 2: According to formula (5), calculate the center points of $Non-dom_{t-1}$ and $Non-dom_t$, C_k^{t-1} and C_k^t , respectively.
Step 3: According to formula (6), predict the non-dominated set in the t+1 time step, $Non-dom1$.

ALGORITHM 1: Calculating the nondominated set by the prediction strategy based on center point.



- Individual in nondominated set
- CTI in nondominated set

FIGURE 5: CTI prediction schematic.

3.3. Random Set. Random set is derived from the adaptive population diversity maintenance strategy. In dynamic optimization, random set plays a very important role. It is based on two reasons.

(1) In the environment, only the fittest individuals can survive. The genes of these individuals are suitable for the environment. Thus, in the environment, the genes of the individuals that survive will gradually become similar, which means that the genes of the population will become single. When the environment changes, these genetically unitary individuals who are used to the previous environment will be easily eliminated. Therefore, we need to increase the diversity in the population. When the environment changes, those individuals whose genes are suitable for the new environment will easily survive and reproduce.

(2) When dealing with dynamic problems with prediction strategies, especially for particularly complex problems, it is easy for prediction strategies to cause inaccurate prediction. Increasing the diversity of the population at this

time will make up for the bad effect caused by inaccurate prediction.

In this paper, an adaptive population diversity maintenance strategy similar to that in [46] was adopted. The strategy adaptively introduced a corresponding number of random individuals according to the difficulty of the problem to be solved. Generally, the more complex the problem is, the greater the possibility of inaccurate prediction will be. Therefore, more random individuals need to be introduced to make up for the bad influence brought by inaccurate prediction to a greater extent. And, at this time in DOPS, there were a big number of random individuals produced. On the contrary, the simpler the problem is, the more accurate the prediction is. At this time, the introduction of more random individuals will bring bad results. At that time in DOPS, the number of random individuals introduced happened to be small, even 0. Algorithm 5 describes the process in detail.

3.4. The Detailed Description of DOPS. DOPS is under the dynamic framework in Section 2.1. When the environment changes, DOPS reinitializes the population to respond quickly to the environmental change. The detailed description of DOPS is shown in Algorithm 6.

4. Test Problems and Performance Indicator

4.1. Test Problems. In [52], DMOPs are generally divided into 4 categories according to the location of changes in the test problems as follows:

- (1) PS changes, but PF does not change.
- (2) PS does not change, but PF changes.
- (3) PS and PF change together.
- (4) PS and PF do not change.

However, the first three are often encountered. These test problems come from three test suites: FDA test suite [52], dMOP test suite [53], and F5-F10 [37]. Among them, FDA4 and F8 have three objectives, but the other problems have two objectives. In addition, the decision variables of FDA and dMOP suites are linearly correlated between each other, while the decision variables of F series are nonlinearly correlated between each other. Therefore, the test problems in F suite are more difficult to optimize than those in FDA and dMOP series. In the F suite, F9 and

Input: Pop_{t-1} and Pop_t , the population in t and $t-1$ time step.
Output: $Non-dom2$, the predicted $t+1$ time step non-dominated set.
Step 1: Layer the population in $t-1$ and t time step, Pop_{t-1} and Pop_t and then select the individuals in the first layer as non-dominated set, $Non-dom_{t-1}$ and $Non-dom_t$;
Step 2: Calculate the boundary points of the non-dominated set in $t-1$ and t time step.
Step 3: Calculate the ideal point H_{t-1} and H_t , according to formula (2).
Step 4: The nearest individuals to H_{t-1} and H_t are identified as CTI_{t-1} and CTI_t in $Non-dom_{t-1}$ and $Non-dom_t$, respectively.
Step 5: According to formula (7), predict the non-dominated set in the $t+1$ time step, $Non-dom2$.

ALGORITHM 2: Calculating the nondominated set by the prediction strategy based on CTI.

Input: $Nmem$, the capacity of memory set; $psize$, the size of the population; $Non-dom1$ and $Non-dom2$, the non-dominated sets obtained by the prediction methods based on the center point and CTI, respectively; $Nnondom_t$, the capacity of the non-dominated set in t time step.
Output: The final non-dominated set in $t+1$ time step, $Non-dom$ and its size, $Nnondom$.

```

1: if  $Nnondom_t + Nmem \leq psize$  then
2:   if  $Nnondom_t \% 2 = 0$  then
3:     Select  $(Nnondom_t/2)$  individuals from  $Non-dom1$ , as  $Non-dom1'$ ;
4:     Select  $(Nnondom_t/2)$  individuals from  $Non-dom2$ , as  $Non-dom2'$ ;
5:
6:      $Non-dom := Non-dom1' + Non-dom2'$ .
7:   end if
8:   if  $Nnondom_t \% 2 = 1$  then
9:     Select  $((Nnondom_t + 1)/2)$  individuals from  $Non-dom1$ , as  $Non-dom1^*$ ;
10:    Select  $((Nnondom_t - 1)/2)$  individuals from  $Non-dom2$ , as  $Non-dom2^*$ ;
11:
12:     $Non-dom := Non-dom1^* + Non-dom2^*$ .
13:   end if
14:    $Nnondom := Nnondom_t$ .
15: end if
16: if  $Nnondom_t + Nmem > psize$  then
17:   if  $(psize - Nmem) \% 2 = 0$  then
18:     Select  $((psize - Nmem)/2)$  individuals from  $Non-dom1$ , as  $Non-dom1'$ ;
19:     Select  $((psize - Nmem)/2)$  individuals from  $Non-dom2$ , as  $Non-dom2'$ ;
20:
21:      $Non-dom := Non-dom1' + Non-dom2'$ .
22:   end if
23:   if  $(psize - Nmem) \% 2 = 1$  then
24:     Select  $(psize - Nmem + 1)/2$  individuals from  $Non-dom1$ , as  $Non-dom1^*$ ;
25:     Select  $(psize - Nmem - 1)/2$  individuals from  $Non-dom2$ , as  $Non-dom2^*$ ;
26:
27:      $Non-dom := Non-dom1^* + Non-dom2^*$ .
28:   end if
29:    $Nnondom := psize - Nmem$ .
30: end if

```

ALGORITHM 3: Calculating the final nondominated set in $t+1$ time step and its size.

F10 are two of the most difficult problems to optimize. In F9, for the most part, the environment changes slightly, but sometimes PS can occasionally jump from one area to another. In F10, the geometry of two consecutive PFs is different.

4.2. Performance Indicator. Two performance metrics, MIGD and MHVD, were used to evaluate the performance of these strategies regarding convergence and distribution.

4.2.1. Modified Inverted Generational Distance (MIGD). MIGD [37, 41, 45, 46] is a widely used indicator. Before MIGD, IGD metric is described firstly [37, 51, 56]. Here, let PF_t be a set of uniformly distributed Pareto optimal points of PF in t time and let P_t be an approximation set of PF in t time. The formula is as follows:

$$IGD(PF_t, P_t) = \frac{\sum_{v \in PF_t} d(v, P_t)}{|PF_t|} \quad (8)$$

Input: $Nmem$, the capacity of memory set; $psize$, the size of the population; Pop_{t-1} and Pop_t , the population in t-1 and t time step.
Output: The final non-dominated set, $Non-dom$ and its size, $Nnondom$.
Step 1: Layer the Pop_t , and select the individuals in the first layer as the non-dominated set, $Non-dom_t$, and get its capacity, $Nnondom_t$;
Step 2: According to Algorithm 1, calculate the non-dominated set, $Non-dom1$;
Step 3: According to Algorithm 2, calculate the non-dominated set, $Non-dom2$;
Step 4: According to Algorithm 3, calculate the final non-dominated set, $Non-dom$, and its size, $Nnondom$;

ALGORITHM 4: Hybrid predictive strategy.

Input: $Nmem$, the capacity of memory set; $psize$, the size of the population; $Nnondom$, the capacity of the non-dominated set; Low and Up , the upper and lower bounds of the decision vector; m , the dimensions of the decision space.
Output: Pop_{rand} , the random set and its size, $Nrand$.
1: **if** $Nnondom + Nmem \leq psize$ **then**
2: $Nrand := psize - Nnondom - Nmem$
3: **for all** $Ind \in Pop_{rand}$ **do**
4: $i := 0$;
5: **while** $i < m$ **do**
6: $Ind_i := \text{rand}(Low_i, Up_i)$
7: **end while**
8: **end for**
9: **else**
10: $Nrand := 0$
11: $Pop_{rand} := \text{NULL}$
12: **end if**

ALGORITHM 5: An adaptive diversity maintenance strategy.

where $d(v, P_t) = \min_{u \in P_t} \|F(v) - F(u)\|$ is the distance between v and P_t . $|PF_t|$ is the cardinality of PF_t .

Because the environment is dynamic, if only using simply IGD, it is hard to judge which algorithm is better between two algorithms in some situations [34]. MIGD is a modified version of IGD, which is defined as the average value of the IGD values in some time steps over a run.

$$MIGD = \frac{1}{|T|} \sum_{t \in T} IGD(PF_t, P_t) \quad (9)$$

where T is a set of discrete time points in a run and $|T|$ is the cardinality of T . MIGD is also a comprehensive performance metric to evaluate the performance of algorithms, such as convergence and distribution.

The smaller the value of MIGD is, the better the performance of the algorithm is.

4.2.2. Modified Hypervolume Difference (MHVD). The Hypervolume Difference (HVD) [34, 41] measures the gap between the hypervolumes of the got PF and the true PF. PF_t denotes a set of uniformly distributed Pareto optimal points of the PF in t time step and P_t is an approximation set of PF in t time step.

$$HVD(PF_t, P_t) = HV(PF_t) - HV(P_t) \quad (10)$$

where $HV(S)$ denotes the hypervolume of a set S .

MHVD is got by modifying HVD like MIGD to IGD. MHVD shows the average of the HVD values in some time steps over a run.

$$MHVD = \frac{1}{|T|} \sum_{t \in T} HVD(PF_t, P_t) \quad (11)$$

where T is a set of discrete time points over a run and $|T|$ is the cardinality of T . The reference point for the computation of hypervolume is $(Z_1^t + 0.5, Z_2^t + 0.5, \dots, Z_M^t + 0.5)$, where Z_j^t is the maximum value of the j-th objective of the true PF in t time; M is the number of objectives. MHVD is also a comprehensive indicator. The smaller the value is, the better the performance of the algorithm is.

5. Experimental Results and Analysis

In this paper, we chose four commonly used strategies for comparison: (1) random initialization strategy (RIS), which will randomly initialize the population when the environment changes, (2) feed-forward prediction strategy (FPS) [36], (3) predictive multiobjective genetic algorithm (PMGA) [39], and (4) population prediction strategy (PPS) [37]. RM-MEDA [51] was used as the optimization algorithm.

5.1. Parameter Setting. In this article, set the number of individuals in the population to 100. The running generation

Initialization: number of time change, $t := 0$; generation counter, $gt := 0$; total generation number, $gmax$.

Step 1: Initialize the population, Pop_t .

Step 2: Detect the environmental change. If no change, go to step 8; else, calculate the non-dominated set in the current population.

Step 3: If $t=1$, go to step 1.

Step 4: Select randomly $Nmem$ individuals from non-dominated set, as memory set, Pop_{mem} ;

Step 5: According to Algorithm 4, get the non-dominated set in $t+1$ time step, $Non-dom$ and its size, N_{nondom} .

Step 6: According to Algorithm 5, get the random set, Pop_{rand} .

Step 7: Get the predicted population in the new environment by equation (*).

$$Pop_{t+1} = Pop_{mem} + Non-dom + Pop_{rand}. \quad (*)$$

Step 8: The optimization algorithm *RM – MEDA* [51] is used to optimize the problem.

Step 9: If $gt > gmax$, output Pop_t , and then end; else, $gt:=gt+1$; go to step 2.

ALGORITHM 6: A hybrid predictive strategy carried through simultaneously from decision space and objective space.

number of the algorithm is set as 2500. And the environment changes once in 25 generations, so the number of environmental changes is 100. The dimension of the decision space is set as 20.

The Parameters in FPS. The number of the predicted individuals is $3(m+1)$ and m is the dimensions of the objective space. Seventy percent of the remaining individuals are inherited from the previous environment. The remaining 30% are randomly generated within the scope of the decision space.

The Parameters in PPS. The length of the historical sequence M is 23; the parameter p is 3.

The Parameters in DOPS. The size of memory set, $Nmem$, is set to 10.

Change Detection. Five percent of the individuals were randomly selected to detect environmental changes. If the objective value of an individual is detected to be inconsistent with the original objective value, the environment is considered to have changed.

In the experiment, we ran the algorithms 20 times on average and then took the average values to reduce the experimental errors.

5.2. Comparison on Performance Evaluation Results. In a dynamic environment, in some cases, when we compare the performance of different strategies, we need to analyze them at different time periods [34, 37, 54, 57]. In this paper, 100 environmental changes were divided into three stages. The first 20 environmental changes were the first stage; the middle 40 were the second stage, and the last 20 were the third stage. The mean and standard deviation of the MIGD values were calculated to compare the performance of the algorithms where the data marked black indicates that the MIGD value of the strategy is the minimum, which means that the strategy has the best performance. Wilcoxon rank sum test [58] was used to analyze the significance of different results at the significance level of 0.05.

5.2.1. Performance Comparison on FDA and dMOP Test Suites. As shown in Table 1, the five strategies were compared for the MIGD values in the FDA and dMOP test suites. In each small cell, this value represents the mean and standard deviation of the MIGD value. Total represents all environmental changes, which means 100 environmental changes.

(1) For the total stage, DOPS is slightly worse than with FPS on FDA2 and dMOP1. But, on other problems, DOPS is better than all other strategies. This shows that DOPS has good overall performance. Among them, RIS has the worst effect; that is, these prediction strategies are better than the random algorithm. This shows that the prediction strategy is not blind and effective in dealing with DMOPs.

(2) For the 1st stage, DOPS was slightly worse than FPS for FDA2 and dMOP1, but better than the other strategies for FDA3 and dMOP2. This shows that DOPS has the ability to respond quickly to environmental changes.

(3) For 2nd and 3rd stages, DOPS was slightly worse than PPS on test problems FDA1, dMOP2, and dMOP3.

(4) For 3-dimensional problem, FDA4, DOPS is better than other prediction strategies at total and 1st stages. This shows that DOPS is effective for multiobjective optimization problems whose objective number is bigger than 2.

MHVD values are shown in Table 2. The results in Table 2 are almost the same with Table 1.

From these experimental results, the reasons can be guessed. PPS needs to learn for a long time to accumulate experience, resulting in the fact that the effect is not good in the early stage. DOPS, on the other hand, does not require too long time to learn but processes simultaneously from decision space and objective space, so that it has a better prediction effect. However, when PPS accumulates enough experience, the prediction effect will be better. The specific performance in 2nd and 3rd stages was better than DOPS on some test problems. FPS showed good effect on FDA1 and dMOP1, because both FDA1 and dMOP1 were invariant PS problems. In FPS, 70% of the newly generated individuals were inherited from the optimal solutions in the previous environment, so FPS showed better results on the two test problems than the other four strategies.

TABLE 1: Mean and Standard Deviation of MIGD values of five strategies on FDA and dMOP test suites. The values in bold face denote having the best effect in these five strategies. ‡ and † indicate that DOPS is significantly better than and is equivalent to the corresponding strategy, respectively.

