

文章编号:1007-5321(2018)05-0052-10

DOI:10.13190/j.jbupt.2018-258

B5G 系统中基于无线大数据的新兴技术

张四海¹, 张建华², 陈颖³, 朱近康¹

(1. 中国科学技术大学 中国科学院无线光电通信重点实验室, 合肥 230027;

2. 北京邮电大学 网络与交换技术国家重点实验室, 北京 100876;

3. 浙江大学 信息与电子工程学院 英国约克大学-浙江大学认知网络与绿色通信联合实验室, 杭州 310058)

摘要: 基于无线大数据(WBD)和人工智能(AI)的通信技术(涵盖物理层、网络层和应用层)被认为是最有前景的研究之一。该领域三项有趣的研究工作包括信道建模、大规模接入和网络拓扑设计。信道建模部分从机器学习在无线信道建模中应用的可行方法入手,介绍了参数估计中的主流方法,即信道多径聚类,该方法对于未来的研究具有重要意义。大规模接入部分关注分形现象及其在无线网络中的可能应用,主要研究了分形 D2D 社交网络的最大容量。网络拓扑设计部分介绍了如何利用移动用户的动态移动性特征来减少超密集网络(UDN)中的无线资源消耗。这些工作被认为是 5G 后有发展前景的研究领域,无线大数据分析为相关研究的未来工作提供了可能线索。

关键词: 无线大数据; Beyond 5G; 大规模接入; 信道建模; 移动性识别

中图分类号: TN929.53

文献标志码: A

Wireless Big Data Enabled Emerging Technologies for Beyond 5G System

ZHANG Si-hai¹, ZHANG Jian-hua², CHEN Ying³, ZHU Jin-kang¹

(1. Key Laboratory of Wireless-Optical Communications, CAS, University of Science and Technology of China, Hefei 230027, China;

2. State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China;

3. York-Zhejiang Laboratory for Cognitive Radio and Green Communications, College of Information Science & Electronic Engineering, Zhejiang University, Hangzhou 310058, China)

Abstract: The fifth generation of mobile communications system (5G) will be deployed in 2020 and provide diverse communication capabilities, but the promising future of beyond 5G is surely necessary for the communication requirements incurred by fast growing information technology in the next decade. Among several roadmaps toward beyond 5G, wireless big data (WBD) plus artificial intelligence (AI) based communication technology, which covers physical layer, network layer and application layer, is considered as one of the most promising ways. Along with this thought, some emerging research works have been published, which further stimulate more researcher to pay more attention in this area. This paper introduces three interesting works toward this aim, which covers channel modelling, huge access, and network topology design. The channel modelling part starts with the feasible ways to apply machine learning to wireless channel modelling, and presents the prevailing methods in parameter estimation, channel multipath clustering, which is of great importance for future research. The huge access part focuses on fractal

收稿日期: 2018-10-19

基金项目: 国家自然科学基金重点项目(61631018); 科技部重大科技专项项目(2018ZX03001031); 北京市自然科学基金重点项目(L172030)

作者简介: 张四海(1974—), 男, 副教授, E-mail: shzhang@ustc.edu.cn.

phenomenon and its possible applications in wireless networks. After introducing the basic concept, this part investigates the maximum capacity fractal D2D social networks. The network topology design part proposes an interesting topic, whose motivation is to utilize the dynamic mobility features of mobile users to decrease the wireless resource consumption in ultra dense networks (UDN). In summary, these three works are considered as promising topics in beyond 5G, which combines wireless big data analysis and may shed light on related future research.

Key words: wireless big data; beyond 5G; huge access; channel modelling; mobility recognition

The fifth generation of mobile communications system will be widely deployed and commercially released in 2020, which can provide diverse communication capability, so that 5G can support not only the massive mobile users for emerging novel services, but also the vertical industries with diverse performance requirements. Thereafter, 5G networks will accommodate more numerous wireless devices, which will increase the data amount generated, collected and stored in 5G networks. Such big data resource, which has been coined as wireless big data (WBD)^[1], will thoroughly incur and enable novel research topics and application areas in the era of beyond 5G system.

