

Research Article

A Survey on the AI and Spectrum Management for Cache-Enabled Internet of Things in Smart Cities

Liming Chen ¹, Xiaoyun Kuang,¹ Fusheng Zhu ², Lijia Lai ³, and David Fan ³

¹Electric Power Research Institute of China Southern Power Grid, Guangzhou, China

²Guangdong New Generation Communication and Network Innovative Institute (GDCNi), Guangzhou, China

³School of Computer Science, Guangzhou University, Guangzhou, China

Correspondence should be addressed to Liming Chen; chenlm1018@foxmail.com, Fusheng Zhu; zhufusheng@gdcni.cn, and David Fan; davfan@126.com

Received 20 May 2021; Accepted 16 June 2021; Published 16 July 2021

Academic Editor: Xin Liu

Copyright © 2021 Liming Chen 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.

In this paper, we provide a comprehensive survey for the artificial intelligence and spectrum management, which are used for cache-enabled Internet of Things (IoT) in smart cities. In smart cities, there emerge a lot of new applications such as data collection and communication, environment monitoring, and real-time processing, which cannot be supported by the conventional wireless transmission techniques. Hence, some new wireless transmission techniques should be developed to support the emerging applications in smart cities. In this survey, we focus on the artificial intelligence, spectrum management, and caching techniques, where the interference arises due to the limited spectrum resources. In particular, we first review the current research status of these new techniques and, then, give some challenges on the system design. We further provide several feasible solutions on these challenges, in order to implement the IoT networks in smart cities. Finally, we conclude the work in the part of conclusions and give some discussions on the future works.

1. Introduction

The deployment and promotion of 5G mobile communication systems have promoted the development of mobile communication and related industries [1–4]. For the further development of mobile communication, researchers have proposed various new wireless transmission technologies such as massive MIMO, millimeter wave, and nonorthogonal multiple access [5–8]. They have also studied network technologies such as network slicing and edge computing [9–11]. These technologies can support a lot of services, such as ultrahigh speed, ultralow latency, and massive connections. However, the existing network communication technology, especially the underlying transmission technology, is still unable to meet the ever-increasing demand for communication services [12–14]. Therefore, it is of vital importance to design a wireless transmission technology that matches the rapidly developing communication network and to deeply exploit the wireless transmission resources of the

system. As a result, various performances including wireless transmission rate, latency, and energy efficiency are improved, and various application requirements and technical requirements of future communication networks can be met [15–17].

Due to the declining storage cost and the increasing storage performance, the network nodes can be equipped with storage resources. The introduction of the cache into the nodes in the networks can help perform data transmission and data storage. The nodes can flexibly adjust the transmit and receive data rate on the basis of cache according to the instantaneous state of the wireless channel [18]. This effectively expands the resource dimension of wireless transmission in the cache-enabled networks. For example, this enables wireless transmission to have both cooperation gain and content distribution gain. This is conducive to the improvement of network wireless transmission. Therefore, in order to effectively tap the potentials of the cache and promote the development of mobile communications and

related industries, the research of cache-enabled networks and the design of wireless collaborative transmission schemes matching the cache are all necessary.

In application scenarios such as 5G/B5G massive connections, the rapidly increasing number of network nodes and data traffic are conflicted with the limited wireless spectrum resources in the system, resulting in multiuser interference in the network. In the cache-enabled networks, multiple characteristics such as the number of interfering sources and transmit power are all affected by the cache, and the interference pattern is affected by the specific communication services and requirements. Existing researches have shown that interfering signal formed by the superimposition of multiple users in the wireless networks often exhibit a certain degree of heavy-tailed characteristics that deviate from the Gaussian distribution, which not only seriously affects the performance of the system wireless transmission but also limits the application and development of cache-enabled networks. Therefore, modeling the statistical distribution of multiuser interference in the cache-enabled networks can help optimize the wireless cooperative transmission of the design system.

This survey is based on the cache-enabled networks under multiuser interference and focuses on key theories and methods for wireless cooperative transmission. Wireless transmission in the cache-enabled networks is severely deteriorated by the interference. The specific services of the communication and the cache affect the shape of the interference in the system, and the interference thus shows a certain degree of heavy trailing characteristics deviating from the Gaussian distribution, which severely limits the advantage of the cache. At present, the internal mechanism of multiuser interference affecting the wireless transmission of the cache-enabled networks is not clear. It is a big challenge about how to design a wireless transmission theory and method compatible with the cache-enabled networks under interference, in order to give full play to the potentials of the cache to improve wireless transmission is a huge challenge. In view of this, the survey will explore the internal mechanism of multiuser interference affecting the wireless transmission of cache-enabled networks and establish a theory and method of wireless cooperative transmission that can effectively suppress interference and adapt to the cache on this basis, so as to give full play to the potentials of cache and greatly improve the system performance of the wireless transmission. The research results of the survey provide important theoretical and technical support for the optimal design of the cache-enabled networks, which provides an important reference for the development of future communication networks.

2. Recent Progress and Challenges

This survey comprehensively investigates the impact of interference and edge caching and studies the theory and method of wireless cooperative transmission of cache-enabled networks under multiuser interference. In the following, the application scenarios and performance requirements of future mobile communication networks, the current research status of the cache-enabled networks, and the research prog-

ress of wireless transmission under interference environments are described in order.

2.1. Wireless Transmission Technology. Compared with the traditional mobile communication networks, 5G/B5G and even future communication networks pay more attention to user experience. The core goals include not only classic technologies such as point-to-point physical-layer transmission and channel coding and decoding but also multipoint, multiuser, multiantenna, and multicell cooperative networking technologies, which strive to improve the system performance from the system architecture. According to Figure 1, there are three typical 5G application scenarios including eMBB (enhanced mobile broadband), URLLC (ultralow latency and high reliability), and mMTC (massive large connections). Among them, eMBB is mainly to increase the transmission rate, requiring to support a downlink rate of 20 Gbps and an uplink rate of 10 Gbps to cope with high-traffic applications such as 4 K/8 K ultrahigh-definition video and VR/AR; URLLC requires that the transmission latency of 5G must be less than 1 ms, in order to deal with low-latency applications such as unmanned driving and smart factories; the mMTC scenario is a large number of connections, which can meet the Internet of Things connection requirements such as more than one million access nodes per square kilometer.

The future communication network must have a new breakthrough in the wireless transmission technology to meet the application requirements of the three major scenarios. Therefore, the assessment of the satisfaction of the transmission technology to the specific requirements of practical communication services should be carried out systematically in combination with multiple performance metrics such as the transmission rate, latency, and energy efficiency, as follows.

2.1.1. Wireless Transmission Rate. Wireless transmission must be able to provide a certain rate of transmission to meet the basic needs of the Internet of Everything. In the application scenario of 5G-based Internet of Things, protocols such as IEEE 802.11ad (i.e., WiGig) can be applied to analyze the throughput performance of the system rate through analytical model and simulation derivation. For the 5G-based backhaul millimeter-wave network, the rate of each hop link can be reasonably allocated according to the needs of system transmission, combined with technologies such as path selection. In the communication systems, such as IEEE 802.11-based wireless local area networks, heterogeneous networks, and multicarrier NOMA, the system parameters can be optimized to maximize the rate of the system to significantly improve the connection performance of the system.

2.1.2. Wireless Transmission Latency. In the latency-sensitive communication application scenarios such as real-time traffic monitoring and Internet of Vehicles, the wireless transmission also needs to meet the key requirement of low latency. Queue theory and coding theory can be applied, and forward error correction codes can be used and optimized to achieve an optimized compromise between

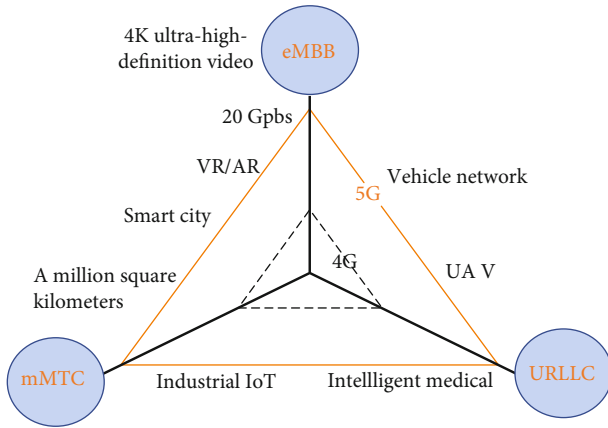


FIGURE 1: Typical application scenarios of 5G/B5G mobile communications.

the latency and throughput, in order to meet the requirement of end-to-end ultralow latency in 5G networks. In the wireless charging network, the transmission latency of the system data package can be analyzed, which serves as a basis for the system enhancement by minimizing the transmission latency of the wireless transmission data. In the communication system of large-scale wireless access and related tasks, this can significantly reduce the wireless transmission latency by optimizing the wireless transmission packet length, the maximum latency of the packet, and the task offloading strategy.