Problems	Statistic	RIS	FPS	PMGA	PPS	DOPS
FDA1	Total	1.3155(0.03030)‡	0.0516(0.00864)‡	0.0581(0.00316)‡	0.0528(0.00913)‡	0.0303(0.00623)
	1st stage	1.2215(0.07518)‡	0.2090(0.04391)‡	0.0634(0.00658)	0.2406(0.04615)‡	0.1102(0.03141)
	2nd stage	1.3310(0.04145)‡	0.0151(0.00117)‡	0.0567(0.00606)‡	0.0102(0.00103)	0.0113(0.00021)
	3rd stage	1.3447(0.04696)‡	0.0134(0.00084)‡	0.0581(0.00316)‡	0.0062(0.00008)	0.0113(0.00028)
FDA2	Total	0.0500(0.00078)‡	0.0085(0.00068)	0.4435(0.01275)‡	0.0097(0.00075)	0.0100(0.00042)
	1st stage	0.0491(0.00115)‡	0.0198(0.00332)	0.5137(0.04302)†	0.0232(0.00317)†	0.0213(0.00237)
	2nd stage	0.0503(0.00123)‡	0.0060(0.00034)	0.4269(0.01171)‡	0.0070(0.00068)	0.0073(0.00015)
	3rd stage	0.0501(0.00130)‡	0.0056(0.00003)	0.4435(0.01275)‡	0.0060(0.00004)	0.0074(0.00009)
FDA3	Total	1.7564(0.06554)‡	0.0645(0.00927)‡	0.7617(0.00886)‡	0.0941(0.01582)‡	0.0388(0.00442)
	1st stage	1.5737(0.12477)‡	0.2084(0.04439)‡	0.7552(0.01274)‡	0.3209(0.07801)‡	0.1302(0.02582)
	2nd stage	1.7709(0.11832)‡	0.0305(0.00339)‡	0.7610(0.01097)‡	0.0420(0.00821)‡	0.0170(0.00160)
	3rd stage	1.8287(0.11044)‡	0.0303(0.00421)‡	0.7617(0.00886)‡	0.0384(0.00675)‡	0.0171(0.00207)
FDA4	Total	0.4566(0.00922)‡	0.1414(0.00337)‡	3.7558(0.00341)‡	0.1307(0.00205)‡	0.0171(0.00207)
	1st stage	0.4390(0.02098)‡	0.1629(0.00852)‡	0.1583(0.00239)‡	0.1660(0.00816)‡	0.1368(0.00365)
	2nd stage	0.4594(0.01512)‡	0.1376(0.00427)‡	4.6121(0.01405)‡	0.1247(0.00278)‡	0.1368(0.00365)
	3rd stage	0.4622(0.01131)‡	0.1386(0.00340)‡	3.7558(0.00341)‡	0.1231(0.00283)‡	0.1368(0.00365)
dMOP1	Total	0.6386(0.01427)‡	0.0072(0.00115)	0.2437(0.01891)‡	0.0379(0.05152)†	0.0105(0.00102)
	1st stage	0.6486(0.03460)‡	0.0195(0.00610)	0.2037(0.02545)‡	0.1413(0.21119)†	0.0234(0.00512)
	2nd stage	0.6383(0.02790)‡	0.0043(0.00008)	0.2724(0.02718)‡	0.0209(0.02865)‡	0.0075(0.00012)
	3rd stage	0.6341(0.01865)‡	0.0043(0.00007)	0.2437(0.01891)‡	0.0057(0.00002)	0.0074(0.00018)
dMOP2	Total	1.6968(0.05407)‡	0.0622(0.00788)‡	0.1451(0.00650)‡	0.0607(0.01023)‡	0.0379(0.00504)
	1st stage	1.6332(0.08749)‡	0.2552(0.03931)‡	0.1774(0.01274)‡	0.2799(0.05178)‡	0.1416(0.02592)
	2nd stage	1.7087(0.09220)‡	0.0170(0.00113)‡	0.1384(0.00727)‡	0.0111(0.00103)	0.0131(0.00029)
	3rd stage	1.7150(0.06717)‡	0.0159(0.00080)‡	0.1451(0.00650)‡	0.0061(0.00006)	0.0134(0.00020)
dMOP3	Total	1.3215(0.03752)‡	0.0523(0.00654)‡	0.0581(0.00316)‡	0.0527(0.01084)‡	0.0312(0.00889)
	1st stage	1.2558(0.08243)‡	0.2124(0.03335)‡	0.0634(0.00658)	0.2403(0.05530)‡	0.1158(0.04631)
	2nd stage	1.3319(0.05383)‡	0.0149(0.00077)‡	0.0567(0.00606)‡	0.0101(0.00093)	0.0112(0.00022)
	3rd stage	1.3423(0.06014)‡	0.0136(0.00076)‡	0.0581(0.00316)‡	0.0062(0.00010)	0.0111(0.00035)

5.2.2. *Performance Comparison on F5-F10.* As shown in Table 3, mean values and standard deviation of MIGD values are put into it. From Table 3, some things can be seen.

For two-dimensional problems, F5, F6, and F7, DOPS is better than the other four strategies in total, 1st, and 2nd stages. This shows that DOPS still has better performance on problems where there are the nonlinear relationships between decision variables. But, in the third stage, the effect is slightly worse than PPS. The explanation in Section 5.2.1 can be used here. PPS needs to accumulate experience; when accumulating enough experience, it will show better results. For particularly complex problems F9 and F10, DOPS performed better at all stages than the other four strategies, possibly because the problems were so complex that the predictions of the FPS and PPS were inaccurate. The prediction in DOPS carried through simultaneously from the decision space and objective space can alleviate this situation. Coupled with the adaptive diversity maintenance strategy in DOPS, the effect of DOPS is better. In addition, in Tables 1 and 3, the standard deviation of DOPS is relatively small, which means that DOPS is robust. This is because DOPS simultaneously predicts from decision space and objective space, so that the

two predictions are complementary to each other, making DOPS more robust to deal with various problems.

Table 4 shows the MHVD values on F5-F10. The results are almost identical with Table 3, but it seems better on F5. For F5, DOPS is better than the other strategies on all stages.

5.3. *Final Population Diagram.* In order to compare the performance of various strategies more intuitively, we selected the final population distribution diagram of five strategies on four test problems with different characteristics. As shown in Figures 6–9, the blue line represents the true PF of the population at time t , and every red dot represents an individual in the population. As the test problem is FDA1 in Figure 6, the PF of the test problem is invariant, so 6 moments were selected to observe the effect, while in other problems 8 moments were selected.

At certain moment, the more individuals are closer to the PF at that moment, indicating that the effect is better. As shown in Figures 6 and 7, the effect is almost identical to that shown in Tables 1 and 2. In the early stage, many individuals in DOPS converge to the PF surface. While in other strategies fewer individuals converge to the PF. In the

TABLE 2: Mean and Standard Deviation of MHVD values of five strategies on FDA and dMOP test suites. The values in bold face denote having the best effect in these five strategies. ‡ and † indicate that DOPS is significantly better than and is equivalent to the corresponding strategy, respectively.

Problems	Statistic	RIS	FPS	PMGA	PPS	DOPS
FDA1	Total	1.2328(0.01073)‡	0.0968(0.01053)‡	0.1326(0.00669)‡	0.0948(0.01448)‡	0.0507(0.00401)
	1st stage	1.2145(0.01900)‡	0.3680(0.05223)‡	0.1424(0.01292)	0.4192(0.07139)‡	0.1566(0.01827)
	2nd stage	1.2353(0.01564)‡	0.0342(0.00249)‡	0.1302(0.00557)‡	0.0222(0.00227)	0.0255(0.00062)
	3rd stage	1.2390(0.01275)‡	0.0305(0.00222)‡	0.1301(0.01206)‡	0.0133(0.00022)	0.0257(0.00062)
FDA2	Total	0.0714(0.00125)‡	0.0320(0.00079)	0.7283(0.01669)‡	0.0325(0.00072)	0.0330(0.00067)
	1st stage	0.0719(0.00221)‡	0.0372(0.00413)	0.9557(0.05030)‡	0.0405(0.00408)†	0.0396(0.00322)
	2nd stage	0.0714(0.00209)‡	0.0306(0.00026)	0.6743(0.00871)‡	0.0301(0.00029)	0.0314(0.00014)
	3rd stage	0.0711(0.00188)‡	0.0310(0.00004)	0.6742(0.00871)‡	0.0310(0.00005)	0.0315(0.00010)
FDA3	Total	1.9361(0.01767)‡	0.7761(0.01825)‡	0.8604(0.00959)‡	0.8420(0.02751)†	0.6712(0.00683)
	1st stage	1.9195(0.04820)‡	1.0479(0.05214)†	0.9852(0.01236)‡	1.2829(0.14896)†	0.7806(0.00898)
	2nd stage	1.9387(0.02533)‡	0.7167(0.01667)‡	0.8310(0.01652)‡	0.7436(0.01854)‡	0.6377(0.01288)
	3rd stage	1.9414(0.02960)‡	0.7064(0.01855)‡	0.8310(0.02220)‡	0.7310(0.02692)‡	0.6528(0.00784)
FDA4	Total	1.3838(0.01837)‡	0.4249(0.01308)‡	1.3108(0.01204)‡	0.3818(0.00756)‡	0.3410(0.00648)
	1st stage	1.3392(0.03975)‡	0.4993(0.03182)‡	0.4503(0.00852)‡	0.5100(0.02893)‡	0.4030(0.01961)
	2nd stage	1.3942(0.02764)‡	0.4051(0.01404)‡	1.5138(0.01953)‡	0.3541(0.01032)‡	0.3239(0.00633)
	3rd stage	1.3945(0.02189)‡	0.4093(0.01419)‡	1.5165(0.01596)‡	0.3487(0.01104)‡	0.3287(0.00605)
dMOP1	Total	1.1531(0.01823)‡	0.1501(0.00097)	0.4267(0.03460)‡	0.1688(0.03108)†	0.1505(0.00098)
	1st stage	1.0885(0.04142)‡	0.1180(0.00484)†	0.3385(0.05409)‡	0.1715(0.09166)†	0.1178(0.00341)
	2nd stage	1.1718(0.03312)‡	0.1577(0.00018)	0.3985(0.02189)‡	0.1780(0.03702)†	0.1586(0.00044)
	3rd stage	1.1651(0.02816)‡	0.1576(0.00020)	0.4968(0.04069)‡	0.1583(0.00015)†	0.1580(0.00086)
dMOP2	Total	1.2672(0.01126)‡	0.2156(0.01314)‡	0.2138(0.01806)‡	0.2190(0.01537)‡	0.1718(0.00203)
	1st stage	1.1689(0.02085)‡	0.4303(0.06200)‡	0.2783(0.02538)‡	0.4484(0.07664)‡	0.2107(0.01138)
	2nd stage	1.2967(0.02139)‡	0.1656(0.00309)‡	0.1933(0.01614)‡	0.1709(0.00237)‡	0.1619(0.00049)
	3rd stage	1.2844(0.02055)‡	0.1636(0.00140)†	0.2036(0.02210)‡	0.1581(0.00015)	0.1632(0.00062)
dMOP3	Total	1.2282(0.00948)‡	0.0972(0.00848)‡	0.1326(0.00669)‡	0.0947(0.01470)‡	0.0521(0.00452)
	1st stage	1.2046(0.02570)‡	0.3695(0.04400)‡	0.1424(0.01292)	0.4193(0.07355)‡	0.1654(0.02333)
	2nd stage	1.2317(0.01575)‡	0.0340(0.00172)‡	0.1302(0.00557)‡	0.0219(0.00206)	0.0254(0.00040)
	3rd stage	1.2359(0.02242)‡	0.0312(0.00184)‡	0.1302(0.01206)‡	0.0133(0.00024)	0.0250(0.00069)

latter stages, DOPS and PPS work best. As can be seen from Figures 8 and 9, the effect of the final population distribution diagram is almost consistent with that in Tables 3 and 4. Both F6 and F10 are complex problems, especially F10. As can be seen from Figure 9, on RIS, FPS, PMGA, and PPS, few individuals converge on PF, while there are more in DOPS, indicating that DOPS has better effect on complex problems.

6. More Discussion

From the previous sections to see, DOPS is effective for DMOPs. However, DOPS is a hybrid prediction strategy: it includes the prediction strategy based on the center point and the prediction strategy based on CTI. In order to more accurately judge the role of the two strategies playing in dealing with DMOPs, the two strategies were compared experimentally. Here, only two strategies are compared and the memory set and the random set are no longer needed. The forms of the two strategies here are similar to formulas (6) and (7). But in formulas (6) and (7) the meanings of Pop_k^t and Pop_k^{t+1} have changed. Instead of representing individuals in a nondominated set, they represent individuals in the whole

population. The prediction strategy based on the center point is called CPS, while the prediction strategy based on CTI is called IPS. CPS and IPS were compared on five test problems with different characteristics. The characteristics of the five test questions are shown in Table 5. As shown in Table 5, the number of objectives, change types, and relationship between decision variables are listed.

The MIGD comparison results of CPS and IPS on the five problems are shown in Table 6. From Table 6, we can observe the following.

(1) On dMOP3, the MIGD values of IPS are lower than CPS at all stages. This means that IPS is better than CPS in the problem of fixed PF. This may be because CTI is a special point on PF. When PF is constant, CTI can be used to judge the evolution direction of the population and the location and distribution of PF in the new environment more accurately.

(2) On F5, the MIGD values of CPS are smaller than IPS at all stages. This shows that the effect of CPS is better than effect of IPS on the problem with nonlinear relationship between decision variables; that is, on relatively complex problems, CPS is better than IPS. This may be because when dealing with nonlinear problems, only CTI is not enough to

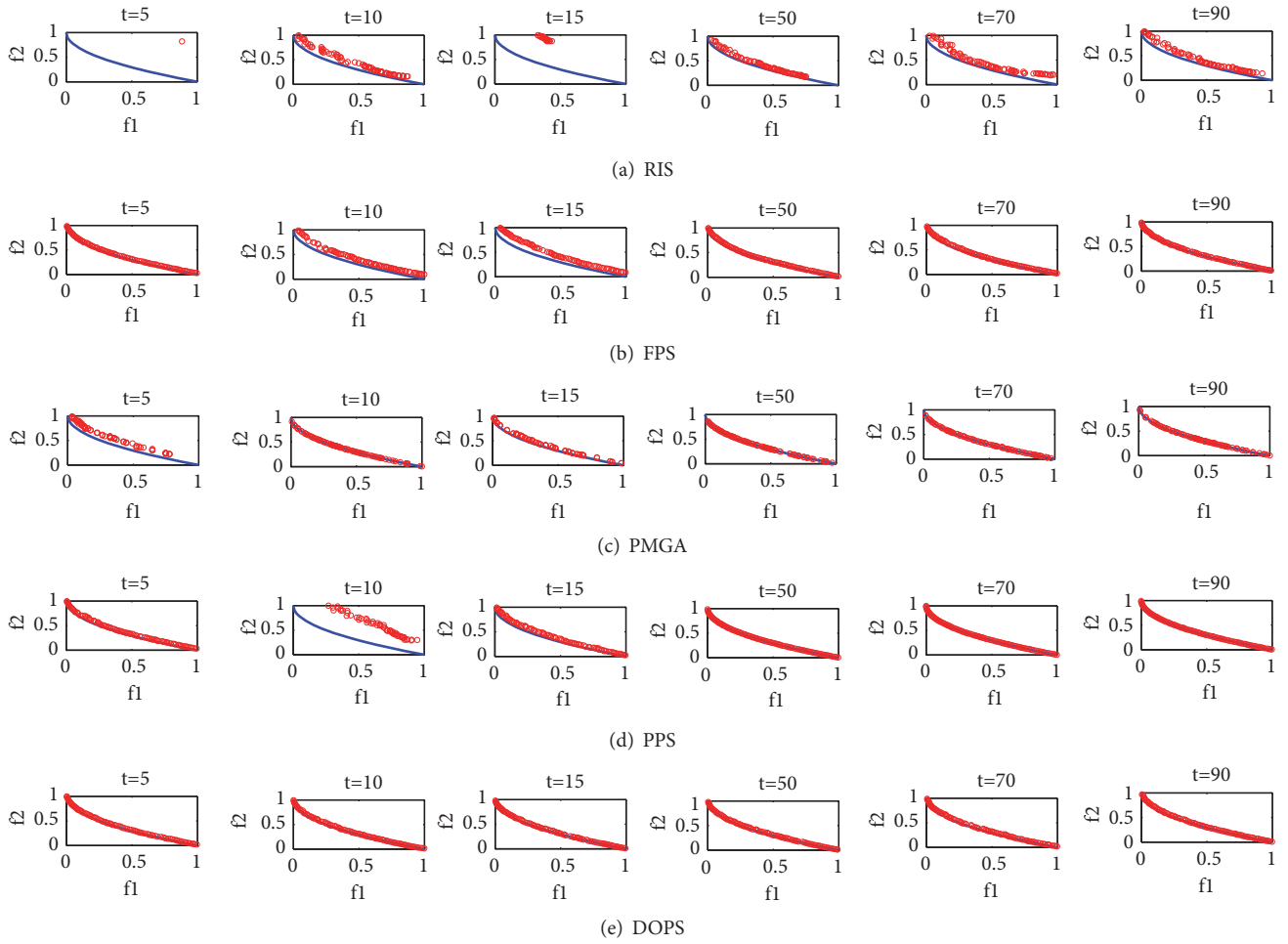


FIGURE 6: Final population distribution of the five strategies on FDA1.