2.1.3. Wireless Transmission Energy Efficiency. Due to the limited energy at the communication nodes, the wireless transmission energy efficiency should meet the demand to extend the life of the node effectively. In the NOMA communication system, the energy efficiency of the system can be significantly improved by optimizing the system's wireless signal beamforming or optimizing the management of resources such as system power allocation and user clustering. Aiming at the Internet of Things communication system based on wireless local area network, Pirayesh et al. gave an optimization scheme for the system energy efficiency by tapping the potentials of Wifi network architecture. In the 5G heterogeneous cellular network, the authors in [19] applied technologies such as millimeter-wave and drone assistance, combined with user QoS and power constraints, and gave a unified optimization framework for the system energy efficiency. In the study of wireless transmission latency and energy efficiency, researchers have proposed a new joint optimization framework for mobile edge computing networks, starting from transmit power, wireless transmission bandwidth, and offloading strategies, which significantly reduced the system latency and energy consumption.

Existing challenges: as discussed in the above literature review, the existing 5G communication network should meet the ever-increasing requirement of various application services. However, with the rapid growth of communication network nodes and data traffic in the future, the wireless transmission of the system is still facing enormous challenges. The key to improve the wireless transmission technology is to dig deep into the system's various wireless resources

such as space, time, frequency, and power to improve the performance of the wireless transmission. At present, the limited wireless communication resources in the network are becoming marginal, which has become a bottleneck for improving the wireless performance of the system. It is urgent to find new dimensions of communication resources, to open up new ways to improve network wireless transmission, and to strongly support the rapid development of mobile communications and related industries.

2.2. Caching Technology. With the continued decline in storage costs and the increasing performance, it is possible to allocate storage resources on the nodes of the 5G/B5G network. Meanwhile, with the explosive growth of real-time video surveillance and various real-time computational tasks, it is difficult for the traditional wireless transmission architecture to support the increasing services. To this end, researchers have proposed a network system with cache as an important core and mobile edge computing as the architecture. As is shown in Figure 2, the introduction of cache can allow the computational tasks to sink to the edge nodes of the network in advance, which makes it possible to implement ultralow latency and ultrareliable wireless transmission and computing, achieving task offloading as soon as possible. In addition to being applied to mobile edge computing networks, caching can also be applied to architecture networks such as wireless cellular networks and wireless Mesh networks. Related researches on the cache-enabled networks have become a hot topic in the academic and industrial fields recently.

For the edge caching networks, the most significant issue is to construct a wireless transmission study of the theoretical caching framework, as a basis to optimize the design of the wireless transmission system architecture. In this area, Maddah-Ali et al. studied a framework for the communication with caching, through proposing a new caching strategy to optimize the local and global cache gain while significantly enhance the performance of the system. Based on Ali's work, researchers further studied the impact of cache on the wireless transmission mechanism and resource scheduling to fully dig the network's cache resources to enhance the system performance. Specifically, in view of the impact of cache-assisted wireless transmission rate, the authors in [20, 21] considered the cloud wireless access system with cache and studied the impact of cache on the system wireless transmission scheme, from the perspective of signal beamforming in the airspace, and also analyzed the wireless transmission performance of the system under the strategies of probabilistic caching and popularity caching. In view of the impact on the wireless transmission latency in the cache-enabled networks, researchers put forward a normalized performance metric based on the transmission latency and the transmission rate. From this, the optimized caching strategy is used to characterize the trade-off and replacement relationship between the cache and latency. Moreover, the existence of cache also has a profound impact on the wireless transmission rate, latency, energy consumption, and other performances in multirelay cooperative cache-enabled networks. For this system, the transmission rate queue and the energy

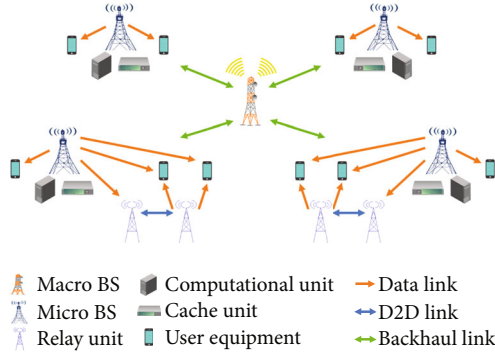


FIGURE 2: Application diagram of the cache-enabled networks.

consumption queue can be optimized to fully tap the potentials of the cache to improve the overall performance of the system wireless transmission. In further, the transmission rate of queue transmission and energy consumption queue can be optimized to fully tap the potentials of the cache to improve the overall performance of the system's wireless transmission.

Another research focus of edge caching networks is the design of caching strategies of the file system, that is, how to cache popular files on the network on limited storage resources. The file caching strategy affects both the content distribution gain and the signal cooperation gain. It is necessary to comprehensively consider the effects of these two gains to design a proper file caching strategy. The two traditional caching strategies are the largest content diversity (LCD) and most popular content (MPC), which can, respectively, maximize the system content distribution gain and signal cooperation gain. Researchers designed a hybrid caching strategy, based on LCD and MPC, for edge caching networks to optimize wireless transmission performance, and achieved a good compromise between content distribution gain and signal cooperation gain. Aiming at the multirelay-assisted cache-enabled networks in the Nakagami-m fading scenario, the literature [22] gives a joint framework for system wireless transmission performance evaluation and caching strategy optimization and optimizes popular files by minimizing the transmission outage probability, which was significantly better than the traditional LCD and MPC caching strategies. In the optimization design of the caching strategy, an intelligent learning algorithm based on the random sampling can be also used for the wireless cooperative transmission of the multirelay assisted cache-enabled networks, which can optimize the system confidential file caching strategy and obtain a significantly better wireless cooperative transmission performance than the traditional LCD, MPC, and other caching strategies. In addition, it is difficult for wireless networks where nodes arrive randomly to apply for the traditional LCD, MPC, and other deterministic file caching strategies.

Existing challenges: as discussed in the above literature review, the existence of cache in the cache-enabled networks has significantly expanded the dimensions of the system communication resources, giving a profound impact on the system performance and wireless transmission design. The system design of the wireless transmission should compre-

hensively consider the content distribution gain provided by the cache and strive to achieve a good compromise between the content distribution gain and the signal cooperation gain, so as to maximize the potentials of the cache and improve wireless transmission performance. Therefore, the research of cache-enabled networks needs to combine the specific services and the system architecture and in-depth consideration of the practical environment of wireless transmission, so as to analyze the impact of cache on the system wireless transmission to design the associated file caching strategy to give full play to the content distribution gain provided by the cache and the signal cooperation gain.

2.3. Impact of Interference on the Wireless Transmission.

With the rapid growth of the industrial scale and data traffic in the wireless networks, the number of wireless access nodes in the system is increasing. We take the large-scale Internet of Things with massively connected applications as an example, where the Internet of Things needs to provide connection services for more than one million access nodes per square kilometer in the future. However, the system wireless spectrum resources are limited, and interference in the network is unavoidable, which has become a bottleneck restricting the wireless transmission performance of the system. The interference of statistical distribution is closely related to many parameters such as the specific network architecture and the characteristics of communication services. Regarding the interfering signal formed by the superposition of multiple nodes, the existing research often assumes that the various characteristics of the interference node are the same. Based on this and the law of large numbers, it is concluded that the statistical characteristics of the interference obey the Gaussian distribution. However, this assumption focuses on the central distribution, ignoring or simplifying the tail characteristics of interfering signal. In practical communication scenarios, the interference node may reach the network according to the statistical law of the Poisson point process. And the distribution characteristics of the interference node itself are no longer the same, for example, the time and power to reach the network are no longer the same. These characteristics profoundly affect the statistical distribution of interference, causing the interference to exhibit heavy-tailed characteristics in a certain, which deviates from the Gaussian distribution. To this end, researchers should combine the specific network architecture with specific operational characteristics, by utilizing the heavy-tailed distributions, such as the Gaussian mixture model and Cauchy distribution, to characterize the distribution characteristics of the network interference more accurately.

The interference has seriously affected the transmission performance of the wireless network and has become a bottleneck restricting the wireless transmission of the system. Specifically, the interference seriously affects the performance of wireless network transmission such as rate, latency, and energy efficiency, detailed as follows.