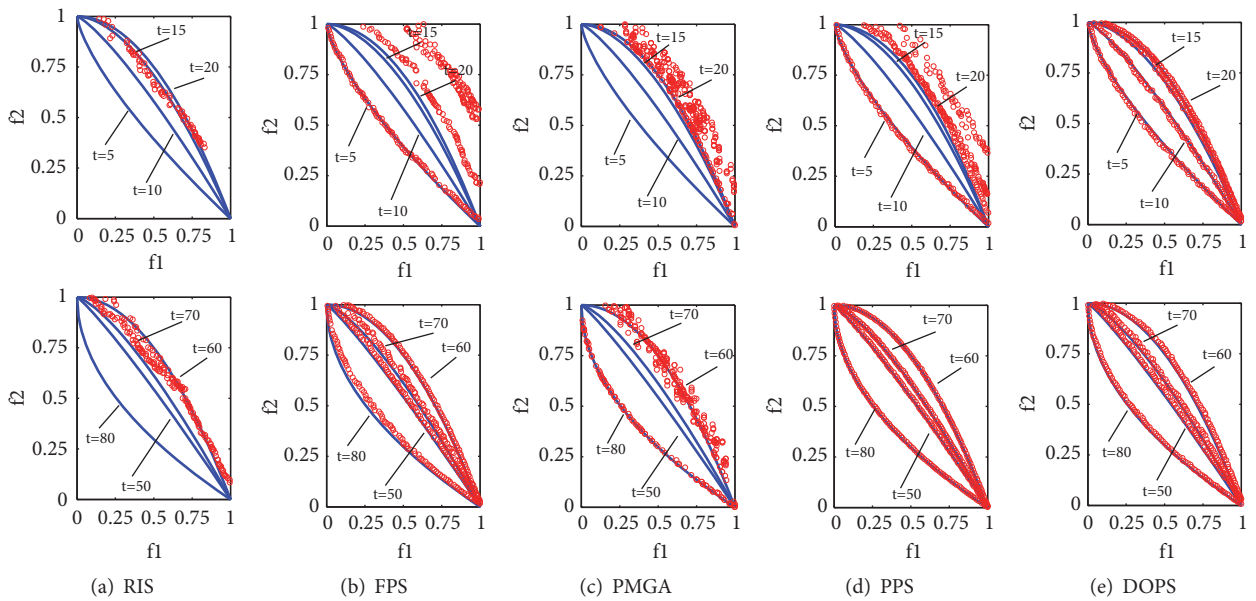


FIGURE 7: Final population distribution of the five strategies on dMOP2.

TABLE 3: Mean and Standard Deviation of MIGD values of five strategies on F5-F10. The values in bold face denote having the best effect in these five strategies. ‡ and † indicate that DOPS is significantly better than and is equivalent to the corresponding strategy, respectively.

Problems	Statistic	RIS	FPS	PMGA	PPS	DOPS
F5	Total	1.1439(0.04184)‡	0.1852(0.08194)‡	3.1110(0.07348)‡	0.2323(0.07728)‡	0.0442(0.01145)
	1st stage	1.1844(0.06775)‡	0.5886(0.41127)‡	2.4771(0.36662)‡	1.0473(0.36248)‡	0.1467(0.05625)
	2nd stage	1.1344(0.04050)‡	0.1088(0.02892)‡	3.2192(0.24029)‡	0.0664(0.02698)‡	0.0201(0.00147)
	3rd stage	1.1341(0.06016)‡	0.0746(0.02526)‡	3.1110(0.07348)‡	0.0169(0.00131)	0.0197(0.00085)
F6	Total	0.5399(0.01228)‡	0.0548(0.01680)‡	1.2521(0.01678)‡	0.0751(0.04238)‡	0.0241(0.00281)
	1st stage	0.6958(0.03924)‡	0.1291(0.07744)‡	1.2312(0.06075)‡	0.3084(0.20490)‡	0.0515(0.01462)
	2nd stage	0.5103(0.01898)‡	0.0404(0.00820)‡	1.2799(0.03064)‡	0.0269(0.00916)‡	0.0171(0.00035)
	3rd stage	0.4956(0.01674)‡	0.0352(0.00504)‡	1.2521(0.01678)‡	0.0143(0.00050)	0.0181(0.00060)
F7	Total	0.6165(0.01535)‡	0.1273(0.02343)‡	1.2892(0.08301)‡	0.1006(0.04024)‡	0.0451(0.01848)
	1st stage	0.6764(0.03353)‡	0.3499(0.10311)‡	1.3816(0.11334)‡	0.4575(0.20361)‡	0.0451(0.01848)
	2nd stage	0.6009(0.01925)‡	0.0879(0.02554)‡	1.2468(0.09617)‡	0.0208(0.00415)‡	0.0160(0.00054)
	3rd stage	0.6037(0.02552)‡	0.0642(0.02467)‡	1.2892(0.08301)‡	0.0133(0.00055)	0.0165(0.00037)
F8	Total	0.9083(0.02482)‡	0.1418(0.00363)†	0.3795(0.00647)‡	0.1455(0.00463)‡	0.1408(0.00249)
	1st stage	0.7666(0.04113)‡	0.1944(0.01645)†	0.3978(0.02375)‡	0.2106(0.02369)‡	0.1760(0.01330)
	2nd stage	0.9473(0.04443)‡	0.1313(0.00293)	0.3708(0.00756)†	0.1335(0.00198)†	0.1316(0.00167)
	3rd stage	0.9366(0.02643)‡	0.1309(0.00169)	0.3795(0.00647)‡	0.1301(0.00250)	0.1333(0.00238)
F9	Total	1.1923(0.03253)‡	0.3542(0.06750)‡	2.5320(0.11486)‡	0.6186(0.19477)‡	0.1305(0.02763)
	1st stage	1.2308(0.10033)‡	0.9770(0.19262)‡	4.2814(0.09784)‡	2.5294(0.95190)‡	0.4243(0.20966)
	2nd stage	1.1787(0.06723)‡	0.2203(0.07491)‡	1.2583(0.00682)‡	0.2632(0.09615)‡	0.0613(0.02554)
	3rd stage	1.1875(0.04852)‡	0.2012(0.08490)‡	2.5320(0.11486)‡	0.0819(0.01601)‡	0.0602(0.02883)
F10	Total	1.0691(0.04720)‡	0.4280(0.05312)‡	0.7665(0.05569)‡	0.5097(0.09978)‡	0.1388(0.03141)
	1st stage	1.1510(0.07724)‡	0.6341(0.12031)‡	1.1785(0.25386)‡	1.5974(0.44011)‡	0.2993(0.09248)
	2nd stage	1.0608(0.04913)‡	0.4188(0.05751)‡	0.6958(0.01573)‡	0.3004(0.08774)‡	0.1162(0.03636)
	3rd stage	1.0385(0.07501)‡	0.3499(0.06662)‡	0.7665(0.05569)‡	0.2150(0.04395)‡	0.0852(0.01222)

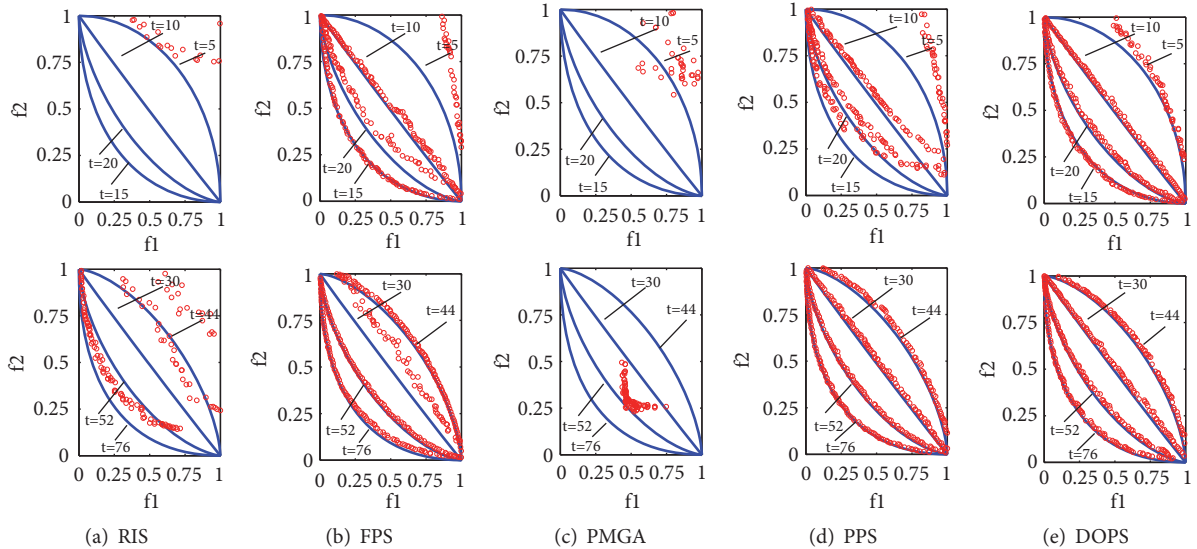


FIGURE 8: Final population distribution of the five strategies on F6.

accurately judge the PF or PS in the new environment, and the error will be smaller if the center point is used.

(3) For FDA2, CPS performed better in the total and 1st stages than IPS. This indicates that CPS is faster and more accurate than IPS in responding to environmental changes

with invariant PS problems. But in the 2nd and 3rd stages, CPS is slightly less effective than IPS. For 3-dimensional problem, FDA4, the effect is similar to that of FDA2. This may be because, with the progress of the prediction process, the judgment of CTI will become more and more accurate,

TABLE 4: Mean and Standard Deviation of MHVD values of five strategies on F5-F10. The values in bold face denote having the best effect in these five strategies. ‡ and † indicate that DOPS is significantly better than and is equivalent to the corresponding strategy, respectively.

Problems	Statistic	RIS	FPS	PMGA	PPS	DOPS
F5	Total	1.3849(0.01419)‡	0.4745(0.07067)‡	1.7364(0.01230)‡	0.4845(0.05354)‡	0.2923(0.01296)
	1st stage	1.5641(0.03251)‡	0.9525(0.29421)‡	1.8196(0.04988)‡	1.3391(0.21636)‡	0.4868(0.06892)
	2nd stage	1.3441(0.01687)‡	0.3951(0.05421)‡	1.7344(0.03510)‡	0.3148(0.04150)‡	0.2463(0.00293)
	3rd stage	1.3406(0.02808)‡	0.3269(0.04382)‡	1.6988(0.01811)‡	0.2483(0.00111)‡	0.2459(0.00219)
F6	Total	0.9995(0.01531)‡	0.3050(0.03246)‡	1.5555(0.00974)‡	0.3436(0.04867)‡	0.2483(0.00122)
	1st stage	1.3182(0.04079)‡	0.4923(0.14787)†	1.7031(0.03430)‡	0.7177(0.22655)‡	0.3684(0.03515)
	2nd stage	0.9273(0.02093)‡	0.2646(0.02006)‡	1.5148(0.01269)‡	0.2626(0.01394)‡	0.2481(0.00091)
	3rd stage	0.9204(0.02104)‡	0.2565(0.00858)†	1.5259(0.03027)‡	0.2468(0.00105)	0.2483(0.00122)
F7	Total	1.0988(0.01522)‡	0.4448(0.04393)‡	1.5220(0.02041)‡	0.3625(0.02877)‡	0.2654(0.00732)
	1st stage	1.3190(0.03876)‡	0.8952(0.17527)‡	1.7286(0.01664)‡	0.8413(0.14139)‡	0.3448(0.03924)
	2nd stage	1.0465(0.02740)‡	0.3599(0.05159)‡	1.4745(0.02450)‡	0.2509(0.00448)‡	0.2469(0.00117)
	3rd stage	1.0464(0.02178)‡	0.3157(0.04522)‡	1.4715(0.02634)‡	0.2467(0.00126)†	0.2463(0.00204)
F8	Total	2.2789(0.02309)‡	0.3797(0.00916)†	0.9777(0.02016)‡	0.3954(0.01577)‡	0.3783(0.00758)
	1st stage	2.1534(0.06297)‡	0.5431(0.04274)‡	1.0026(0.06811)‡	0.6076(0.07844)‡	0.4932(0.03185)
	2nd stage	2.3030(0.02905)‡	0.3413(0.00900)	0.9828(0.04316)‡	0.3504(0.00781)†	0.3474(0.00564)
	3rd stage	2.3144(0.03879)‡	0.3405(0.00750)	0.9609(0.01666)‡	0.3397(0.00780)	0.3547(0.00668)
F9	Total	1.3930(0.01395)‡	0.6393(0.08717)‡	1.7018(0.01392)‡	0.6481(0.07278)‡	0.3643(0.00475)
	1st stage	1.5966(0.03298)‡	1.2334(0.16517)‡	1.9478(0.00387)‡	1.6530(0.26199)‡	0.7094(0.08916)
	2nd stage	1.3371(0.01545)‡	0.5259(0.10356)‡	1.7470(0.02775)‡	0.5175(0.10465)‡	0.2814(0.01612)
	3rd stage	1.3523(0.03077)‡	0.4705(0.12764)‡	1.5398(0.00912)‡	0.3014(0.02611)†	0.2834(0.02490)
F10	Total	1.3534(0.01879)‡	0.8905(0.07059)‡	1.1394(0.02832)‡	0.7476(0.05592)‡	0.4360(0.04524)
	1st stage	1.5587(0.03333)‡	1.2692(0.11566)‡	1.7628(0.10585)‡	1.5369(0.09760)‡	0.7386(0.11058)
	2nd stage	1.3022(0.02225)‡	0.8541(0.09269)‡	1.0906(0.03339)‡	0.6028(0.09317)‡	0.3880(0.06687)
	3rd stage	1.3071(0.02902)‡	0.7471(0.10528)‡	0.8921(0.00826)‡	0.5175(0.07129)‡	0.3403(0.01635)

TABLE 5: Five test problems with different characteristics.

Problems	Number of objectives	Changes types	Relationship between decision variables
FDA2	2	PF changes; PS is fixed	Linear
FDA4	3	PF is fixed; PS changes	Linear
dMOP3	2	PF is fixed; PS changes	Linear
F5	2	PF changes; PS changes	Nonlinear
F10	2	PF changes; PS changes	Nonlinear

TABLE 6: Mean and Standard Deviation of MIGD values of CPS and IPS. The values in bold face denote having the better effect in these two strategies.