2.3.1. The Impact of Interference on Wireless Transmission Rate.

Researchers have studied the transmission rate of the wireless network in the interference environments, analyzed

the impact of the interference on the system transmission performance, and gave an optimization framework for the system transmission rate. In this area, for the NOMA system, the interference between far and near users also has an adverse effect on the transmission rate of the system. The wireless transmission performance of the system can be improved through the allocation of the transmit power and rate. In addition, in the scenario of multiple interfering users, the power allocation among interfering users will also affect the wireless transmission rate of the system. In the research on the impact of interference on the wireless transmission rate of the network, the researchers employed an iterative detection scheme based on a deep convolutional neural network and maximum likelihood detector for the multi-antenna transmission system in the correlated interference environments, which improved the system transmission rate effectively.

2.3.2. The Impact of Interference on Wireless Transmission Latency. Regarding the multiuser interference, some techniques such as interference alignment can be used to improve the degree of freedom of network, from which the wireless transmission rate-latency tradeoff can be evaluated. In the multihop relay network under the interference environments, Alvandi et al. proposed to optimize the latency of wireless transmission through network coding and interference cancellation techniques. Moreover, for the downlink NOMA communication system, under the premise of satisfying the statistical latency of wireless transmission, the system performance can be characterized by the effective channel capacity, which started to reveal the comprehensive impact of the interference between NOMA users on the system latency and rate.

2.3.3. The Impact of Interference on Wireless Transmission Energy Efficiency. In the fog access network, the authors in [23] analyzed the impact of interference on the system energy efficiency of wireless transmission, and based on this, an optimization framework of system power resources was provided from the perspective of interference management and control to improve the energy efficiency of the network. Similarly, for the cognitive networks with limited interference, the system energy efficiency of wireless transmission can be improved by optimizing the power resources of the system. In the uplink and downlink NOMA system, the energy efficiency of the uplink and downlink transmission can be improved in combination with the mixed Gaussian interference of the system, and the internal mechanism of interference affecting the energy efficiency of the NOMA system can be revealed.

2.4. Artificial Intelligence for IoT Networks. Although the conventional optimization and resource management have enhanced the communication quality significantly, the communication resources in the network have not been exploited fully. Moreover, the system optimization often involves many complicated operations, causing that the conventional optimization schemes fail to work. To solve these problems, artificial intelligence-based algorithms have been proposed

to exploit the system resources of the communication networks. In this area, the authors in [24, 25] considered the correlated interference and utilized the deep convolutional neural network (DCNN) to improve the signal detection performance. For the non-Gaussian noise, a deep learning-based algorithm was proposed to approximate the probability density function of noise under impulsive noise environments [26]. For the mobile edge computing networks, deep Q-learning networks could be used to optimize the system offloading as well as the power allocation and bandwidth allocation [27–29]. Moreover, the federated learning framework could be used to optimize the system design in a distributed manner, which can help protect the privacy of users' data [30]. In further, the transfer learning can be used to optimize the offloading strategy for the MEC networks, when the system operates in a dynamic condition [31–33].

From the above analysis of the current research, we can see that the key point to the research on wireless transmission of communication networks is to dig the various system resources such as space, time, frequency, and power [34–36]. The introduction of storage resources in the cache-enabled networks expands the dimensions of traditional communication resources, creating favorable conditions for significantly improving and enhancing the system's wireless transmission performance [37, 38]. However, with the rapid growth of network nodes and data traffic, the increasing scarcity of wireless spectrum resources makes it hard to avoid multiuser interference in the system. The existence of the cache and the specific characteristics of the communication service have profoundly affected the statistical distribution of interference, causing the interference to deviate from the Gaussian distribution, showing a heavy-tail distribution characteristic in a certain [39–41]. The existence of interference seriously affects the performance of the system's wireless transmission of rate, latency, and energy efficiency, becoming a bottleneck restricting the wireless transmission of the cache-enabled networks. At present, the internal mechanism of interference affecting the wireless transmission of the cache-enabled networks is still unclear. It is still a big challenge to design an effective wireless cooperative transmission mechanism for the cache-enabled networks in the interference environments to fully tap the potentials of the cache to improve system performance.

This survey takes cache-enabled networks as the research object, focusing on the impact of multiuser interference, analyzing the mechanism of system wireless transmission, and designing the associated wireless cooperative transmission mechanism. Firstly, we analyze the impact of caching and communication-specific services on the interference distribution characteristics, using statistical methods and data fitting to characterize the center and tail distributions of interference. Secondly, under multiuser interference, we analyze the replacement relationship between the cache and the wireless transmission rate, latency, and energy efficiency of the system to reveal the internal mechanism of interference affecting the wireless transmission of the cache-enabled networks. Thirdly, we optimize the system design of the cache-enabled networks by devising the wireless cooperative transmission in the spatiotemporal dimensions of the cache and

signal transmission, which can fully tap the cache resources to suppress the interference. Finally, we build a complete wireless cooperative transmission framework for the cache-enabled networks. The research results provide a significant theory for the development of cache-enabled networks and provide an important reference for the development of future communication networks.

3. Feasible Solutions on the Challenges

3.1. Statistical Characteristics of Multiuser Interference in Cache-Enabled Networks. Firstly, this survey takes nonorthogonal multiple access (NOMA) and other large-scale Internet of Things as typical application scenarios to collect the data of multiuser interference, from two aspects of computer simulation and IoT platform test [42, 43]. The computer simulation starts from the statistical distribution and uses the Monte Carlo method to simulate the instantaneous parameter of the interference. The practical measurement of the IoT platform was carried out in a joint IoT laboratory. The laboratory is equipped with a USRP platform, radio frequency antennas, various signal generators, and other equipments. By building a wireless transmission platform with a cache architecture, and combining the specific needs of different scenarios, we obtain the multiuser interference data in multiple connections from the network.

Assume that there are M interfering users in the wireless transmission system of the cache-enabled networks, where the transmission signal of the m th interfering source is $x_{m,I}$, the transmit power is $P_{m,I}$, and the channel parameter between the interference and the target user is $h_{m,I}$. Combined with specific communication application scenarios, the interfering signal of the M interferers is

$$I = \sum_{m=1}^M \sqrt{P_{m,I}} h_{m,I} x_{m,I}. \quad (1)$$

It can be seen from (1) that the form of interference is closely related to the number of interfering sources M , channel parameters $h_{m,I}$, transmit power $P_{m,I}$, and other parameters. The different requirements of cache and practical communication services have a profound impact on the interference parameters. Specifically, it is possible that the network nodes store the content of different users in advance because of the introduction of cache. In particular, the popularity of the stored files will profoundly change the number of interfering source M during transmission and significantly affect the shape distribution of interference. At the same time, the introduction of cache makes nearby communication become possible. Under the premise of ensuring transmission performance, the transmit power $P_{m,I}$ can be appropriately reduced to improve the energy efficiency of network nodes, which will inevitably have a nonnegligible impact on the distribution of multiuser interference. In addition, different services of practical communication have different requirements for transmission rate, latency, and energy efficiency, which will inevitably affect the transmit

power $P_{m,I}$ of interference and change the distribution of complicated interference in the system. Moreover, the different locations of network nodes, such as near or far users of NOMA systems, will inevitably affect the channel parameters $h_{m,I}$ of interference, which can change the form of interference.

To sum up, cache and different requirements of practical communication services have brought profound changes to various parameters of multiuser interference, which significantly affects the form of interference in the network. When the user arrives at the network with the statistical law of the Poisson point process, the existing research shows that the statistical distribution of interference presents a certain heavy-tailed property, which invalidates the traditional Gaussian distribution model. Consequently, it is urgent to model the statistical distribution of complicated interference in the cache-enabled networks, which lays the foundation for systematical analysis and system design of the cache-enabled networks.

Secondly, statistical modeling is performed on the heavy-tailed multiuser interference in the cache-enabled networks. After collecting the interference data in the cache-enabled networks, it is necessary to analyze the center distribution and tail distribution of the interference, which model the complicated interference data in the cache-enabled networks reasonably and effectively. The tail characteristic of interference can be characterized by the associated characteristic function $\psi_I(w)$ in the statistical distribution $f_I(x)$. A typical form of $\psi_I(w)$ was proposed by Swami and Haenggi, where the expression is given by

$$\psi_I(w) = e^{-\beta_I^{\alpha_I} |w|^{\alpha_I}}, \quad (2)$$

where $\alpha_I \in [1, 2]$ is the characteristic coefficient of the interference. In particular, $\alpha_I = 1$ corresponds to the Cauchy distribution, while $\alpha_I = 2$ corresponds to the Gaussian distribution. Notation $\beta_I > 0$ is the scale coefficient of the interference. Starting from $\psi_I(w)$, the statistical distribution $f_I(x)$ of the interference can be calculated by the inverse Laplace transform and compared with the distribution presented by the practical data. When $\beta_I = 1$, the calculation result of $f_I(x)$ under different characteristic coefficients α_I are shown in Figure 3.

As shown in Figure 4, as the interference characteristic coefficient α_I becomes smaller, the tail distribution of interference becomes more obvious, which significantly deviates from the Gaussian distribution associated to $\alpha_I = 2$. Since the central distribution of the interference shows the concentrated characteristics of the Gaussian distribution, while the tail distribution shows a certain tail property, where the mixed form of the Gaussian distribution and the typical tail property can be used to characterize the statistical distribution of multiuser interference in the cache-enabled networks, which is a reasonable and feasible choice. We use $f_{\text{Guass}}(x)$ to represent the Gaussian distribution and $f_{\text{Tail}}(x)$ to represent the tail distribution. The specific form of $f_{\text{Guass}}(x)$ is

$$f_{\text{Gauss}}(x) = \frac{1}{\sqrt{2\pi}\beta_I} e^{-\frac{x^2}{2\beta_I^2}}. \quad (3)$$

Note that $f_{\text{Tail}}(x)$ can be characterized by using the typical heavy-tailed distributions such as Cauchy distribution $f_{\text{Cauchy}}(x)$ and Laplace distribution $f_{\text{Lap}}(x)$. The specific expressions of $f_{\text{Cauchy}}(x)$ and $f_{\text{Lap}}(x)$ are as follows

$$f_{\text{Cauchy}}(x) = \frac{\beta_I}{\pi(x^2 + \beta_I^2)}, \quad (4)$$

$$f_{\text{Lap}}(x) = \frac{1}{2\beta_I} e^{-\frac{|x|}{\beta_I}}. \quad (5)$$

From (3)–(5), a linear combination of Gaussian distribution and tail distribution can be used to approximate the complicated interfering signal in the cache-enabled networks

$$f_I(x) = \varepsilon f_{\text{Tail}}(x) + (1 - \varepsilon) f_{\text{Gauss}}(x), \quad (6)$$

where $\varepsilon \in [0, 1]$ is the mixing coefficient, which reflects the degree of the tail in the interference distribution. In particular, $\varepsilon = 0$ means that the interference obeys Gaussian distribution without a serious tail, while $\varepsilon = 1$ means that the interference tail is serious, and its distribution is mainly characterized by the heavy-tailed distribution. From (6), we take the Laplace transform on both sides to obtain the approximate expression of the characteristic domain $\psi_I(w)$. By minimizing the approximation error in the characteristic domain w within the range of the characteristic domain, we can obtain the estimation of the mixing coefficient ε . Finally, an approximate expression of the statistical distribution of multiuser interference in the cache-enabled networks is realized.

The impact of cache and communication services on the statistical distribution of interference can be evaluated. From the analytical expression of interference in (6), it can be seen that the characteristic parameter combination of interference distribution is $(\alpha_I; \beta_I, \varepsilon)$, where α_I and ε reflect the tail degree of interference, and β_I determines the average intensity of interference. The characteristics of cache and multiple connections with different services have a profound impact on system interference, which significantly affects the interference characteristic combination parameter $(\alpha_I; \beta_I, \varepsilon)$. This survey will start from the cache and various services characteristic parameters of multiple connections and analyze the impact on the statistical distribution of interference. On one hand, in view of the impact of cache, we analyze the impact of the characteristic parameters of the cache on the interference distribution, including the number of cache nodes N_I of interfering files, the cache capacity K_I of interference nodes, and the popularity coefficient of interfering files ρ_I . Qualitatively, the larger the cache capacity K_I of the interference node, the more the interfering files can be cached. The interfering files can be communicated nearby during the transmission, and there is no need to use a high transmit

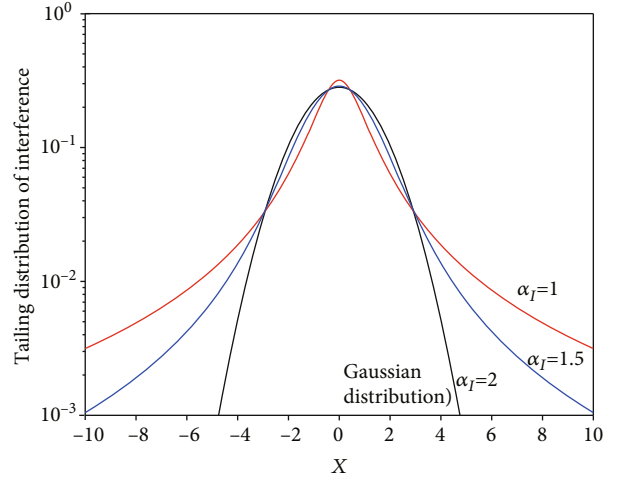


FIGURE 3: The statistical distribution of several typical interferences.

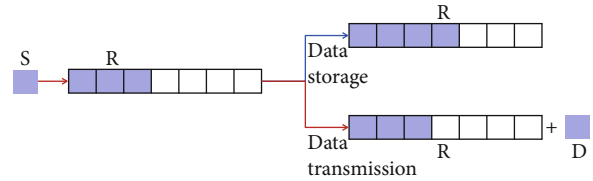


FIGURE 4: Cache storage and transmission.

power to transmit files, which will reduce the average intensity β_I of interference. In addition, the higher the popularity coefficient ρ_I of the interference file, the greater the probability that the network node will store the file. Based on this, the relationship will be established between the interference characteristic parameter $(\alpha_I; \beta_I, \varepsilon)$ and the cache parameters (N_I, K_I, ρ_I) , where a multidimensional changing surface to describe the impact of the cache will be drawn on the interference statistical distribution. On the other hand, we further analyze the impact of its characteristics on the statistical distribution of interference for different services of multiple connections in practical communications. According to the different specific application scenarios of 5G/B5G, it is necessary to support the service transmission of multiple connections with different characteristics. For example, the video surveillance service uses a long packet transmission and imposes high requirements on the energy consumption of the node, while the transmission of control signals uses a short packet transmission and imposes a strict requirement on the latency. In a multiuser system based on NOMA, users far and near can correspond to different transmission services, and the signal transmit power and energy efficiency between users will also be different. These different services characteristics will profoundly change the statistical characteristics of system interference and significantly affect the interference characteristic parameter $(\alpha_I; \beta_I, \varepsilon)$. Two typical application scenarios will be selected from a long packet and short packet transmission, respectively. The impact of multiple connections of different services will be analyzed on the interference characteristic parameter $(\alpha_I; \beta_I, \varepsilon)$, where we draw a multidimensional change surface to

describe the statistical distribution of interference between multiple connections and different services.

3.2. Wireless Transmission Performance of Cache-Enabled Networks in Interference Environments. This survey studied the wireless transmission capacity and outage probability of the system in the interference environments. In the cache-enabled networks, the multiuser interference of the system exhibits a certain tail characteristic, which deviates from the Gaussian distribution. At this time, the capacity analysis of the communication system under traditional Gaussian noise is difficult to directly apply, and it is necessary to design a performance evaluation method such as a transmission rate suitable for the interference. When the interfering signal is I , we use x_s to represent the transmission signal of the network wireless node, and z is used to represent the associated received signal. For a fading-free communication system with $z = x_s + I$, starting from the entropy theory of information theory, the transmission capacity C of the communication system under interference can be obtained as

$$C = \max_{f_{x_s}(x)} \text{Entr}(z) - \text{Entr}(I), \quad (7)$$

where $\text{Entr}(z) = - \int f_z(z) \log_2 f_z(z) dz$ is the information entropy of the received signal z , and $\text{Entr}(I) = - \int f_I(I) \log_2 f_I(I) dI$ is the information entropy of the interference. Starting from the approximate expression of the interference statistical distribution in (6), the solution of $\text{Entr}(I)$ can be obtained. In the environment of Gaussian noise, the optimal expression form of the x_s distribution $f_{x_s}(x)$ is the Gaussian distribution. However, this result cannot be applied to the interference environments. Since the optimal expression form of the x_s distribution $f_{x_s}(x)$ is unknown, it is difficult to find an accurate value of $\text{Entr}(z)$, which brings difficulties to the accurate solution of the transmission capacity under interference. For this reason, we turn to analyze and solve the upper and lower bounds of C in order to measure C within a certain range as accurate as possible. The relevant research results of the existing information theory show that the upper and lower bounds of C can be expressed in many forms. A feasible upper and lower bound is

$$C^{\text{lb}} = C(I_0), \quad (8)$$

$$C^{\text{ub}} = C(I_0) + \text{Entr}_{I_0}(I), \quad (9)$$

where I_0 is the interference of Gaussian distribution and its variance is equal to the variance of I . Notation $C(I_0)$ represents the capacity of the communication system under Gaussian interference I_0 , and $\text{Entr}_{I_0}(I)$ is the entropy of the interference in the Gaussian interference space, expressed as $\int f_I(x) \log_2(f_I(x)/f_{I_0}(x)) dx$. From this result and equations (8) and (9), the upper and lower bounds of the channel capacity C can be obtained. In addition to (8) and (9), other forms can be integrated to characterize the upper and lower bounds of C to obtain a more compact upper and lower bounds of channel capacity. In this way, the characterization

of the transmission capacity of the communication system can be realized under multiuser interference.