Problems	Strategy	Total	1st stage	2nd stage	3rd stage
FDA2	CPS	0.0087(0.00138)	0.0165(0.00138)	0.0073(0.00180)	0.0063(0.00120)
	IPS	0.0087(0.00064)	0.0187(0.00228)	0.0068(0.00133)	0.0058(0.00048)
FDA4	CPS	0.1521(0.00348)	0.1679(0.00649)	0.1489(0.00323)	0.1479(0.00513)
	IPS	0.1530(0.00408)	0.1797(0.00843)	0.1468(0.00648)	0.1467(0.00477)
dMOP3	CPS	0.1551(0.01011)	0.2850(0.03400)	0.1210(0.02425)	0.1275(0.01648)
	IPS	0.1349(0.01312)	0.2160(0.04473)	0.1074(0.01234)	0.1239(0.02219)
F5	CPS	0.5321(0.07584)	0.6463(0.24566)	0.5106(0.11876)	0.4995(0.10182)
	IPS	0.6699(0.21041)	0.7706(0.29603)	0.7780(0.41179)	0.5141(0.08521)
F10	CPS	2.1766(0.21428)	1.6762(0.30120)	2.4446(0.16529)	2.1462(0.31530)
	IPS	1.9126(0.16360)	1.4834(0.31371)	2.0169(0.43282)	2.0120(0.45767)

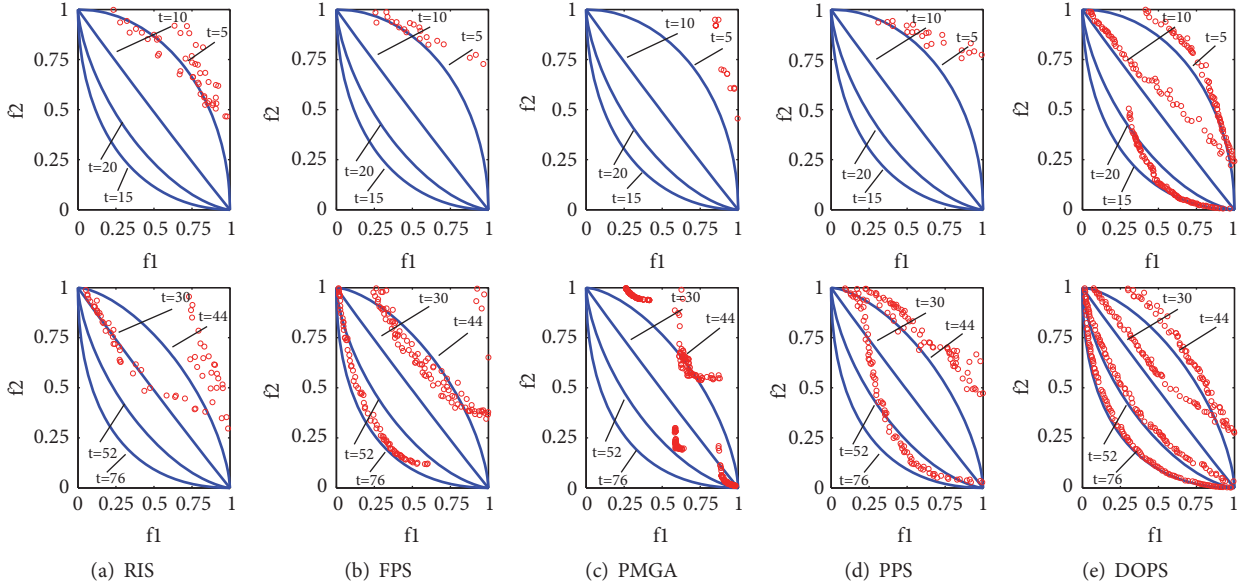


FIGURE 9: Final population distribution of the five strategies on F10.

so the prediction with CTI will become more and more accurate.

(4) For F10, we can see that IPS is better than CPS at all stages. This is a very strange phenomenon because F10 is the most complex among these problems. Its two consecutive PFs have different shapes and CTI is a point on PF. IPS is better may be because of two reasons as follows.

(a) Due to the complexity of F10, the prediction method based on the center point is not effective.

(b) The PF shape of F10 is suitable for IPS to predict.

In order to more intuitively observe the effect of CPS and IPS, we drew the trend chart of IGD of two strategies on dMOP3 and F5 with the change of environment. The smaller the IGD value is, the better the algorithm performance is. As shown in Figure 10, the abscissa represents the time step, which means that in which change the environment lies, and the ordinate represents the IGD value at that time step. The effect shown in Figure 10 is almost identical to that shown in Table 6. For example, on dMOP3, in the early stage, IPS is much better than CPS; in the later stage, IPS is slightly better than CPS.

7. Conclusion and Future Work

In this paper, a hybrid prediction strategy is proposed, which is carried through in decision space and objective space simultaneously. Mixed prediction strategy includes center point-based strategy and CTI-based strategy. The two strategies predict simultaneously from the decision space and objective space and obtain the predicted nondominant set of the next environment.

Nondominant set is the main body of the predicted population and plays an important role in guiding the evolution of population. In addition, in order to deal with the problem with periodic changes, some optimal individuals in the current population are introduced into the predicted

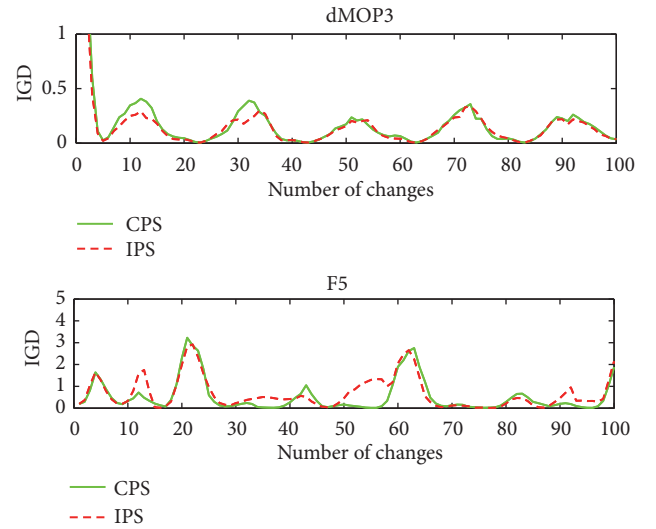


FIGURE 10: IGD trend comparison of CPS and IPS over number of changes for 20 runs on dMOP3 and F5.

population as memory set. Finally, a random set is introduced into the population. The number of the random individuals in the random set can vary adaptively according to the difficulty degree of the problem. The more difficult the problem is, the more random the individuals are in the random set and vice versa. Then, the two prediction strategies in the hybrid prediction strategy were compared to study the role of each strategy in solving the problem. Through experiments, it can be observed that both strategies have their own test problems that are suitable to solve. The two prediction strategies are just combined, so that they are complementary to each other to deal with DMOPs.

By comparing DOPS with four classical prediction strategies on 13 test problems with different characteristics, it can

be seen that DOPS is effective in dealing with DMOPs. In addition, we will try to apply DOPS to practical problems, like scheduling resources in integrating 5G networks and the Internet of things, and make contributions to human development.

Data Availability

The data used to support the findings of this study are included within the article.

Ethical Approval

This article does not contain any studies with human participants or animals performed by any of authors.

Consent

Informed consent was obtained from all individual participants included in the study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors wish to thank the support of the National Key Research and Development Program of China, Research and demonstration on key technologies for intelligent monitoring and early warning of supervision sites (2017YFC0803400) and Shandong Provincial Natural Science Foundation, China (No. ZR2018MF003).

References

- [1] X. Ben, C. Gong, P. Zhang, R. Yan, Q. Wu, and W. Meng, "Coupled bilinear discriminant projection for cross-view gait recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, 2019.
- [2] J. Chi, S. Gao, Y. Zhang, and C. Zhang, "A region-based expression tracking algorithm for spacetime faces," *Computers and Graphics*, vol. 51, pp. 98–105, 2015.
- [3] X. Ben, C. Gong, P. Zhang, X. Jia, Q. Wu, and W. Meng, "Coupled patch alignment for matching cross-view gaits," *IEEE Transactions on Image Processing*, vol. 28, no. 6, pp. 3142–3157, 2019.
- [4] P. Di Barba, "Dynamic multiobjective optimization: a way to the shape design with transient magnetic fields," *IEEE Transactions on Magnetics*, vol. 44, no. 6, pp. 962–965, 2008.
- [5] S. Han, Y. Zhang, W. Meng, and H. Chen, "Self-interference-cancellation-based SLNR precoding design for full-duplex relay-assisted system," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 9, pp. 8249–8262, 2018.
- [6] A. K. Hutzschenreuter, P. A. Bosman, and H. La Poutré, "Evolutionary multiobjective optimization for dynamic hospital resource management," in *Evolutionary Multi-Criterion Optimization*, vol. 5467 of *Lecture Notes in Computer Science*, pp. 320–334, Springer, Berlin, Germany, 2009.
- [7] J. Tang, S. Alam, C. Lokan, and H. A. Abbass, "A multi-objective evolutionary method for dynamic airspace re-sectorization using sectors clipping and similarities," in *Proceedings of the 2012 IEEE Congress on Evolutionary Computation, CEC 2012*, pp. 3565–3572, June 2012.
- [8] K. Deb and S. Karthik, "Dynamic multi-objective optimization and decision-making using modified NSGA-II: a case study on hydro-thermal power scheduling criterion optimization," in *Proceedings of the International Conference on Evolutionary Multi-Criterion Optimization*, pp. 803–817, Springer, 2007.
- [9] M. B. Abello, L. T. Bui, and Z. Michalewicz, "An adaptive approach for solving dynamic scheduling with time-varying number of tasks - Part II," in *Proceedings of the 2011 IEEE Congress of Evolutionary Computation, CEC 2011*, pp. 1711–1718, June 2011.
- [10] M. B. Abello, L. T. Bui, and Z. Michalewicz, "An adaptive approach for solving dynamic scheduling with time-varying number of tasks C Part I," in *Proceedings of the IEEE Congress on Evolutionary Computation (CEC 2011)*, pp. 1703–1710, 2011.
- [11] A. Isaacs, V. Puttige, T. Ray, W. Smith, and S. Anavatti, "Development of a memetic algorithm for dynamic multi-objective optimization and its applications for online neural network modeling of UAVs," in *Proceedings of the 2008 International Joint Conference on Neural Networks, IJCNN 2008*, pp. 548–554, June 2008.
- [12] L. T. Bui and Z. Michalewicz, "An evolutionary multi-objective approach for dynamic mission planning," in *Proceedings of the IEEE Congress on Evolutionary Computation (CEC 2010)*, pp. 1–8, 2010.
- [13] A. R. da Cruz, R. T. Cardoso, and R. H. Takahashi, "Multiobjective dynamic optimization of vaccination campaigns using convex quadratic approximation local search," in *Evolutionary Multi-Criterion Optimization*, vol. 6576 of *Lecture Notes in Computer Science*, pp. 404–417, Springer, Berlin, Germany, 2011.
- [14] P. P. Wu, D. Campbell, and T. Merz, "Multi-objective four-dimensional vehicle motion planning in large dynamic environments," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 41, no. 3, pp. 621–634, 2011.
- [15] B. Andrés-Toro, J. M. Girón-Sierra, P. Fernández-Blanco, J. López-Orozco, and E. Besada-Portas, "Multiobjective optimization and multivariable control of the beer fermentation process with the use of evolutionary algorithms," *Journal of Zhejiang University SCIENCE A*, vol. 5, no. 4, pp. 378–389, 2004.
- [16] Z. Zhang, "Multiobjective optimization immune algorithm in dynamic environments and its application to greenhouse control," *Applied Soft Computing*, vol. 8, no. 2, pp. 959–971, 2008.
- [17] S. Zeng, S. Chen, J. Zhao, A. Zhou, Z. Li, and H. Jing, "Dynamic constrained multi-objective model for solving constrained optimization problem," in *Proceedings of the 2011 IEEE Congress on Evolutionary Computation (CEC)*, pp. 2041–2046, New Orleans, LA, USA, June 2011.
- [18] S. Zeng, S. Chen, J. Zhao, A. Zhou, Z. Li, and H. Jing, "Dynamic constrained multi-objective model for solving constrained optimization problem," in *Proceedings of the 2011 IEEE Congress of Evolutionary Computation, CEC 2011*, pp. 2041–2046, June 2011.
- [19] K. Kim, R. I. McKay, and B.-R. Moon, "Multiobjective evolutionary algorithms for dynamic social network clustering," in *Proceedings of the 12th Annual Genetic and Evolutionary Computation Conference, GECCO-2010*, pp. 1179–1186, July 2010.
- [20] B. S. Rabil, R. Sabourin, and E. Granger, "Watermarking stack of grayscale face images as dynamic multi-objective optimization problem," in *Proceedings of the 6th International Conference*

- on *Mass Data Analysis of Signals and Images in Medicine, Chemistry, Biotechnology and Food Industry, MDA 2011*, pp. 63–77, September 2011.
- [21] M. Linnala, E. Madetoja, H. Ruotsalainen, and J. Hämäläinen, “Bi-level optimization for a dynamic multiobjective problem,” *Engineering Optimization*, vol. 44, no. 2, pp. 195–207, 2012.
 - [22] X. Ben, P. Zhang, Z. Lai, R. Yan, X. Zhai, and W. Meng, “A general tensor representation framework for cross-view gait recognition,” *Pattern Recognition*, vol. 90, pp. 87–98, 2019.
 - [23] J. Chi, C. Tu, and C. Zhang, “Dynamic 3D facial expression modeling using Laplacian smooth and multi-scale mesh matching,” *The Visual Computer*, vol. 30, no. 6-8, pp. 649–659, 2014.
 - [24] S. Han, S. Xu, W. Meng, and C. Li, “Dense-device-enabled cooperative networks for efficient and secure transmission,” *IEEE Network*, vol. 32, no. 2, pp. 100–106, 2018.
 - [25] S. Han, Y. Huang, W. Meng, C. Li, N. Xu, and D. Chen, “Optimal power allocation for scma downlink systems based on maximum capacity,” *IEEE Transactions on Communications*, vol. 67, no. 2, pp. 1480–1489, 2019.
 - [26] J. Chi, S. Gao, and C. Zhang, “Interactive facial expression editing based on spatio-temporal coherency,” *The Visual Computer*, vol. 33, no. 6-8, pp. 981–991, 2017.
 - [27] L. T. Bui, Z. Michalewicz, E. Parkinson, and M. B. Abello, “Adaptation in dynamic environments: a case study in mission planning,” *IEEE Transactions on Evolutionary Computation*, vol. 16, no. 2, pp. 190–209, 2012.
 - [28] K. Deb, S. Agrawal, A. Pratap et al., “A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II,” in *Proceedings of the International Conference on Parallel Problem Solving From Nature*, pp. 849–858, Springer, 2000.
 - [29] E. Zitzler, M. Laumanns, and L. Thiele, “SPEA 2: Improving the Strength Pareto Evolutionary algorithm,” Tech. Rep. 103, Computer Engineering and Networks Laboratory, ETH, Zurich, Switzerland. Priddeleau, redevance prleve surlusager, 2001.
 - [30] C. A. C. Coello, G. B. Lamont, and D. A. van Veldhuizen, *Evolutionary Algorithms for Solving Multi-Objective Problems*, Springer, New York, NY, USA, 2007.
 - [31] L. Chun’an and W. Yuping, “Multiobjective evolutionary algorithm for dynamic nonlinear constrained optimization problems,” *Journal of Systems Engineering and Electronics*, vol. 20, no. 1, pp. 204–210, 2009.
 - [32] C. Liu and Y. Wang, “New evolutionary algorithm for dynamic multiobjective optimization problems,” in *Proceedings of the International Conference on Natural Computation*, pp. 889–892, Springer, 2006.
 - [33] M. Greeff and A. P. Engelbrecht, “Solving dynamic multi-objective problems with vector evaluated particle swarm optimisation,” in *Proceedings of the IEEE Congress on Evolutionary Computation (IEEE World Congress on Computational Intelligence)*, pp. 2917–2924, June 2008.
 - [34] A. Zhou, Y. Jin, Q. Zhang et al., “Prediction-based population re-initialization for evolutionary dynamic multi-objective optimization,” in *Proceedings of the International Conference on Evolutionary Multi-Criterion Optimization*, pp. 832–846, Springer, 2007.
 - [35] I. Hatzakis and D. R. Wallace, “Topology of anticipatory populations for evolutionary dynamic multi-objective optimization,” in *Proceedings of the 11th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, pp. 1944–1953, September 2006.
 - [36] I. Hatzakis and D. Wallace, “Dynamic multi-objective optimization with evolutionary algorithms: a forward-looking approach,” in *Proceedings of the 8th Annual Genetic and Evolutionary Computation Conference*, pp. 1201–1208, ACM, July 2006.
 - [37] A. Zhou, Y. Jin, and Q. Zhang, “A Population prediction strategy for evolutionary dynamic multiobjective optimization,” *IEEE Transactions on Cybernetics*, vol. 44, no. 1, pp. 40–53, 2014.
 - [38] Y. Wu, Y. Jin, and X. Liu, “A directed search strategy for evolutionary dynamic multiobjective optimization,” *Soft Computing*, vol. 19, no. 11, pp. 3221–3235, 2015.
 - [39] Y. Wu, X. X. Liu, and C. Z. Chi, “Predictive multiobjective genetic algorithm for dynamic multiobjective optimization problems,” *Control Decis*, vol. 28, no. 5, pp. 677–682, 2013.
 - [40] J.-H. Zheng, Z. Peng, J. Zou, and R.-M. Shen, “A prediction strategy based on guide-individual for dynamic multi-objective optimization,” *Acta Electronica Sinica*, vol. 43, no. 9, pp. 1816–1825, 2015.
 - [41] S. Jiang and S. Yang, “A steady-state and generational evolutionary algorithm for dynamic multiobjective optimization,” *IEEE Transactions on Evolutionary Computation*, vol. 21, no. 1, pp. 65–82, 2017.
 - [42] L. B. Said, S. Bechikh, and K. Ghédira, “The r-dominance: a new dominance relation for interactive evolutionary multicriteria decision making,” *IEEE Transactions on Evolutionary Computation*, vol. 14, no. 5, pp. 801–818, 2010.
 - [43] J. Molina, L. V. Santana, A. G. Hernández-Díaz, C. A. Coello Coello, and R. Caballero, “g-dominance: reference point based dominance for multiobjective metaheuristics,” *European Journal of Operational Research*, vol. 197, no. 2, pp. 685–692, 2009.
 - [44] L. Thiele, K. Miettinen, P. J. Korhonen, and J. Molina, “A preference-based evolutionary algorithm for multi-objective optimization,” *Evolutionary Computation*, vol. 17, no. 3, pp. 411–436, 2009.
 - [45] Q. Li, J. Zou, S. Yang, J. Zheng, and G. Ruan, “A predictive strategy based on special points for evolutionary dynamic multi-objective optimization,” *Soft Computing*, vol. 1, pp. 1–17, 2018.
 - [46] J. Zou, Q. Li, S. Yang, H. Bai, and J. Zheng, “A prediction strategy based on center points and knee points for evolutionary dynamic multi-objective optimization,” *Applied Soft Computing*, vol. 61, pp. 806–818, 2017.
 - [47] X. Zhang, Y. Tian, and Y. Jin, “A knee point driven evolutionary algorithm for many-objective optimization,” *IEEE Transactions on Evolutionary Computation*, vol. 19, no. 6, pp. 761–776, 2015.
 - [48] I. Das, “On characterizing the “knee” of the Pareto curve based on normal-boundary intersection,” *Journal of Structural Optimization*, vol. 18, no. 2-3, pp. 107–115.
 - [49] J. Branke, K. Deb, H. Dierolf, and M. Osswald, “Finding knees in multi-objective optimization,” in *Proceedings of the International Conference on Parallel Problem Solving from Nature*, vol. 3242, pp. 722–731, Springer, 2004.
 - [50] K. Deb and S. Gupta, “Understanding knee points in bicriteria problems and their implications as preferred solution principles,” *Engineering Optimization*, vol. 43, no. 11, pp. 1175–1204, 2011.
 - [51] Q. F. Zhang, A. M. Zhou, and Y. C. Jin, “RM-MEDA: a regularity model based multiobjective estimation of distribution algorithm,” *IEEE Transactions on Evolutionary Computation*, vol. 12, no. 1, pp. 41–63, 2008.

- [52] M. Farina, K. Deb, and P. Amato, "Dynamic multiobjective optimization problems: test cases, approximations, and applications," *IEEE Transactions on Evolutionary Computation*, vol. 8, no. 5, pp. 425–442, 2004.
- [53] C.-K. Goh and K. C. Tan, "A competitive-cooperative coevolutionary paradigm for dynamic multiobjective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 1, pp. 103–127, 2009.
- [54] Z. Peng, J. Zheng, J. Zou, and M. Liu, "Novel prediction and memory strategies for dynamic multiobjective optimization," *Soft Computing*, vol. 19, no. 9, pp. 2633–2653, 2015.
- [55] Z. Peng, J. Zheng, and J. Zou, "A population diversity maintaining strategy based on dynamic environment evolutionary model for dynamic multiobjective optimization," in *Proceedings of the 2014 IEEE Congress on Evolutionary Computation, CEC 2014*, pp. 274–281, July 2014.
- [56] M. R. Sierra and C. A. C. Coello, "Improving PSO-based multi-objective optimization using crowding, mutation and ϵ -dominance," in *Proceedings of the Evolutionary Multi-Criterion Optimization (EMO 2005)*, vol. 3410 of *Lecture Notes in Computer Science*, pp. 505–519, Springer, 2005.
- [57] G. Ruan, G. Yu, J. Zheng, J. Zou, and S. Yang, "The effect of diversity maintenance on prediction in dynamic multi-objective optimization," *Applied Soft Computing*, vol. 58, pp. 631–647, 2017.
- [58] F. Wilcoxon, "Individual comparisons by ranking methods," *Biomedical Bulletin*, vol. 1, no. 6, pp. 80–83, 1945.

Research Article

Performance Analysis for Downlink MIMO-NOMA in Millimeter Wave Cellular Network with D2D Communications

Jianguo Li , Xiangming Li , Aihua Wang, and Neng Ye 

School of Information and Electronics, Beijing Institute of Technology, Beijing, China

Correspondence should be addressed to Xiangming Li; xmli@bit.edu.cn

Received 4 April 2019; Accepted 3 June 2019; Published 19 June 2019

Guest Editor: Xi Chen

Copyright © 2019 Jianguo Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Enabling nonorthogonal multiple access (NOMA) in device-to-device (D2D) communications under the millimeter wave (mmWave) multiple-input multiple-output (MIMO) cellular network is of critical importance for 5G wireless systems to support low latency, high reliability, and high throughput radio access. In this paper, the closed-form expressions for the outage probability and the ergodic capacity in downlink MIMO-NOMA mmWave cellular network with D2D communications are considered, which indicates that NOMA outperforms TDMA. The influencing factors of performance, such as transmission power and antenna number, are also analyzed. It is found that higher transmission power and more antennas in the base station can decrease the outage probability and enhance the ergodic capacity of NOMA.

1. Introduction

With the rapid growth of a variety of 5th-generation wireless systems commercial requirements, like 8k video, cloud VR, unmanned driving, smart city, etc., the demands on high data rates, low latency, and high reliability of wireless communication are increasing rapidly. As the core technologies of 5G, the wide bandwidth and large scale antenna arrays can be used by mmWave MIMO to provide high data rates. Meanwhile, NOMA can increase the spectral efficiency of transmission so that more data is transmitted over the same bandwidth spectrum. At the same time, due to the short communication distance, D2D communication can obtain lower transmission power and lower delay to ensure the battery availability and system reliability. Therefore, combined with MIMO-NOMA and mmWave in D2D communications, the throughput of the entire cell system can be improved.

D2D communication was noticed as a viable candidate for some certain applications such as proximity services, content sharing, multiparty games, and coverage discovered with business purposes in the 3rd-Generation Partnership Project (3GPP) Long-term Evolution (LTE) 12th and 13th editions [1, 2]. By sharing the same resource blocks with downlink cellular users, D2D communication is a useful technology which can enhance the spectrum efficiency and

system capacity successfully. Due to the short communication distance, D2D pairs can communicate with each other directly without being relayed by the base station which can reduce transmission power, increase transmission reliability, and extend system range [3, 4]. Since the cellular user and the D2D user coexist, if the D2D user and the cellular user are not well coordinated, it will not bring any benefits and may affect the communication of the normal cellular user. According to the literature [5], based on the overlapping coalition formation game theory, the authors have proposed a method to conduct joint interference management and resource allocation in the D2D communications. In addition, a new interference management strategy has been discussed to enhance the overall sum rate of cellular networks and D2D pairs which has combined the conventional mechanism and δ_D -interference limited area control scheme [6]. Furthermore, a graph theoretic approach for transmission-order optimization scheme in bidirectional D2D communications underlying cellular TDD networks has been introduced in the [7]. Meanwhile, the authors in [8] have considered a continuous beamforming vector design for all cellular users and the D2D pairs association vector search algorithm to maximize the capacity of the cellular users and D2D pairs. In order to reduce the interference between the cellular users and D2D pairs, a combining call admission control and power

control scheme under guaranteeing QoS of every users in the cellular network has been presented in [9]. Besides, it is significant that the NOMA and MU-MIMO can also improve the overall system throughput in the cellular network with underlaid D2D communications, which have been proposed in [10].

As the core technology of 5G, NOMA can serve multiple users in the same resource block, such as a time slot, frequency channel, or spreading code. Compared to the orthogonal multiple access, NOMA can provide a set of visible benefits such as greater spectrum efficiency and higher system throughput [11–15]. There are two available NOMA technology categories, namely, power-domain and code-domain NOMA, widely used in the cellular network. This paper focuses on the power-domain NOMA which divides the users in power-domain. On the transmitter sides, signals from multiple users are transmitted at the same resource blocks. On the receiver sides, the multiuser detection algorithms, such as successive interference cancellation, have been utilized to detect the signal, regarding the weak signal as the interference, and decode the strong signal while cancelling the strong signal to detect the weak signal [16, 17]. Aiming at enhancing system capacity and user fairness, the authors have presented a novel resource and power allocation technique which can provide a flexible balancing between capacity and fairness maximization in [18]. Meanwhile, the optimum received power levels of uplink NOMA signal have been proposed in [19]. Furthermore, comparing with OMA in massive connectivity, NOMA can supply more users than OMA because of the limited number of supported users by the amount of the available resources [20–22]. In addition, the analytical outage probability expression for each user has been derived in the downlink cooperative NOMA network over Nakagami- m fading channels in [23]. Since the perfect channel state information at the transmitter side is nearly impractical for many communication scenarios, the literature [24] presented a practical downlink NOMA system over the Nakagami- m fading channels with statistical CSI associated with each other. At the same time, the authors attempted to combine the NOMA and MIMO to accomplish high spectral efficiency [25, 26]. Additionally, the outage probability of the massive MIMO-NOMA system has been derived with perfect user-ordering and limited feedback [27]. Besides, the scheme discussed in [28] required that the number of antennas at the receiver is larger than that at the transmitter instead of the CSI at the transmitter. In [29], the optimal and low complexity suboptimal power allocation schemes have been proposed to maximize the ergodic capacity of MIMO-NOMA system with statistical channel state information in the transmitter over Rayleigh fading channel. Also, the authors in [30] have proposed a cluster beamforming strategy which can optimize beamforming vectors and power allocation coefficients in MIMO-NOMA system to decrease the total power. In addition, NOMA can improve the achievable rate greatly in the D2D-aided cooperative relaying system in the literature [31].

Nowadays, as spectrum resources become increasingly scarce, the mmWave band with a wider spectrum becomes a natural choice for large-volume content services [32]. Meanwhile, benefitting from the small wavelength of mmWave,

the large scale antenna arrays can be adopted easily. Combining the high gain directional antenna and beamforming technology enabled by massive antennas, the high data rates can be reached within a 200-meter mmWave cellular network which has high path loss in mmWave band [33]. In the literature [34], the authors have proposed that the base station which is equipped with a large antenna array can serve a set of users through the users precoding with massive MIMO. Furthermore, the literature [35] has proved that the orders of magnitude are increased in spectral efficiency with massive MIMO which can offer more multiuser gain. However, as for MIMO in conventional cellular frequency band, the beamforming precoding is totally achieved in the digital domain which can cancel the interference between different beams. In order to realize the precoder in the digital domain, a dedicated RF chain is needed by every antenna. But it is difficult for the mmWave base station with large scale antennas whose energy consumption is a large part of the total energy consumption at mmWave frequencies due to the wide bandwidth [36]. At the same time, the authors in [36] have discussed a successive interference cancellation-based hybrid precoding method with near-optimal performance and low complexity which can maximize the achievable substrate of each subantenna array and avoid the need for singular value decomposition and matrix inversion. Additionally, the literature [37] has proposed an interference-aware beam selection which can avoid serious multiuser interferences to reduce obvious performance loss. Besides, the authors in [38] have proposed a practical design of hybrid precoders and combiners with low-resolution phase shifters in mmWave MIMO systems adopting an iterative algorithm for hybrid precoders and combiners to optimize the spectral efficiency. Moreover, the MIMO capacity has been computed over the Nakagami- m fading channel in [39] and the closed-form expressions for outage probabilities achieved by NOMA users in a multicell downlink mmWave network have been obtained over a Poisson cluster point process. Moreover, a fine-grained performance analysis over Poisson bipolar model of the mmWave D2D communication networks was provided in [40]. The literature [41] has presented an in-depth capacity analysis for the integrated NOMA mmWave-massive-MIMO systems which can achieve significant capacity improvements.

Because of the large link attenuation and weak coverage in the MIMO-NOMA cellular network, the D2D communication can be used to enhance service for cell edge users. In this paper, the outage probability and the ergodic capacity are proposed in the downlink MIMO-NOMA cellular network with D2D communications. The remaining part of this paper is organized as follows. In Section 2, we introduce a system model of the downlink MIMO-NOMA cellular network with multiple direct D2D pairs underlying communication. The closed-form expressions of the performance analysis including the outage probability and the ergodic capacity are given in Section 3. Finally, the numerical results in Section 4 validate the theoretical analysis and demonstrate that the system capacity can be improved by the integrated MIMO-NOMA in the mmWave network with D2D communications.

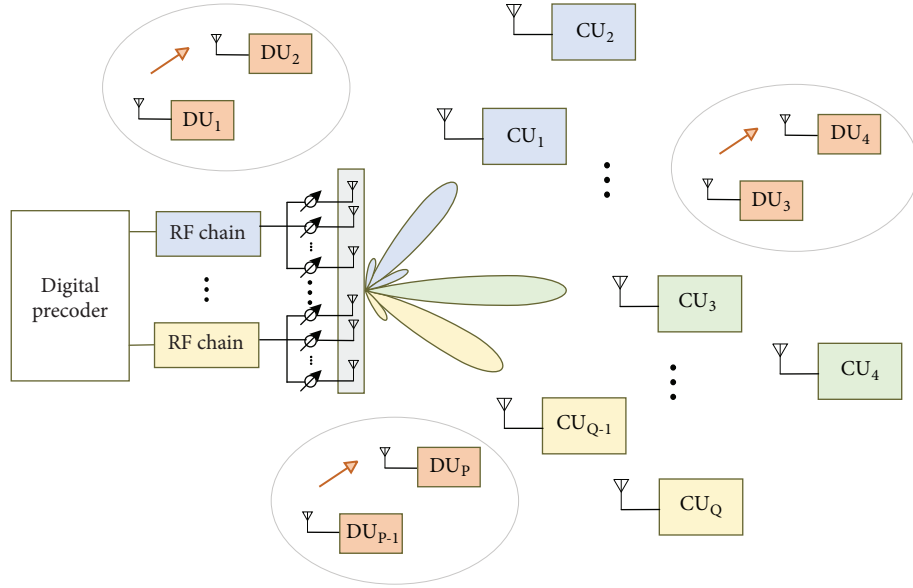


FIGURE 1: System model of D2D-aided mmWave MIMO-NOMA.

2. System Model

The paper considers a downlink NOMA, MU-MIMO mmWave cellular network with multiple direct D2D pairs underlying communications where the cellular users are randomly distributed. The same mmWave resources are used by the cellular user and D2D pairs with NOMA. The mmWave MIMO-NOMA D2D communication system model is shown in Figure 1. The base station is equipped with multiple antennas, which can generate high directional and high gain beams for cellular users. There are Q cellular users with signal antenna and P D2D users with signal antenna in the cellular network, which are denoted as CU_1, CU_2, \dots, CU_Q and DU_1, DU_2, \dots, DU_P . The D2D pairs are randomly distributed in the edge of the cellular network and there is no direct link between the base station and the D2D pairs [10].

2.1. NOMA Signal. In beam n , we assume that $u(n, 1), u(n, 2), \dots, u(n, K)$ are scheduled on the same radio resource with NOMA, $K \geq 2$, where CU of k_{th} in the beam n is denoted as $u(n, k)$. x_n is the transmitted signal by the base station in the beam n , which is the sum of all K user signals.

$$x_n = \sum_{i=1}^K \sqrt{\lambda_{u(n,i)} P_n} S_{u(n,i)}, \quad (1)$$

where $\lambda_{u(n,k)}$ is the power ratio of k_{th} user, $\|\lambda_{u(n,1)}\| \leq \|\lambda_{u(n,2)}\| \leq \dots \leq \|\lambda_{u(n,K)}\|$, and $\sum_{k=1}^K \lambda_{u(n,k)} = 1$. P_n is the total power in the beam n , $S_{u(n,k)}$ is the normalized transmitted signal of k_{th} user in the beam n , and $\mathbb{E}(\|S_{u(n,k)}\|^2) = 1$.

2.2. Channel Model. As for the large scale fading, the mmWave link is similar to that used in [33]; the large scale fading $L(r)$, in dB, is modeled as

$$L(r) = \rho + 10\alpha \log(r), \quad (2)$$

where $\rho = 32.4 + 20 \log(f_c)$, f_c is the carrier frequency, α is the path loss exponents, and r is the distance from transmitter to receiver.