The above analysis of transmission capacity is performed without channel fading. In a fading channel environment, it is also necessary to analyze the outage probability of the communication system from the abovementioned channel capacity to evaluate the transmission performance of the system more comprehensively. Considering the classic fading communication system $z = hx_s + I$, where h is the instantaneous parameter of the fading channel. Starting from the solution result of the abovementioned channel capacity C , the outage probability P_{out} of the wireless transmission can be analyzed for the cache-enabled networks under interference. We use a_S to indicate the transmission rate of the source node and use R_{th} to indicate the target transmission rate under interference, where the outage probability P_{out} is given by $P_{\text{out}} = \text{Prob}(a_S < R_{\text{th}})$. It can be seen that P_{out} characterizes the evaluation result of the system transmission performance over a long period of time.

From the above analytical expressions of transmission capacity and transmission outage probability, the quantitative impact of cache and interference on the system transmission performance will be analyzed. In further, communication scenarios such as high signal-to-interference ratio can be considered, and the asymptotic form of P_{out} can be analyzed and solved. Starting from the asymptotic expression, the main impact of the main parameters of the system on P_{out} is analyzed. In particular, we can study the cache-related parameters, including file cache node number N , node cache capacity K , file popularity coefficient ρ , and parameters $(\alpha_I, \beta_I, \epsilon)$ of system interference related. The main constraints of these important parameters are analyzed and evaluated on the system transmission performance, revealing the internal mechanism that affects the wireless transmission rate and outage performance in the cache-enabled networks.

We further analyze the impact of cache and interference on the wireless transmission latency of cache-enabled networks. The utilization of cache on the network nodes provides the possibility for the scheduling between the system data storage and transmission, which gives the possibility to improve the wireless transmission performance under interference and also brings transmission latency in a certain. In practical communication environments, the tolerance to transmission latency is also different according to different communication services. For example, the Internet of Vehicles, unmanned driving, and other services related to control have particularly high requirement on the latency. Therefore, it is of vital significance to analyze the wireless transmission latency of the cache-enabled networks under multiuser interference.

We take a relay-assisted double-hop relay network as an example. The source node S transmits its data to the destination D through the relay node R. Both R and D are affected by the interference, which severely restricts the performance of the system transmission. The time resource provided by the cache at the relay R makes it possible for the system to schedule the data storage and transmission. As shown in Figure 4, when the S-R link quality is good, R can choose to receive

data from S and store it in the cache. On the contrary, when the R-D link quality is good, R can choose to transmit the data in the cache to D.

Let γ_d and $\bar{\gamma}_d$ represent the instantaneous value and statistical average of the packet transmission latency, respectively. Starting from Little's law, the expression of $\bar{\gamma}_d$ can be obtained as

$$\bar{\gamma}_d = \frac{1}{R_{\text{eff}}} \bar{Q}, \quad (10)$$

where R_{eff} is the effective transmission rate of the R-D link under interference and \bar{Q} is the average length of the queue in the cache. From the expressions of capacity C and outage probability P_{out} under interference, we can write R_{eff} as $R_{\text{eff}} = R_{\text{RD}}(1 - P_{\text{out,RD}})$, where P_{RD} and $P_{\text{out,RD}}$ represent the wireless transmission rate and outage probability of the R-D link, respectively. Let $[Q(1), Q(2), \dots, Q(T)]$ denote the cache queue in time slot T , and the state of the cache queue at time t is q_i . Starting from the Markov state transition model, we can analyze the probability of queue state transition. Let π_i represent the probability of staying in the state q_i , and then, the average length \bar{Q} of the queue in the cache is

$$\bar{Q} = \sum_{i=1}^{L_s} \pi_i \text{Len}(q_i), \quad (11)$$

where L_s is the number of all possible states, and $\text{Len}(q_i)$ is the data length of the cache in state q_i .

It can be seen from (10) and (11) that cache and interference profoundly affect the wireless transmission latency of the cache-enabled networks. Specifically, the data storage space provided by the cache makes it possible for the system to schedule in the time dimension, which profoundly affects the length and status of the cache queue. At the same time, interference profoundly affects the reception and transmission performance of the relay node, which significantly affects the state transition of the cache queue. Therefore, both cache and interference have a profound impact on the system latency. Based on (10) and (11), we will measure the characteristic parameters of the cache (including the number of cache nodes of the file N , the cache capacity of the node K , and the file popularity coefficient ρ) and the characteristic parameter $(\alpha_I, \beta_I, \varepsilon)$ of interference when it is used for wireless transmission. The analytical expression of the associated relationship is established to quantitatively describe the impact of cache and interference on the latency. Further, draw the multidimensional change surface of the latency versus the cache characteristic parameters, and the interference characteristic parameters, and explore the internal mechanism of the mutual replacement of the cache and transmission latency under multiuser interference.

Finally, the energy efficiency of wireless transmission of cache-enabled networks is analyzed under multiuser interference. The wireless transmission energy consumption of nodes in the cache-enabled networks is the key factor that restricts the service life. Therefore, it is significant for the net-

work nodes to improve the wireless transmission efficiency to extend their service life as much as possible. We take the classical double-hop relay network S->R->D as an example and use $P_S(t)$ and $P_R(t)$ to represent the transmit power of the source node S and the relay R at time t , where $t \in [1, T]$. Due to the storage resource brought by the cache, the relay R can choose to store data or transmit data to the target node D. Let $\delta(t)$ represent the working status of the relay node R. Specifically, $\delta(t) = 0$ represents that the relay chooses to receive data from the source node S and store it in the cache, while $\delta(t) = 1$ represents that the relay R transmits the data to D. By ignoring the scarce conditions such as the cache is empty or full, the energy efficiency γ_{ee} of the wireless transmission of the cache-enabled networks in the interference environments can be expressed as

$$\gamma_{\text{ee}} = \frac{\sum_{t=1}^T \delta(t) a_R(t) (1 - P_{\text{out,R}})}{\sum_{t=1}^T \delta(t) P_R(t) + (1 - \delta(t)) P_S(t)}, \quad (12)$$

where $a_R(t)$ is the transmission rate of relay R under interference and $P_{\text{out,R}}$ is the associated outage probability. This work will start from the expression of γ_{ee} in (12) and analyze the impact of the characteristic parameters of the cache and the characteristic parameters of interference on the energy efficiency of the system. Specifically, N , K , and ρ are used to represent the number of cache nodes, cache capacity, and file popularity coefficient of the file, respectively; α_I , β_I , and ε are used to represent the characteristic parameters of interference, respectively. From a qualitative point of view, the existence of the cache makes it possible to communicate with each other nearby, which reduces the adverse effects of interference in a multihop network, so that the network nodes do not have to transmit data with a high power, which effectively improves the transmission energy efficiency γ_{ee} . In particular, when the number of file cache nodes N becomes larger, there are more nodes in the network which have prestored files to be transmitted, and the diversity gain provided by multinode cooperation can significantly reduce the demand on the transmit power. On the other hand, as the file popularity coefficient ρ becomes larger, the file popularity distribution becomes more concentrated, and some highly popular files will be stored by more nodes, and the associated transmission energy efficiency γ_{ee} will increase. Under multiuser interference, as the interference intensity β_I increases, the transmission quality of the channel will deteriorate, which will inevitably affect the $a_R(t)$ and $P_{\text{out,R}}$ in γ_{ee} . Similarly, as α_I becomes smaller or ε becomes larger, the tail degree of interference increases, causing $a_R(t)$ to become smaller or $P_{\text{out,R}}$ to become larger, which will inevitably have an adverse effect on the system transmission energy efficiency. We will start from (12) and try to establish the expression between the system transmission energy efficiency γ_{ee} and the characteristic parameter $(N; K; \rho)$ of the cache, as well as the characteristic parameter $(\alpha_I, \beta_I, \varepsilon)$ of the interference, so as to quantitatively analyze the impact of cache and interference on the system transmission energy efficiency. At the same time, we draw a multidimensional surface map of

γ_{ee} versus these characteristic parameters and discover the internal mechanism of cache and interference affecting the wireless transmission energy efficiency of cache-enabled networks.