As for the small scale fading, the Nakagami- m fading is considered for each link. $h_{u(n,k)}, h_{pu(n,k)}, h_{p'p}$ are denoted as the link of base station to cellular user, D2D user to cellular user, and D2D user to D2D user whose modular square is normalized Gamma random variable [39]. And when H is normalized Gamma random variable, which is denoted as $H \sim \text{Gamma}(\omega, \psi)$, the probability density function $f(x)$ and Cumulative Distribution Function $F(x)$ of H are

$$f(x) = \frac{1}{\psi^\omega \Gamma(\omega)} x^{\omega-1} e^{-x/\psi}, \quad (3)$$

$$F(x) = \begin{cases} 1 - \sum_{j=0}^{\omega-1} \frac{1}{j! \psi^j} x^j e^{-x/\psi}, & x \geq 0 \\ 0, & \text{otherwise.} \end{cases} \quad (4)$$

2.3. Directional Beamforming. The base station with mmWave band is equipped with multiple antennas which can generate high directional and high gain beams. The actual antenna pattern is modeled as the sectorized antenna model approximately for the sake of mathematical tractability [42]. Generally the maximum power gain is adopted to replace the array gain within the half-power beam width (main lobe gain), and the first minor maximum gain is used to replace the gains of the other DoAs (side lobe gain). According to the literature [39, 40], when the antenna pattern is a planar square, the total array gain from base station to the user is $G_{(n,k)}$ where

$$G_{(n,k)} = \begin{cases} T, & \text{mainlobe} \\ \tau, & \text{sidelobe,} \end{cases} \quad (5)$$

where $T = L$, $\tau = 1/\sin^2(3\pi/2\sqrt{L})$ and L is the number of antennas.

2.4. Received Signal. For downlink MIMO-NOMA transmission, the cellular user $u(n, K)$ will receive the sum of signal from base station and the signal from D2D transmitter at the same time. In addition to receiving the D2D transmitted signal, the D2D receiver DU_p will also receive the interference signal from the other D2D users. Without loss of generality, we assume there are N beams and M D2D pairs in the cellular network; then the received signal of cellular user $y_{u(n,k)}$ and the received signal of D2D user y_{D_p} can be formulated as

$$y_{u(n,k)} = h_{u(n,k)} \sum_{n=1}^N G_{(n,k)} x_n + \sum_{p=1}^M \sqrt{P_D} h_{pu(n,k)} S_p + n_{u(n,k)}, \quad (6)$$

$$y_{D_p} = \sum_{p'=1}^M \sqrt{P_D} h_{p'p} S_{p'} + n_{D_p}, \quad (7)$$

where $h_{u(n,k)}$ is the channel gain for downlink $u(n, K)$, $\|h_{u(n,1)}\| \geq \|h_{u(n,2)}\| \geq \|h_{u(n,K)}\|$, $h_{pu(n,k)}$ is the channel gain between DU_p and $u(n, K)$, and $h_{p'p}$ is the channel gain from $DU_{p'}$ transmitter to DU_p receiver. $G_{(n,k)}$ is the total beamforming array gain from base station to $u(n, K)$. x_n is the superimposed signal by the total K $u(n, k)$ in the beam n . P_D is the transmitted power of DU_p . S_p is the signal transmitted by DU_p , and $\mathbb{E}(\|S_p\|^2) = 1$. Meanwhile $n_{u(n,k)}$ and n_{D_p} are the i.i.d. white Gaussian noise with zero mean and one variance at cellular user $u(n, k)$ and D2D user DU_p , which is denoted as $n_{u(n,k)}, n_{D_p} \sim \mathcal{CN}(0, 1)$.

We denote the SINRs of $u(n, k)$ and DU_p in the downlink NOMA-MIMO cellular network as $\gamma_{u(n,k)}, \gamma_{DU_p}$. Without loss of generality, the interbeam interference is ignored in this paper. In the meantime, the perfect SIC is used to prevent error propagation in the NOMA users in the paper. Using (1) in (6) and (7), $\gamma_{u(n,k)}$ and γ_{DU_p} can be formulated as

$$\gamma_{u(n,k)} = \frac{\lambda_{u(n,k)} P_n \|h_{u(n,k)} G_{(n,k)}\|^2}{I_{u(n,k)}^N + I_{u(n,k)}^D + \sigma_n^2}, \quad (8)$$

$$\gamma_{D_p} = \frac{P_D \|h_{pp}\|^2}{I_{D_p}^D + \sigma_n^2}, \quad (9)$$

where

$$I_{u(n,k)}^N = \sum_{k'=1}^{k-1} \lambda_{u(n,k')} P_n \|h_{u(n,k)} G_{(n,k)}\|^2, \quad (10)$$

$$I_{u(n,k)}^D = \sum_{p'=1}^M P_D \|h_{p'u(n,k)}\|^2, \quad (11)$$

$$I_{D_p}^D = \sum_{p'=1, p' \neq p}^M P_D \|h_{p'p}\|^2. \quad (12)$$

3. Performance Analysis

In this section, we present the performance analysis of D2D-aided mmWave MIMO-NOMA system. Specifically, the closed-form expressions for the performance metrics (i.e., the outage probability and the ergodic capacity) are presented in the following. Without loss of generality, in the beam n , k -th CUs are adopted with NOMA in one beam, $k \in \{1, 2\} \|h_{u(n,1)}\| \geq \|h_{u(n,2)}\|$, and one DU is randomly distributed at the edge of the beam. Three events are considered in this system.

Event 1. According to the NOMA successive interference cancellation (SIC) principle, user 1 obtains the information intended for user 2 with $\gamma_{1 \rightarrow 2}$ and removes it. When decoding the information intended for user 2, user 1 cancels it successfully with γ_1 .

$$\begin{aligned} \gamma_{1 \rightarrow 2} &= \frac{\lambda_{u(n,2)} P_n \|h_{u(n,1)} G_{(n,1)}\|^2}{\lambda_{u(n,1)} P_n \|h_{u(n,1)} G_{(n,1)}\|^2 + P_D \|h_{pu(n,1)}\|^2 + \sigma_n^2}, \quad (13) \end{aligned}$$

$$\gamma_1 = \frac{\lambda_{u(n,1)} P_n \|h_{u(n,1)} G_{(n,1)}\|^2}{P_D \|h_{pu(n,1)}\|^2 + \sigma_n^2}. \quad (14)$$

Event 2. User 2 decodes the signal with γ_2 , treating user 1 as interference.

$$\gamma_2 = \frac{\lambda_{u(n,2)} P_n \|h_{u(n,2)} G_{(n,2)}\|^2}{\lambda_{u(n,1)} P_n \|h_{u(n,2)} G_{(n,2)}\|^2 + P_D \|h_{pu(n,2)}\|^2 + \sigma_n^2}. \quad (15)$$

Event 3. The DU receiver only receives signals from the DU transmitter, whose SINR is denoted as γ_{D_p} :

$$\gamma_{D_p} = \frac{P_D \|h_{pp}\|^2}{\sigma_n^2}. \quad (16)$$

3.1. Outage Probability. In this section, we study the outage probability of CU and DU. The outage probability of the user 1, user 2, and DU is given by $P_1^{\text{out}}, P_2^{\text{out}}, P_D^{\text{out}}$:

$$\begin{aligned} P_1^{\text{out}} &= P(\log(1 + \gamma_{1 \rightarrow 2}) < R_2 \text{ or } \log(1 + \gamma_1) < R_1) \\ &= 1 \end{aligned} \quad (17)$$

$$- P(\log(1 + \gamma_{1 \rightarrow 2}) \geq R_2) P(\log(1 + \gamma_1) \geq R_1),$$

$$P_2^{\text{out}} = P(\log(1 + \gamma_2) < R_2), \quad (18)$$

$$P_D^{\text{out}} = P(\log(1 + \gamma_{D_p}) < R_D), \quad (19)$$

where R_1, R_2, R_D are the target rates of user 1, user 2, and DU.

Firstly, we consider the outage probability of user 1 P_1^{out} ; then $P(\log(1 + \gamma_{1 \rightarrow 2}) \leq R_2)$ can be rewritten as $P(\gamma_{1 \rightarrow 2} \leq 2^{R_2} - 1)$, so we set

$$F(m, a, b, c, d) = P\left(\frac{aH_1}{bH_1 + dH_2 + c} \leq m\right). \quad (20)$$

According to (13),(20), we can get

$$P(\gamma_{1 \rightarrow 2} \leq 2^{R_2} - 1) = F(m, a, b, c, d) = F_{\gamma_{1 \rightarrow 2}}(m), \quad (21)$$

where $a = \lambda_{u(n,2)} P_n \|G_{(n,1)}\|^2$, $b = \lambda_{u(n,1)} P_n \|G_{(n,1)}\|^2$, $d = P_D$, $c = \sigma_n^2$, $m = 2^{R_2} - 1$, and $H_1 = \|h_{u(n,1)}\|^2 \sim \text{Gamma}(\omega, \psi)$, $H_2 = \|h_{pu(n,1)}\|^2 \sim \text{Gamma}(\eta, \theta)$; then

$$\begin{aligned} F_{\gamma_{1 \rightarrow 2}}(m) &= P\left(\frac{aH_1}{bH_1 + dH_2 + c} \leq m\right) \\ &= \int_0^\infty P\left(\frac{ay}{by + dH_2 + c} \leq m\right) f_{H_1}(y) dy \\ &= \int_0^\infty P\left(\frac{ay - bmy - cm}{dm} \leq H_2\right) f_{H_1}(y) dy. \end{aligned} \quad (22)$$

In order to determine if $(ay - bmy - cm)/dm \leq 0$, we set $\Phi = cm/(a - bm)$. When $\Phi \leq 0$, $(ay - bmy - cm)/dm \leq 0$, so $P((ay - bmy - cm)/dm \leq H_2) = 1$; therefore

$$F_{\gamma_{1 \rightarrow 2}}(m) = 1. \quad (23)$$

When $\Phi > 0$, which is $m < a/b$,

$$\begin{aligned} F_{\gamma_{1 \rightarrow 2}}(m) &= \int_0^\Phi f_{H_1}(y) dy \\ &+ \int_\Phi^\infty \left(1 - F_{H_2}\left(\frac{ay - bmy - cm}{dm}\right)\right) f_{H_1}(y) dy \\ &= \int_\Phi^\infty \sum_{j=0}^{\eta-1} \frac{1}{j! \theta^j} \left(\frac{a - bm}{dm} y - \frac{c}{d}\right)^j \\ &\cdot e^{-((a-bm)/dm\theta)y - c/d\theta} \frac{1}{\psi^\omega \Gamma(\omega)} y^{\omega-1} e^{-y/\psi} dy \\ &+ \int_0^\Phi f_{H_1}(y) dy. \end{aligned} \quad (24)$$

We set $\alpha = (a - bm)/dm$, $\beta = -c/d$, and

$$\begin{aligned} F_{\gamma_{1 \rightarrow 2}}(m) &= \int_0^\Phi f_{H_1}(y) dy + \frac{1}{\psi^\omega \Gamma(\omega)} \sum_{j=0}^{\eta-1} \frac{1}{j! \theta^j} \\ &\cdot \int_\Phi^\infty (\alpha y + \beta)^j e^{-((\alpha/\theta)y + \beta/\theta)} y^{\omega-1} e^{-y/\psi} dy. \end{aligned} \quad (25)$$

As for $(x + a)^k = \sum_{j=0}^k \binom{k}{j} x^j a^{k-j}$, then

$$\begin{aligned} F_{\gamma_{1 \rightarrow 2}}(m) &= \frac{1}{\psi^\omega \Gamma(\omega)} \sum_{j=0}^{\eta-1} \frac{1}{j! \theta^j} \sum_{i=0}^j \binom{j}{i} \alpha^i \beta^{j-i} e^{-\beta/\theta} \\ &\cdot \int_\Phi^\infty (y)^{\omega+i-1} e^{-((\alpha/\theta+1/\psi)y)} dy \\ &+ \int_0^\Phi f_{H_1}(y) dy. \end{aligned} \quad (26)$$

We set

$$J(a, n, x) = e^{ax} \sum_{k=0}^n \frac{(-1)^k k! \binom{n}{k}}{a^{k+1}} x^{n-k}. \quad (27)$$

Then by substituting (27) into (26), $F_{\gamma_{1 \rightarrow 2}}(m)$ can be denoted as

$$\begin{aligned} F_{\gamma_{1 \rightarrow 2}}(m) &= \frac{1}{\psi^\omega \Gamma(\omega)} \sum_{j=0}^{\eta-1} \frac{1}{j! \theta^j} \sum_{i=0}^j \binom{j}{i} \alpha^i \beta^{j-i} e^{-\beta/\theta} \\ &\cdot \left(-J\left(-\left(\frac{\alpha}{\theta} + \frac{1}{\psi}\right), \omega + j - 1, \Phi\right)\right) \\ &+ \int_0^\Phi f_{H_1}(y) dy = \frac{1}{\psi^\omega \Gamma(\omega)} \sum_{j=0}^{\eta-1} \frac{1}{j! \theta^j} \\ &\cdot \sum_{i=0}^j \binom{j}{i} \alpha^i \beta^{j-i} \\ &\cdot e^{-\beta/\theta} \left(-J\left(-\left(\frac{\alpha}{\theta} + \frac{1}{\psi}\right), \omega + j - 1, \Phi\right)\right) \\ &+ \frac{1}{\psi^\omega \Gamma(\omega)} \left(J\left(-\left(\frac{1}{\psi}\right), \omega - 1, \Phi\right)\right) \\ &- J\left(-\left(\frac{1}{\psi}\right), \omega - 1, 0\right). \end{aligned} \quad (28)$$

Similarly, by adopting different parameters with a, b, c, d , which is shown in Table 1, we can obtain $F_{\gamma_{1 \rightarrow 2}}(m), F_{\gamma_1}(m), F_{\gamma_2}(m)$.

In addition, when $H_3 = \|h_{pp}\|^2 \sim \text{Gamma}(\omega, \phi)$, $P_D \|h_{pp}\|^2 / \sigma_n^2 \sim \text{Gamma}(\omega, (P_D / \sigma_n^2) \phi)$. Hence $F_{\gamma_D}(m)$ is denoted as

$$F_{\gamma_D}(m) = 1 - \sum_{j=0}^{\omega-1} \frac{1}{j! ((P_D / \sigma_n^2) \phi)^j} x^j e^{-x \sigma_n^2 / P_D \phi}. \quad (29)$$

Finally, the outage probability of user 1, user 2, and DU can be evaluated by $F_{\gamma_{1 \rightarrow 2}}(m), F_{\gamma_1}(m), F_{\gamma_2}(m), F_{\gamma_D}(m)$, which are formulated as

$$\begin{aligned} P_1^{out} &= 1 - (1 - F_{\gamma_{1 \rightarrow 2}}(2^{R_2} - 1)) \\ &\quad * (1 - F_{\gamma_1}(2^{R_1} - 1)), \end{aligned} \quad (30)$$

$$P_2^{out} = F_{\gamma_2}(2^{R_2} - 1), \quad (31)$$

$$P_D^{out} = F_{\gamma_D}(2^{R_D} - 1). \quad (32)$$

3.2. Ergodic Capacity. The ergodic capacity is the average capacity of channel which can be defined as the instantaneous end-to-end mutual information expectations and denoted as

$$\begin{aligned} C_{erg} &= \mathbb{E}[\log_2(1 + \gamma_1)] + \mathbb{E}[\log_2(1 + \gamma_2)] \\ &\quad + \mathbb{E}[\log_2(1 + \gamma_D)]. \end{aligned} \quad (33)$$

TABLE 1: Parameters of the outage probability.

	a	b	c	d
$F_{\gamma_{1 \rightarrow 2}}(m)$	$\lambda_{u(n,1)} P_n \ G_{(n,1)}\ ^2$	$\lambda_{u(n,2)} P_n \ G_{(n,2)}\ ^2$	σ_n^2	P_D
$F_{\gamma_1}(m)$	$\lambda_{u(n,1)} P_n \ G_{(n,1)}\ ^2$	0	σ_n^2	P_D
$F_{\gamma_2}(m)$	$\lambda_{u(n,2)} P_n \ G_{(n,2)}\ ^2$	$\lambda_{u(n,1)} P_n \ G_{(n,1)}\ ^2$	σ_n^2	P_D

The ergodic capacity of the system can be obtained by substituting (14), (15), and (16) into (33), which is formulated as

$$\begin{aligned}
C_{erg} &= \mathbb{E} \left[\log_2 \left(1 + \frac{\lambda_{u(n,1)} P_n \|h_{u(n,1)G_{(n,1)}}\|^2}{P_D \|h_{pu(n,1)}\|^2 + \sigma_n^2} \right) \right] \\
&+ \mathbb{E} \left[\log_2 \left(1 + \frac{\lambda_{u(n,2)} P_n \|h_{u(n,2)G_{(n,2)}}\|^2}{\lambda_{u(n,1)} P_n \|h_{u(n,2)G_{(n,2)}}\|^2 + P_D \|h_{pu(n,2)}\|^2 + \sigma_n^2} \right) \right] + \mathbb{E} \left[\log_2 \left(1 + \frac{P_D \|h_{pp}\|^2}{\sigma_n^2} \right) \right] \\
&= \mathbb{E} \left[\log_2 \left(1 + \frac{\lambda_{u(n,1)} P_n \|G_{(n,1)}\|^2}{\sigma_n^2} \|h_{u(n,1)}\|^2 + \frac{P_D}{\sigma_n^2} \|h_{pu(n,1)}\|^2 \right) \right] \\
&- \mathbb{E} \left[\log_2 \left(1 + \frac{P_D}{\sigma_n^2} \|h_{pu(n,1)}\|^2 \right) \right] \\
&+ \mathbb{E} \left[\log_2 \left(1 + \left(\frac{\lambda_{u(n,2)} P_n \|G_{(n,2)}\|^2}{\sigma_n^2} \|h_{u(n,2)}\|^2 + \frac{P_D}{\sigma_n^2} \|h_{pu(n,2)}\|^2 \right) \right) \right] \\
&- \mathbb{E} \left[\log_2 \left(1 + \frac{P_D}{\sigma_n^2} \|h_{pu(n,2)}\|^2 + \left(\frac{\lambda_{u(n,1)} P_n \|G_{(n,2)}\|^2}{\sigma_n^2} \|h_{u(n,2)}\|^2 \right) \right) \right] + \mathbb{E} \left[\log_2 \left(1 + \frac{P_D \|h_{pp}\|^2}{\sigma_n^2} \right) \right].
\end{aligned} \tag{34}$$

In order to compute (34), we first compute the first item of the formula. As we set before, $a = \lambda_{u(n,1)} P_n \|G_{(n,1)}\|^2$, $d = P_D$, $c = \sigma_n^2$, $H_1 = \|h_{u(n,1)}\|^2 \sim \text{Gamma}(\omega, \psi)$, $H_2 = \|h_{pu(n,1)}\|^2 \sim \text{Gamma}(\eta, \theta)$, we can get

$$\begin{aligned}
&\mathbb{E} \left[\log_2 \left(1 + \frac{\lambda_{u(n,1)} P_n \|G_{(n,1)}\|^2}{\sigma_n^2} \|h_{u(n,1)}\|^2 + \frac{P_D}{\sigma_n^2} \|h_{pu(n,1)}\|^2 \right) \right] = \mathbb{E} \left[\log_2 \left(1 + \frac{a}{c} H_1 + \frac{d}{c} H_2 \right) \right].
\end{aligned} \tag{35}$$

According to the literature [43, 44], we can get

$$\mathbb{E} [\ln(1+x)] \approx \ln(1 + \mathbb{E}[x]) - \frac{\mathbb{E}[x^2] - (\mathbb{E}[x])^2}{2(1 + \mathbb{E}[x])^2}. \tag{36}$$

Based on (36), we start with $\mathbb{E}[x]$ and $\mathbb{E}[x^2]$, which can be written as

$$\begin{aligned}
&\mathbb{E} \left[\frac{a}{c} H_1 + \frac{d}{c} H_2 \right] \\
&= \int_0^\infty \int_0^\infty \left(\frac{a}{c} x + \frac{d}{c} y \right) f(x) f(y) dx dy \\
&= \int_0^\infty \left(\frac{a}{c} x \right) f(x) dx + \int_0^\infty \left(\frac{d}{c} y \right) f(y) dy \\
&= \frac{a}{c} \mathbb{E}(x) + \frac{d}{c} \mathbb{E}(y) = \frac{a}{c} \omega \psi + \frac{d}{c} \eta \theta,
\end{aligned} \tag{37}$$

$$\begin{aligned}
&\mathbb{E} \left[\left(\frac{a}{c} H_1 + \frac{d}{c} H_2 \right)^2 \right] \\
&= \int_0^\infty \int_0^\infty \left(\frac{a}{c} x + \frac{d}{c} y \right)^2 f(x) f(y) dx dy \\
&= \int_0^\infty \left(\frac{ax}{c} \right)^2 f(x) dx + \int_0^\infty \left(\frac{dy}{c} \right)^2 f(y) dy \\
&\quad + \int_0^\infty \int_0^\infty \frac{2adxy}{c^2} f(x) f(y) dx dy \\
&= \left(\frac{a^2}{c^2} \right) (\omega + 1) \omega (\psi)^2 + \left(\frac{d^2}{c^2} \right) (\eta + 1) \eta (\theta)^2 \\
&\quad + \frac{2ad}{c^2} \omega \eta \psi \theta.
\end{aligned} \tag{38}$$

By substituting (37) and (38) into (35), we can obtain the first item of formula (34), which can be denoted as

$$\begin{aligned}
&\mathbb{E} \left[\log_2 \left(1 + \frac{a}{c} H_1 + \frac{d}{c} H_2 \right) \right] = \log_2(e) \\
&\cdot \left(\ln \left(1 + \frac{a}{c} \omega \psi + \frac{d}{c} \eta \theta \right) \right. \\
&\quad \left. - \frac{(a^2/c^2) \omega (\psi)^2 + (d^2/c^2) \eta (\theta)^2}{2(1 + (a/c) \omega \psi + (d/c) \eta \theta)^2} \right).
\end{aligned} \tag{39}$$

Similarly, we can set a, b, c, d as different parameters to obtain the other items of the formula (34) and then pull everything together. The asymptotic result for the ergodic capacity of the consider system can be obtained.

4. Numerical Results

In this section, the outage probability and the ergodic capacity of MIMO-NOMA mmWave cellular network with D2D communications are investigated. The effects of different parameters on the probability of outage and ergodic capacity are analyzed such as the base station transmission power, the number of base station antennas, the power ratio of NOMA user, and the distance between D2D users. In order to verify the performance of the system, the traditional TDMA theme is adopted as the comparison between the two users of each beam. In particular, the time slot is equally divided by the two users. Hence the capacity of this theme is R_{TDMA} , which is denoted as

$$R_{TDMA} = \frac{1}{2} (\log(1 + \gamma_1) + \log(1 + \gamma_2)). \quad (40)$$

A simplified cellular network system is discussed for the performance analyzed here. The carrier frequency is 28GHz which is commonly used for wireless broadband service. There are 16 antennas in the base station whose coverage radius is 100m. There is single antenna with D2D user. And the distance between D2D users is 30m. Meanwhile, the transmission power of base station and D2D users is 5 dbm. In addition, there are 8 NOMA users and 4 D2D users in the cellular network. The path loss exponent is set as 3. Furthermore, the small scale fading is denoted as $H \sim \text{Gamma}(2, 1)$ which is simplified for the simulation.

4.1. Outage Probability. In this section, we consider the outage probability of the NOMA far user and near user.

Figure 2 depicts the outage probability in the different base station transmission power with $R_1 = 5$ bit/s/Hz and $R_2 = 3.32$ bit/s/Hz. As the base station transmission power increases, it can be seen that the outage probability of the NOMA users decreases with the exponential form. Furthermore, the performance of each user's outage probability in the NOMA scheme is significantly better than the TDMA, and the closed-form solution obtained is consistent with the Monte Carlo simulation results.

In Figure 3, the impact of antenna number in the base station on the outage probability ($R_1 = 5.64$ bit/s/Hz and $R_2 = 4$ bit/s/Hz) is presented. The simulation results effectively verify that the number of antennas of the base station can decrease the users' outage probability in the MIMO-NOMA mmWave cellular network, thereby improving the throughput of the system under the limited time-frequency resources. As can be seen from Figure 3, the number of antennas has a greater impact on user 2 than user 1. When the number of antennas is 36, the outage probability of the system is satisfied which can balance the number of RF chains and the system performance. Then the NOMA scheme performs better than the traditional TDMA in the mmWave MIMO cellular network with D2D communications.

In Figure 4, we discuss the influence of the power ratio coefficient in the cellular network between the NOMA users. It can be seen that the outage probability ($R_1 = 4$ bit/s/Hz and $R_2 = 3$ bit/s/Hz) of the two users in NOMA is balanced when the power ratio coefficient is approximately 0.2.

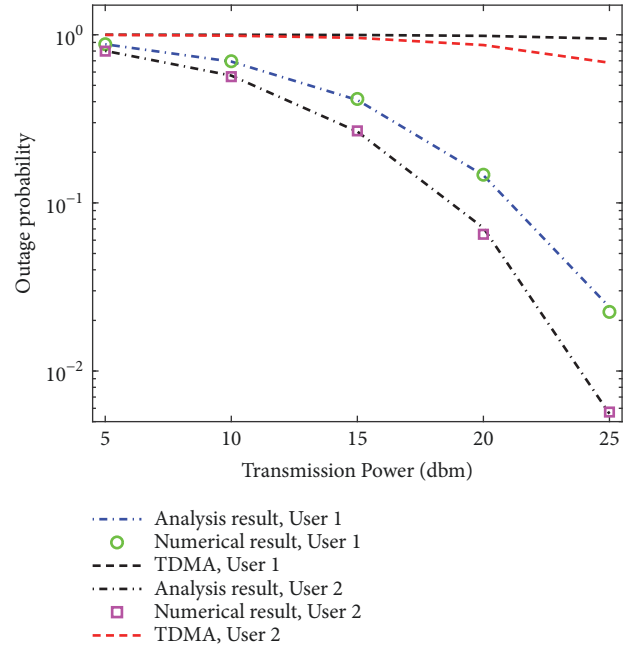


FIGURE 2: Impact of transmission power on outage probability.

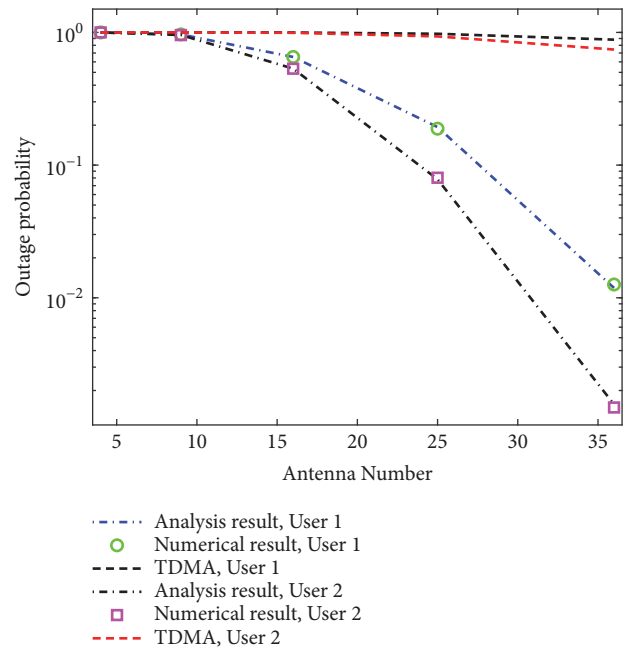


FIGURE 3: Impact of antenna number on outage probability.

In Figure 5, the effect of the distance between D2D users is considered. The figure indicates that the outage probability ($R_1 = 5$ bit/s/Hz and $R_2 = 3.32$ bit/s/Hz) of the NOMA users is reduced in the form of an exponent when the distance of D2D user is linear growth. Since the distance between D2D users is increasing, the interference from the D2D transmitter to the NOMA user is weak. Hence, the throughput of the NOMA users is improved, while the outage probability is dropping.

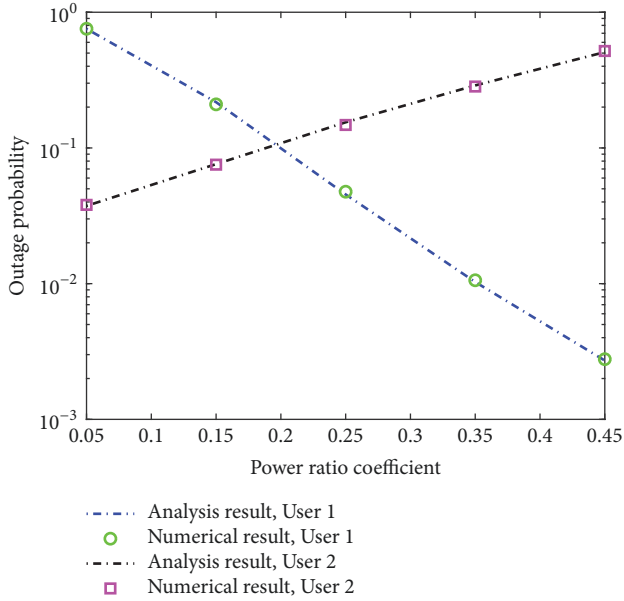


FIGURE 4: Impact of power ratio coefficient on outage probability.

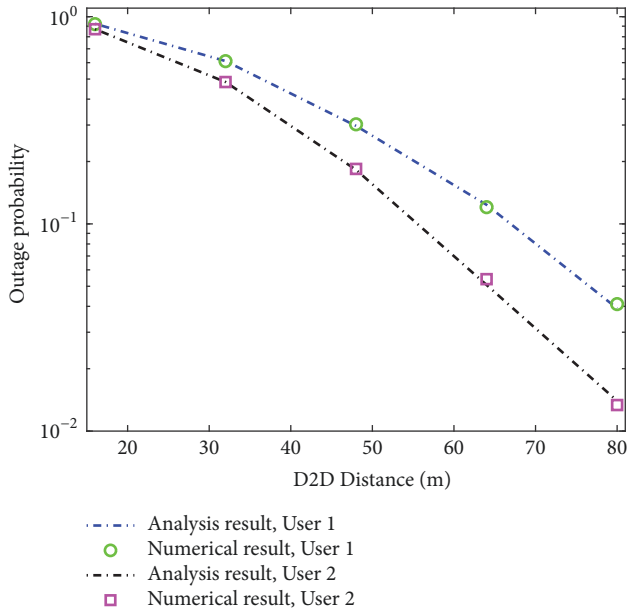


FIGURE 5: Impact of D2D distance on outage probability.

4.2. Ergodic Capacity. In this section, the total ergodic capacity is considered in the MIMO-NOMA mmWave cellular network with D2D communications. It can be seen that the numerical results are consistent with the closed-form solution which is better than the traditional TDMA. Meanwhile, the transmission power and the number of base station antennas have a greater impact, and the power ratio coefficient and the distance between D2D users have less effect.

In Figure 6, the impact of the transmission power of base station on the ergodic capacity is considered in the MIMO-NOMA mmWave cellular network with D2D communications. It is shown that the ergodic capacity is growing linearly

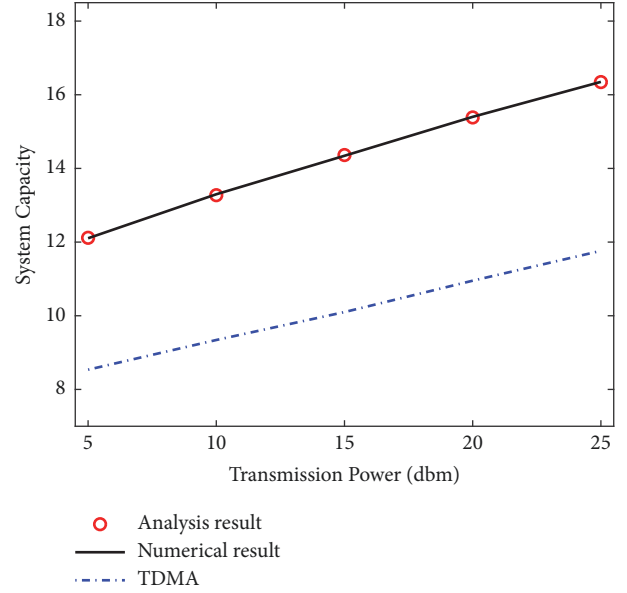


FIGURE 6: Impact of transmission power on ergodic capacity.

with the increase of transmission power. In addition, the ergodic capacity of the system we proposed is higher than the traditional TDMA. Hence, in order to improve the ergodic capacity of the system, we can increase the base station transmission power as much as possible without affecting others.