3.3. Joint Optimization of Wireless Cooperative Transmission in Cache-Enabled Networks. This survey takes the double-hop relay network under the cache architecture as a typical system as an example to study how to optimize the design in the signal space. The research results can also be extended to other topological forms of cache-enabled networks. In the S->R->D double-hop relay network, the source node S transmits information to the target node D with the assistance of the relay node R. The relay node R is equipped with a cache, which can store the files of the source node S in advance, and makes it possible to enhance the wireless transmission of the cache-enabled networks. We use M_1, M_2, M_3 , to denote the number of antennas at the source node S, relay node R, and target node D, respectively. The spatial resources provided by multiple antennas and the time resources provided by the cache in the time domain can be used to improve the wireless transmission performance of the cache-enabled networks. For this reason, it has become an inevitable choice to use spatiotemporal transceivers at the network nodes to exploit the spatiotemporal resources of the system to improve the wireless transmission performance.

The selected spatiotemporal transceiver should be able to effectively tap the spatiotemporal characteristics of signal transmission and have a simple structure, which is easy to implement in practice. Based on the existing research results, the fractional transversal filter can be selected. Therefore, we will refer to the idea of the fractional transversal filters to design the spatiotemporal receivers at the network nodes. We denote the spatiotemporal transceiver coefficients at the network nodes S, R, and D by $W_1(t), W_2(t),$ and $W_3(t)$, respectively. From the analysis of the wireless transmission performance of the system under the aforementioned interference, combined with the spatiotemporal transceiver coefficients $W_1(t), W_2(t),$ and $W_3(t)$, the expressions of the system wireless transmission rate $a(t)$, latency γ_d , and energy efficiency γ_{ee} can be obtained. Then, by optimizing the comprehensive performance of the cache-enabled networks, we can solve $W_1(t), W_2(t), W_3(t)$, which lead to the design of the spatiotemporal transceiver of the network node is realized. Specifically, for multiobjective optimization problems with multiple performance metrics, one metric can be concentratedly optimized under the premise of ensuring other performance metrics. For example, under the premise of ensuring the transmission rate and latency, maximize the energy efficiency of the system:

$$\begin{aligned} & \max_{\{W_i(t)|i=1,2,3\}} \gamma_{ee} \\ & \text{s.t. } a(t) \geq a_{th} \\ & \quad \gamma_d < \gamma_d^{th}, \end{aligned} \quad (13)$$

where a_{th} and γ_d^{th} are the thresholds of the wireless transmission rate and transmission latency, respectively. Similarly, the

latency γ_d of wireless transmission can also be minimized under the premise of ensuring transmission rate and energy efficiency. We can also maximize the wireless transmission rate $a(t)$ under the premise of ensuring the energy efficiency and latency performance. We take (13) to maximize the transmission energy efficiency as an example. This optimization problem is essentially a time series optimization problem. For this type of time series optimization problem, the Lyapunov optimization framework can be used to solve it effectively. Therefore, we choose to use the Lyapunov optimization to solve the spatiotemporal transmission and reception coefficients $W_1(t), W_2(t),$ and $W_3(t)$ in equation (13). Specifically, we first establish a cache state queue $\{Q_B | 1 \leq t \leq T\}$, a channel state queue $\{Q_C | 1 \leq t \leq T\}$, and a node energy consumption queue $\{Q_E | 1 \leq t \leq T\}$. Then, we construct a joint queue of cache, channel, and energy consumption $\{Q_J | 1 \leq t \leq T\}$, where $Q_J(t)$ is

$$Q_J(t) = [Q_B(t), Q_C(t), Q_E(t)]. \quad (14)$$

From this equation, we can obtain the Lyapunov quadratic function $L(Q_J(t))$ and the Lyapunov conditional drift factor $\Delta(Q_J(t))$ as

$$L(Q_J(t)) = \frac{1}{2} Q_B^2(t) + \frac{\mu_1}{2} Q_C^2(t) + \frac{\mu_2}{2} Q_E^2(t), \quad (15)$$

$$\Delta(Q_J(t)) = E[L(Q_J(t+1)) - L(Q_J(t)) | L(Q_J(t))], \quad (16)$$

where μ_1 and μ_2 are two nonnegative control parameters used to adjust the weights of the three terms in (15). In (16), $E[\bullet]$ is the expected operation, that is, the random variables such as the channel state are statistically averaged under the current queue state $Q_J(t)$. The Lyapunov conditional drift factor $\Delta(Q_J(t))$ reflects the expected change of the Lyapunov function in a time slot and reflects the incremental characteristics of the joint queue state $Q_J(t)$. Starting from (15) and (16), and combining the target performance parameters in (13) such as transmission rate $a(t)$, latency γ_d , and energy efficiency γ_{ee} , the Lyapunov drift plus penalty function is established. In further, by minimizing the upper limit and combining convex optimization algorithms and other algorithms, the spatiotemporal transceiving coefficient $\{W_i(t) | i = 1, 2, 3\}$ can be finally solved, which realizes the design of the spatiotemporal transceiver of the wireless node for the cache-enabled networks.

We further optimize the file caching strategy by using the wireless cooperative transmission characteristics of the cache-enabled networks. It has been shown that caching has profoundly changed the storage of files on network nodes and has significantly affected the various performances of wireless transmission of edge caching networks under interference. We will design the caching strategy from multiple performance metrics such as the transmission rate, latency, and energy efficiency. The file caching strategy is mainly based on the file popularity and the transmission characteristics of network nodes to obtain an appropriate compromise

between the cooperation gain of wireless transmission and the cache gain of files, so as to optimize the overall performance of the cache-enabled networks wireless transmission. Two typical caching strategies include the most popular content (MPC) and largest content diversity (LCD), as shown in Figure 5. We can find from Figure 5 that MPC can achieve the best cooperation gain, while LCD can achieve the best content diversity gain.

Without loss of generality, we use the Zipf distribution to model the popularity of L_F files in the network. The popularity coefficient is ρ , and then, the popularity $f_p(k)$ of the k th file is

$$f_p(k) = \frac{k^{-\rho}}{\sum_{k_1=1}^{L_F} k_1^{-\rho}}. \quad (17)$$

In addition to the popularity $f_p(k)$, the statistical characteristics of the wireless transmission of the cache-enabled networks under interference should also be considered. We use $E\{a_n(t)\}$, $E\{\gamma_d(n)\}$, and $E\{\gamma_e(n)\}$ to represent the statistical value of the wireless transmission rate, latency, and energy efficiency of the network node n . Starting from $f_p(k)$, $E\{a_n(t)\}$, $E\{\gamma_d(n)\}$, $E\{\gamma_e(n)\}$, we can design the popularity metric $f_j(n, k)$ that can incorporate the wireless transmission statistical characteristics of cache-enabled networks under interference as

$$f_j(n, k) = g \left[f_p(k) E\{a_n(t)\}, E\{\gamma_d(n)\}, E\{\gamma_e(n)\} \right], \quad (18)$$

where $g[\bullet]$ is the function to be constructed and the output value is in the interval $[0, 1]$. In particular, a larger $f_p(k)$ indicates that the k th file is more popular and has a higher priority to be prestored on the network node. Similarly, a larger $E\{a_n(t)\}$ indicates that node n has a higher transmission capacity and can store more files. Further, a smaller $E\{\gamma_d(n)\}$ or larger $E\{\gamma_e(n)\}$ indicates that node n has better transfer characteristics and can transmit files more efficiently, which leads to a larger $f_j(n, k)$.

We further consider the comprehensive metric $f_j(n, k)$, which integrates the file popularity and wireless transmission characteristics to design a caching strategy for L_F files on N nodes of the network. Specifically, based on two typical caching strategies, MPC and LCD, L_F files are grouped and N nodes are clustered according to the value of $f_j(n, k)$ in order to obtain a suitable compromise between popularity and wireless transmission characteristics. We set two thresholds $f_{1,\text{th}}$, $f_{2,\text{th}}$, both on the interval $[0, 1]$ and satisfying $f_{1,\text{th}} > f_{2,\text{th}}$. Based on the threshold values, a feasible caching strategy can be proposed as follows:

- (i) When $f_j(n, k) \geq f_{1,\text{th}}$, the k th file will be stored on all nodes in the network
- (ii) When $f_j(n, k) \in [f_{2,\text{th}}, f_{1,\text{th}})$, the k th file will be cached on nodes N_1 in the network, where $N_1 < N$

		Caching space				
		C ₁	C ₂	C ₃	C ₄	C ₅
Relays	R ₁	1	2	3	4	5
	R ₂	1	2	3	4	5
	R ₃	1	2	3	4	5
	R ₄	1	2	3	4	5

		Caching space				
		C ₁	C ₂	C ₃	C ₄	C ₅
Relays	R ₁	1	5	9	13	17
	R ₂	2	6	10	14	18
	R ₃	3	7	11	15	19
	R ₄	4	8	12	16	20

FIGURE 5: Two typical caching strategies.

- (iii) When $f_j(n, k) < f_{2,\text{th}}$, the k th file will not be stored on the nodes of the network

In this strategy, $f_{1,\text{th}}$, $f_{2,\text{th}}$, and N_1 can be optimized and adjusted to dig deeper into the system transmission cooperation gain and cache content distribution gain and maximize the advantage of caches in improving the wireless transmission performance of cache-enabled networks.

In addition to the above strategy, other forms of caching strategies will be considered comprehensively. For example, it is difficult to apply the above strategy without knowing the prior popularity of files. In this case, the user request and the system wireless transmission characteristics can be estimated by using some machine learning algorithms, such as deep reinforcement learning, in order to obtain the popularity of network files and determine whether it is appropriate to store the files on the network nodes and make associated decisions. Through these caching strategies, we can maximize the gains brought by caching to effectively suppress the impact of interference on the wireless transmission of the edge caching network and significantly improve the comprehensive performance of the edge caching network in terms of the wireless coordinated transmission rate, latency, and energy efficiency.

Finally, the joint optimization of wireless coordinated transmission and caching strategy is carried out, and a new framework of wireless coordinated transmission for the edge caching networks under interference is constructed. As mentioned in the optimization process of wireless cooperative transmission in the cache-enabled networks, the cache profoundly changes the architecture of the wireless cooperative transmission of the system. On the other hand, the optimization of the cache must fully consider the statistical characteristics of wireless cooperative transmission, so that the advantages of the cache can be effectively used, to improve the comprehensive performance of the wireless cooperative transmission of the cache-enabled networks. The key to establish a new framework of wireless coordinated transmission for cache-enabled networks under multiuser interference is wireless coordinated transmission and caching strategies. Therefore, a joint optimization of wireless coordinated transmission and caching strategy is an inevitable choice of establishing a new framework of wireless coordinated transmission for the cache-enabled networks under multiuser interference.

As shown in Figure 6, wireless cooperative transmission is carried out under a caching architecture. The storage location and number of files on the network node can be determined by parameters $\{f_{1,\text{th}}, f_{2,\text{th}}, N_1\}$ and other parameters, which come from the caching optimization. Based on these data, the spatiotemporal transceiver for wireless cooperative

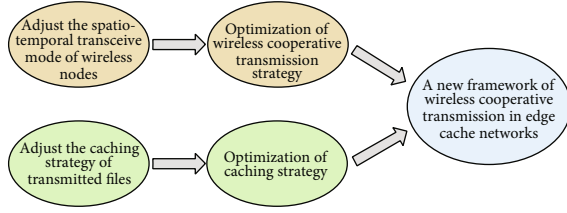


FIGURE 6: Joint optimization framework of wireless cooperative transmission and caching strategy.

transmission of the cache-enabled networks can be designed, and the associated spatiotemporal transceiver coefficient $\{W_i(t) | i = 1, 2, 3\}$ can be solved. On the other hand, the statistical characteristics of $\{a_n(t), \gamma_d(n), \gamma_{ee}(n)\}$ and other performances optimized by the wireless cooperative transmission strategy also affect the storage of files on the network node. The optimization of the caching strategy should fully consider the statistical characteristics of these performances and solve the associated caching parameters $\{f_{1,th}, f_{2,th}, N_1\}$. Finally, by jointly optimizing wireless cooperative transmission and caching strategies, the comprehensive performance of cache-enabled networks under interference has been greatly improved, and then, a new framework wireless cooperative transmission adapted to the cache has been established for the cache-enabled networks.

4. Conclusions

In this paper, we provided a comprehensive survey for the artificial intelligence and spectrum management, which were used for cache-enabled IoT networks in smart cities. In smart cities, there emerged a lot of new applications such as data collection and communication, environment monitoring, and real-time processing, which cannot be supported by the conventional wireless transmission techniques. Hence, some new wireless transmission techniques should be developed to support the emerging applications in smart cities. In this survey, we focused on the artificial intelligence, spectrum management, and caching techniques, where the interference arose due to the limited spectrum resources. In particular, we first reviewed the current research status of these new techniques and, then, gave some challenges on the system design. We further provided several feasible solutions on these challenges, in order to implement the IoT networks in smart cities. In future works, we will take a more comprehensive literature review on some other wireless transmission techniques such as the intelligent reflecting surface [44–46] or artificial intelligent learning-based algorithms [47–49], in order to further enhance the system performance of cache-enabled IoT networks.

Data Availability

The data in this work can be available through the email to the authors of this paper.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Key-Area Research and Development Program of Guangdong Province (No.2018B010124001), by the Science & Technology Projects of China Southern Power Grid (No. SEPRI-K205020), by the Natural Science Foundation of Guangdong Province (No. 2021A1515011392), and in part by the research program of Guangzhou University (No. YK2020008/YJ2021003).

References

- [1] G. Gui, F. Liu, J. Sun, J. Yang, Z. Zhou, and D. Zhao, “Flight delay prediction based on aviation big data and machine learning,” *IEEE Transactions on Vehicular Technology*, vol. 69, no. 1, pp. 140–150, 2020.
- [2] X. Li, M. Zhao, Y. Liu, L. Li, Z. Ding, and A. Nallanathan, “Secrecy analysis of ambient backscatter noma systems under i/q imbalance,” *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 12286–12290, 2020.
- [3] X. Liu and X. Zhang, “Noma-based resource allocation for cluster-based cognitive industrial internet of things,” *IEEE Transactions on Industrial Informatics*, vol. 16, no. 8, pp. 5379–5388, 2020.
- [4] J. Xia, D. Deng, and D. Fan, “A note on implementation methodologies of deep learning-based signal detection for conventional MIMO transmitters,” *IEEE Transactions on Broadcasting*, vol. 66, no. 3, pp. 744–745, 2020.
- [5] X. Liu, X. B. Zhai, W. Lu, and C. Wu, “Qos-guarantee resource allocation for multibeam satellite industrial internet of things with NOMA,” *IEEE Transactions on Industrial Informatics*, vol. 17, no. 3, pp. 2052–2061, 2021.
- [6] Y. Zhang and J. Wang, “CV-3DCNN: complex-valued deep learning for CSI prediction in FDD massive MIMO systems,” *IEEE Wireless Communications Letters*, vol. 10, no. 2, pp. 266–270, 2021.
- [7] X. Li, H. Mengyan, Y. Liu, V. G. Menon, A. Paul, and Z. Ding, “I/q imbalance aware nonlinear wireless-powered relaying of b5g networks: security and reliability analysis,” *IEEE Transactions on Network Science and Engineering*, pp. 1–1, 2020.
- [8] S. Tang, “Battery-constrained federated edge learning in UAV-enabled IoT for B5G/6G networks,” *Phys. Commun.*, vol. 47, no. 101381, pp. 1–9, 2021.
- [9] X. Liu and X. Zhang, “Rate and energy efficiency improvements for 5g-based iot with simultaneous transfer,” *IEEE Internet of Things Journal*, vol. 6, no. 4, pp. 5971–5980, 2019.
- [10] W. Zhou, “PSO based offloading strategy for cache-enabled mobile edge computing UAV networks,” *Cluster Computing*, vol. PP, no. 99, pp. 1–8, 2021.
- [11] M. Liu and F. Tang, “6G: opening new horizons for integration of comfort, security, and intelligence,” *IEEE Wireless Communications*, vol. 27, no. 5, pp. 126–132, 2020.
- [12] W. Lu, Y. Gong, X. Liu, J. Wu, and H. Peng, “Collaborative energy and information transfer in green wireless sensor networks for smart cities,” *IEEE Trans. Ind. Informatics*, vol. 14, no. 4, pp. 1585–1593, 2018.