As the number of the base station antennas is increasing, it is indicated that the ergodic capacity can be improved better than the traditional TDMA in Figure 7. Benefiting from the length of mmWave, more and more antennas can be equipped for the base station. At the same time, we need to balance the improvement in ergodic capacity brought by the increasing of the number of antennas and the power consumption and hardware requirements of the increase in RF chains to determine the final number of antennas.

In Figure 8, the ergodic capacity is affected by the change of the power ratio coefficient of the NOMA users in the MIMO-NOMA mmWave cellular network with D2D communications. It can be seen that the total ergodic capacity changes slowly with the increase of power ratio.

In Figure 9, since the interference from the D2D users is decreasing, the total ergodic capacity is improved with the increase of the distance between the D2D users in the MIMO-NOMA mmWave cellular network. It can also be seen that the ergodic capacity in the MIMO-NOMA mmWave cellular network is always better than the traditional TDMA.

5. Conclusion

In this paper, the outage probability and the ergodic capacity of the NOMA in the MIMO-NOMA mmWave cellular network with D2D communications are studied. The closed-form solutions of the outage probability and the ergodic capacity are obtained, which are consistent with the numerical results. Meanwhile, the performance of NOMA is shown

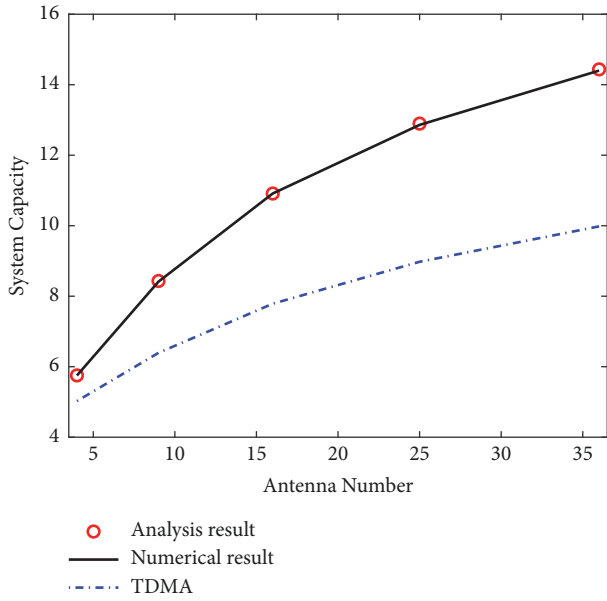


FIGURE 7: Impact of antenna number on ergodic capacity.

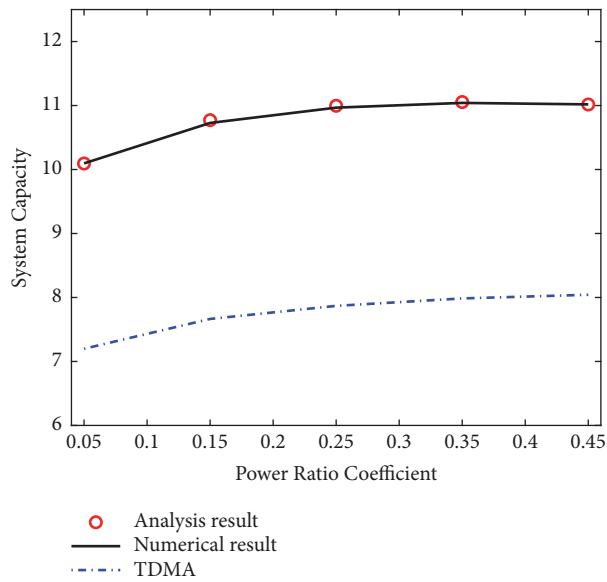


FIGURE 8: Impact of power ratio coefficient on ergodic capacity.

to be better than traditional TDMA in the MIMO mmWave cellular network with D2D communications. Furthermore, the higher transmission power of base station and the larger antenna array can also improve system performance.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

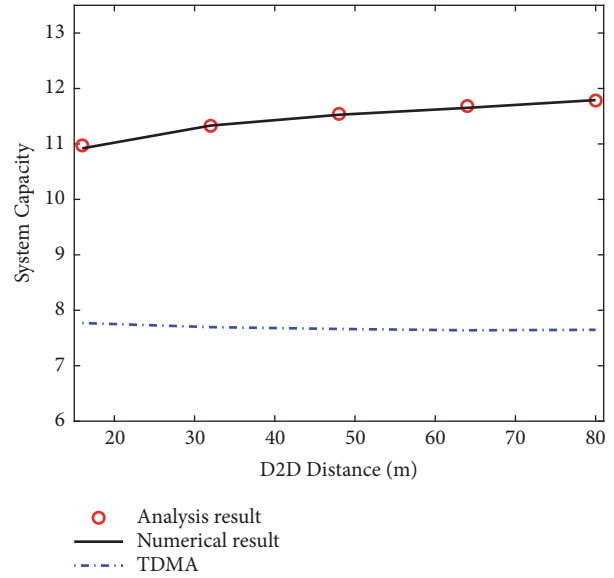


FIGURE 9: Impact of D2D distance on ergodic capacity.

Acknowledgments

This work was supported by Advance Research Projects of 13th Five-Year Plan of Civil Aerospace Technology (B0105) and the National Natural Science Foundation of China (61771051).

References

- [1] M. Tehrani, M. Uysal, and H. Yanikomeroglu, "Device-to-device communication in 5G cellular networks: challenges, solutions, and future directions," *IEEE Communications Magazine*, vol. 52, no. 5, pp. 86–92, 2014.
- [2] S.-Y. Lien, C.-C. Chien, G. S.-T. Liu, H.-L. Tsai, R. Li, and Y. J. Wang, "Enhanced LTE device-to-device proximity services," *IEEE Communications Magazine*, vol. 54, no. 12, pp. 174–182, 2016.
- [3] L. Lei, Z. D. Zhong, C. Lin, and X. M. Shen, "Operator controlled device-to-device communications in LTE-advanced networks," *IEEE Wireless Communications Magazine*, vol. 19, no. 3, pp. 96–104, 2012.
- [4] A. Asadi and V. Mancuso, "Network-assisted outband D2D-clustering in 5G cellular networks: theory and practice," *IEEE Transactions on Mobile Computing*, vol. 16, no. 8, pp. 2246–2259, 2017.
- [5] J. Hu, W. Heng, Y. Zhu, G. Wang, X. Li, and J. Wu, "Overlapping coalition formation games for joint interference management and resource allocation in D2D communications," *IEEE Access*, vol. 6, pp. 6341–6349, 2018.
- [6] H. Min, J. Lee, S. Park, and D. Hong, "Capacity enhancement using an interference limited area for device-to-device uplink underlaying cellular networks," *IEEE Transactions on Wireless Communications*, vol. 10, no. 12, pp. 3995–4000, 2011.
- [7] Z. Uykan and R. Jantti, "Transmission-order optimization for bidirectional device-to-device (D2D) communications underlaying cellular TDD networks—a graph theoretic approach," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 1, pp. 1–14, 2016.

- [8] L. L. Wei, R. Q. Hu, T. He, and Y. Qian, "Device-to-device(d2d) communications underlying MU-MIMO cellular networks," in *Proceedings of the IEEE Global Communications Conference (GLOBECOM '13)*, pp. 4902–4907, IEEE, Atlanta, Ga, USA, December 2013.
- [9] X. Li, W. Zhang, H. Zhang, and W. Li, "A combining call admission control and power control scheme for D2D communications underlying cellular networks," *China Communications*, vol. 13, no. 10, pp. 137–145, 2016.
- [10] H. Sun, Y. Xu, and R. Q. Hu, "A NOMA and MU-MIMO supported cellular network with underlaid D2D communications," in *Proceedings of the 2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*, pp. 1–5, Nanjing, China, May 2016.
- [11] Z. Ding, X. Lei, G. K. Karagiannidis, R. Schober, J. Yuan, and V. K. Bhargava, "A survey on non-orthogonal multiple access for 5G networks: research challenges and future trends," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 10, pp. 2181–2195, 2017.
- [12] N. Ye, X. Li, H. Yu, A. Wang, W. Liu, and X. Hou, "Deep learning aided grant-free noma towards reliable low-latency access in tactile internet of things," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 5, pp. 2995–3005, 2019.
- [13] J. An, K. Yang, J. Wu, N. Ye, S. Guo, and Z. Liao, "Achieving sustainable ultra-dense heterogeneous networks for 5G," *IEEE Communications Magazine*, vol. 55, no. 12, pp. 84–90, 2017.
- [14] N. Ye, A. Wang, X. Li, H. Yu, A. Li, and H. Jiang, "A random non-orthogonal multiple access scheme for mmTc," in *Proceedings of the 2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, pp. 1–6, June 2017.
- [15] K. Yang, N. Yang, N. Ye, M. Jia, Z. Gao, and R. Fan, "Non-orthogonal multiple access: achieving sustainable future radio access," *IEEE Communications Magazine*, vol. 57, no. 2, pp. 116–121, 2019.
- [16] S. M. R. Islam, N. Avazov, O. A. Dobre, and K.-S. Kwak, "Power-domain non-orthogonal multiple access (NOMA) in 5G systems: potentials and challenges," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 721–742, 2017.
- [17] N. Ye, A. Wang, X. Li, W. Liu, X. Hou, and H. Yu, "On constellation rotation of noma with sic receiver," *IEEE Communications Letters*, vol. 22, no. 3, pp. 514–517, 2018.
- [18] M. Hojeij, C. A. Nour, J. Farah, and C. Douillard, "Joint resource and power allocation technique for downlink power-domain non-orthogonal multiple access," in *Proceedings of the 2018 IEEE Conference on Antenna Measurements & Applications (CAMA)*, pp. 1–4, September 2018.
- [19] F. A. Rabee, K. Davaslioglu, and R. Gitlin, "The optimum received power levels of uplink non-orthogonal multiple access (NOMA) signals," in *Proceedings of the 18th IEEE Wireless and Microwave Technology Conference, WAMICON 2017*, pp. 1–4, USA, April 2017.
- [20] Y. Li and G. A. A. Baduge, "Noma-aided cell-free massive mimo systems," *IEEE Wireless Communications Letters*, vol. 7, pp. 950–953, 2018.
- [21] N. Ye, A. Wang, X. Li et al., "Rate-adaptive multiple access for uplink grant-free transmission," *Wireless Communications and Mobile Computing*, vol. 2018, Article ID 8978207, 21 pages, 2018.
- [22] N. Ye, H. Han, L. Zhao, and A.-H. Wang, "Uplink nonorthogonal multiple access technologies toward 5G: a survey," *Wireless Communications and Mobile Computing*, vol. 2018, Article ID 6187580, 26 pages, 2018.
- [23] Y. Liu, W.-J. Lu, S. Shi et al., "Performance analysis of a downlink cooperative noma network over nakagami-m fading channels," *IEEE Access*, vol. 6, pp. 53034–53043, 2018.
- [24] X. Wang, J. Wang, L. He, and J. Song, "Outage analysis for downlink noma with statistical channel state information," *IEEE Wireless Communications Letters*, vol. 7, no. 2, pp. 142–145, 2018.
- [25] A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bölcskei, "An overview of MIMO communications—a key to gigabit wireless," *Proceedings of the IEEE*, vol. 92, no. 2, pp. 198–217, 2004.
- [26] W. Cai, C. Chen, L. Bai, Y. Jin, and J. Choi, "User selection and power allocation schemes for downlink NOMA systems with imperfect CSI," in *Proceedings of the 2016 IEEE 84th Vehicular Technology Conference (VTC-Fall)*, pp. 1–5, Montreal, QC, Canada, September 2016.
- [27] Z. Ding and H. V. Poor, "Design of massive-MIMO-NOMA with limited feedback," *IEEE Signal Processing Letters*, vol. 23, no. 5, pp. 629–633, 2016.
- [28] Z. Ding, F. Adachi, and H. V. Poor, "The application of MIMO to non-orthogonal multiple access," *IEEE Transactions on Wireless Communications*, vol. 15, no. 1, pp. 537–552, 2016.
- [29] Q. Sun, S. Han, I. Chin-Lin, and Z. Pan, "On the ergodic capacity of MIMO NOMA systems," *IEEE Wireless Communications Letters*, vol. 4, no. 4, pp. 405–408, 2015.
- [30] J. Ding, J. Cai, and C. Yi, "An improved coalition game approach for MIMO-NOMA clustering integrating beamforming and power allocation," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 2, pp. 1672–1687, 2019.
- [31] J.-B. Kim, I.-H. Lee, and J. Lee, "Capacity scaling for D2D aided cooperative relaying systems using NOMA," *IEEE Wireless Communications Letters*, vol. 7, no. 1, pp. 42–45, 2018.
- [32] S. Papaioannou, G. Kalfas, C. Vagionas et al., "5G mm Wave Networks Leveraging Enhanced Fiber-Wireless Convergence for High-Density Environments: The 5G-PHOS Approach," in *Proceedings of the 2018 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, pp. 1–5, Valencia, Spain, June 2018.
- [33] S. A. Naqvi and S. A. Hassan, "Combining NOMA and mmWave technology for cellular communication," in *Proceedings of the 2016 IEEE 84th Vehicular Technology Conference (VTC-Fall)*, pp. 1–5, Montreal, QC, Canada, September 2016.
- [34] F. Rusek, D. Persson, B. K. Lau et al., "Scaling up MIMO: opportunities and challenges with very large arrays," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, 2013.
- [35] T. Bai, A. Alkhateeb, and R. W. Heath, "Coverage and capacity of millimeter-wave cellular networks," *IEEE Communications Magazine*, vol. 52, no. 9, pp. 70–77, 2014.
- [36] X. Gao, L. Dai, S. Han, I. Chih-Lin, and R. W. Heath, "Energy-efficient hybrid analog and digital precoding for MmWave MIMO systems with large antenna arrays," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 998–1009, 2016.
- [37] X. Gao, L. Dai, Z. Chen, Z. Wang, and Z. Zhang, "Near-optimal beam selection for beamspace mmwave massive mimo systems," *IEEE Communications Letters*, vol. 20, no. 5, pp. 1054–1057, 2016.
- [38] Z. Wang, M. Li, Q. Liu, and A. L. Swindlehurst, "Hybrid precoder and combiner design with low-resolution phase shifters in mmWave MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 2, pp. 256–269, 2018.

- [39] Y. Sun, Z. Ding, and X. Dai, "On the performance of downlink NOMA in multi-cell mmWave networks," *IEEE Communications Letters*, vol. 22, no. 11, pp. 2366–2369, 2018.
- [40] N. Deng and M. Haenggi, "A fine-grained analysis of millimeter-wave device-to-device networks," *IEEE Transactions on Communications*, vol. 65, no. 11, pp. 4940–4954, 2017.
- [41] D. Zhang, Z. Zhou, C. Xu, Y. Zhang, J. Rodriguez, and T. Sato, "Capacity analysis of NOMA with mmWave massive MIMO systems," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 7, pp. 1606–1618, 2017.
- [42] S. Singh, M. N. Kulkarni, A. Ghosh, and J. G. Andrews, "Tractable model for rate in self-backhauled millimeter wave cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 10, pp. 2191–2211, 2015.
- [43] X. Yan, H. Xiao, C.-X. Wang, and K. An, "On the ergodic capacity of NOMA-based cognitive hybrid satellite terrestrial networks," in *Proceedings of the 2017 IEEE/CIC International Conference on Communications in China, ICCIC 2017*, pp. 1–5, China, October 2017.
- [44] Y. Huang, F. Al-Qahtani, C. Zhong, Q. Wu, J. Wang, and H. Alnuweiri, "Performance analysis of multiuser multiple antenna relaying networks with co-channel interference and feedback delay," *IEEE Transactions on Communications*, vol. 62, no. 1, pp. 59–73, 2014.