- [13] Y. Wang and F. Adachi, "Multi-task learning for generalized automatic modulation classification under non-Gaussian noise with varying SNR conditions," *IEEE Trans. Wireless Commu.*, vol. 20, no. 6, pp. 3587–3596, 2021.
- [14] X. Li, Q. Wang, M. Liu et al., "Cooperative wireless-powered NOMA relaying for B5G IoT networks with hardware impairments and channel estimation errors," *IEEE Internet of Things Journal*, vol. 8, no. 7, pp. 5453–5467, 2021.
- [15] W. Lu, S. Hu, X. Liu, C. He, and Y. Gong, "Incentive mechanism based cooperative spectrum sharing for OFDM cognitive IoT network," *IEEE Trans. Netw. Sci. Eng.*, vol. 7, no. 2, pp. 662–672, 2020.
- [16] C. Cheng and L. Guo, "Machine learning-aided trajectory prediction and conflict detection for internet of aerial vehicles," *IEEE Internet Things J.*, vol. PP, no. 99, pp. 1–12, 2021.
- [17] X. Li, M. Zhao, M. Zeng et al., "Hardware impaired ambient backscatter NOMA systems: reliability and security," *IEEE Transactions on Communications*, vol. 69, no. 4, pp. 2723–2736, 2021.
- [18] X. Lai, "Secure mobile edge computing networks in the presence of multiple eavesdroppers," *IEEE Trans. Commun.*, vol. PP, pp. 1–12, 2021.
- [19] J. Chakareski, S. Naqvi, N. Mastronarde, J. Xu, F. Afghah, and A. Razi, "An energy efficient framework for uav-assisted millimeter wave 5G heterogeneous cellular networks," *IEEE Transactions on Green Communications and Networking*, vol. 3, no. 1, pp. 37–44, 2019.
- [20] J. Xia, L. Fan, N. Yang et al., "Opportunistic access point selection for mobile edge computing networks," *IEEE Transactions on Wireless Communications*, vol. 20, no. 1, pp. 695–709, 2021.
- [21] J. Xia, "Secure cache-aided multi-relay networks in the presence of multiple eavesdroppers," *IEEE Transactions on Communications*, vol. 67, no. 11, pp. 7672–7685, 2019.
- [22] L. Fan, N. Zhao, X. Lei, Q. Chen, N. Yang, and G. K. Karagianidis, "Outage probability and optimal cache placement for multiple amplify-and-forward relay networks," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 12, pp. 12 373–12 378, 2018.
- [23] Y. Qiu, H. Zhang, K. Long, Y. Huang, X. Song, and V. C. M. Leung, "Energy-efficient power allocation with interference mitigation in mmwave-based fog radio access networks," *IEEE Wireless Communications*, vol. 25, no. 4, pp. 25–31, 2018.
- [24] K. He, "Ultra-reliable MU-MIMO detector based on deep learning for 5G/B5G-enabled IoT," *Phys. Commun.*, vol. 43, pp. 1–7, 2020.
- [25] Z. Wang, "An adaptive deep learning-based UAV receiver design for coded MIMO with correlated noise," *Phys. Commun.*, vol. 45, no. 101295, pp. 1–8, 2021.
- [26] K. He, "Learning based signal detection for MIMO systems with unknown noise statistics," *IEEE Transactions on Communications*, vol. 69, no. 5, pp. 3025–3038, 2021.
- [27] R. Zhao, "Deep reinforcement learning based mobile edge computing for intelligent internet of things," *Phys. Commun.*, vol. 43, pp. 1–7, 2020.
- [28] S. Lai, "Intelligent secure mobile edge computing for beyond 5G wireless networks," *Phys. Commun.*, vol. 45, no. 101283, pp. 1–8, 2021.
- [29] C. Li, "Dynamic offloading for multiuser multi-CAP MEC networks: a deep reinforcement learning approach," *IEEE Trans. Vehic. Tech.*, vol. 70, no. 3, pp. 2922–2927, 2021.
- [30] Y. Guo, "Efficient and flexible management for industrial internet of things: a federated learning approach," *Computer Networks*, vol. 192, pp. 1–9, 2021.
- [31] X. Lai, "Cybertwin-driven mobile edge computing for internet of everything with cochannel interference," *IEEE Trans. Ind. Informatics*, vol. PP, no. 99, pp. 1–12, 2021.
- [32] S. Lai and Y. Guo, "Distributed machine learning for multiuser mobile edge computing systems," *IEEE J. Sel. Top. Signal Process.*, vol. PP, no. 99, pp. 1–12, 2021.
- [33] L. Chen, "Intelligent ubiquitous computing for future UAV-enabled MEC network systems," *Cluster Computing*, vol. PP, no. 99, pp. 1–8, 2021.
- [34] B. Wang, F. Gao, S. Jin, H. Lin, and G. Y. Li, "Spatial- and frequency-wideband effects in millimeter-wave massive MIMO systems," *IEEE Trans. Signal Processing*, vol. 66, no. 13, pp. 3393–3406, 2018.
- [35] X. Hu, C. Zhong, Y. Zhu, X. Chen, and Z. Zhang, "Programmable metasurface-based multicast systems: design and analysis," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1763–1776, 2020.
- [36] X. Hu, C. Zhong, Y. Zhang, X. Chen, and Z. Zhang, "Location information aided multiple intelligent reflecting surface systems," *IEEE Transactions on Communications*, vol. 68, no. 12, pp. 7948–7962, 2020.
- [37] J. Zhao and N. Shanjin, "Multiband cooperation for 5G Het-Nets: a promising network paradigm," *IEEE Vehicular Technology Mag.*, vol. PP, no. 99, pp. 1–10, 2019.
- [38] J. Zhao, X. Sun, Q. Li, and X. Ma, "Edge caching and computation management for real-time internet of vehicles: an online and distributed approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 4, pp. 2183–2197, 2021.
- [39] J. Zhao, Q. Li, Y. Gong, and K. Zhang, "Computation offloading and resource allocation for cloud assisted mobile edge computing in vehicular networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 8, pp. 7944–7956, 2019.
- [40] J. Yang, D. Ruan, J. Huang, X. Kang, and Y. Shi, "An embedding cost learning framework using GAN," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 839–851, 2020.
- [41] H. Xie, F. Gao, S. Zhang, and S. Jin, "A unified transmission strategy for TDD/FDD massive MIMO systems with spatial basis expansion model," *IEEE Trans. Vehicular Technology*, vol. 66, no. 4, pp. 3170–3184, 2017.
- [42] K. Du, H. Wu, and G. Huang, "Condensation-free radiant cooling with double-skin infrared-transparent membranes," *Building and Environment*, vol. 193, p. 107660, 2021.
- [43] J. Yang, H. Wu, and X. Xu, "Regional climate effects on the optimal thermal resistance and capacitance of residential building walls," *Energy and Buildings*, vol. 244, no. 111030, 2021.
- [44] X. Hu, J. Wang, and C. Zhong, "Statistical CSI based design for intelligent reflecting surface assisted MISO systems," *Science China: Information Science*, vol. 63, no. 12, p. 222303, 2020.
- [45] J. Zhang, Y. Zhang, C. Zhong, and Z. Zhang, "Robust design for intelligent reflecting surfaces assisted MISO systems," *IEEE Communications Letters*, vol. 24, no. 10, pp. 2353–2357, 2020.
- [46] Q. Tao, J. Wang, and C. Zhong, "Performance analysis of intelligent reflecting surface aided communication systems," *IEEE Communications Letters*, vol. 24, no. 11, pp. 2464–2468, 2020.

- [47] Z. Zhao, "System optimization of federated learning networks with a constrained latency," *IEEE Trans. Vehic. Tech.*, vol. PP, no. 99, pp. 1–5, 2021.
- [48] S. Tang, "Dilated convolution based CSI feedback compression for massive MIMO systems," *IEEE Trans. Vehic. Tech.*, vol. PP, no. 99, pp. 1–5, 2021.
- [49] J. Xia and L. Fan, "Computational intelligence and deep reinforcement learning for next-generation industrial IoT," *IEEE Trans. Network Science and Engineering*, vol. PP, no. 99, pp. 1–12, 2021.