Note that there is a strong connection between 5G and wireless big data in several aspects. First, as mentioned above, the wireless devices connected to 5G will be at least 10 000 times more than that in 4G era, which will generate surprisingly huge data. Second, with the advancing 5G, internet of things (IoT) is rapidly gaining ground. The application of IoT is widespread now, including intelligent transportation, environment protection, public security, industrial monitoring, individual health and etc. Due to the wide application of IoT and advanced sensor, the data volume of IoT is rapidly increasing. Finally, the technology development in wireless communications and networking, including the growing bandwidth, massive antennas and the implementation of software-defined networking (SDN) and network function virtualization (NFV), not only makes the data amount growing feasible, but also enables the convenient data collection and preprocessing in 5G and beyond era. Thereafter, the telecommunication operators are paying more attention on wireless big data so that the big data value will be really cherished for both commercial usage and sci-

entific research.

In the past several years, the wireless big data research has attracted more researchers' interests, especially in China. The two main reasons might be, more and more researchers realize that, the improvement of communication system may come from the power of computing, that is, so called the computing communications, and the huge data generated in China and relatively not strict data protection policy encourages the collaboration among researchers and wireless big data holders. Up till now, in our opinions, such research direction can be divided into two areas, physical layer^[2] and network layer^[3].

In the physical layer, with the development of big data and wireless communication, lots of research works are trying to utilize machine learning (ML) and big data to the physical technique, like multiple input multiple output (MIMO) detection^[4-5], channel estimation^[6], channel modeling^[7] and channel decoding and demodulation^[8-9], deep learning related MIMO channel topics^[10-11]. Therefore, wireless big data emerges and its related technologies are employed to traditional communication research.

The wireless channel is essentially a physical electromagnetic wave, and the current 5G channel model research follows the traditional way. In the 5G mobile communication system, one of the most difficult challenges is the complex and versatile propagation channel. In the different carrier frequencies, propagation environments, antenna structures, and bandwidths, the wireless channel will present complex space-time-frequency characteristics^[12]. As is known, channel coding, modulation, multi-antenna (MIMO), etc. all need to confront the versatile wireless channel. Therefore, a precise channel model with low complexity will

benefit the performance of wireless communication system a lot. With the increased antenna number, huge bandwidth and versatile application scenarios, the channel measurement data will always present in big volume^[7].

In the network layer, besides that network behavior recognition works^[13-14] have been quite popular, content caching in wireless edge^[15] is another important field, whose objective is to decrease the network latency and consume less network resources to meet the content requests from massive mobile users. One key element to improve the caching performance is to understand and utilize human mobility^[16-17]. Another key issue is how to design and manage the beyond 5G networks via the wireless big data analysis.

As 5G will be deployed in 2020, more and more institutions and researchers propose beyond 5G or potential sixth generation (6G). In Feb. 2018, China also announces to start early research of 6G. Globally beyond 5G will be targeted higher data rate, deeply merging with IoT and ever higher frequency bands. And all of those continuously bring high data volume and WBD will be in key role for 5G beyond.

This paper is not intended to present all aspects of wireless big data for 5G beyond, but only focuses on three emerging and promising research works. The first work is about the channel modelling, which starts from the introduction on combination of ML and wireless channel modelling, then presents the prevailing methods in parameter estimation, channel multipath clustering, which is of great importance for future research. The second work is about the huge access, which focuses on fractal phenomenon and its possible applications in wireless networks. After introducing the basic concept, this part investigates the maximum capacity fractal D2D social networks. The last work discusses the network topology design, whose motivation is to utilize the dynamic mobility features of mobile users to decrease the wireless resource consumption in ultra dense networks (UDN).

The rest of this paper is organized as follows. Section 1 introduces the efficient transmission under complex channels. Section 2 presents the novel under-

standing on the fractal phenomenon of wireless networks. Section 3 discusses the interesting question on better network design via user mobility recognition. Finally, section 4 concludes this paper.

1 Efficient transmission in complex channel

The progress of the channel modelling stimulated by data mining and ML is shown in Fig. 1. Firstly, the channel data with a specific frequency in one scenario is collected, stored and pre-processed^[12,18]. Thus, the big database by gathering data from various scenarios and frequency is constructed.

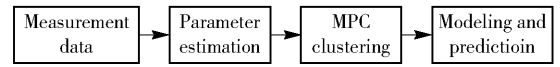


Fig. 1 Channel modelling procedure

1.1 Parameter estimation

Post-processing of the measurement channel data in the analysis of channel characteristics is carried out. Channel parameters including delay, angles of departure and arrival in the azimuth and elevation domains, Doppler frequency and the complex amplitude are extracted. Parametric and non-parametric estimation algorithms are two kinds of classical ones which are widely used. Classical beamforming and Capon beamforming are two conventional non-parametric algorithms, proposed in 1950s and 1960s in Refs. [19-20]. However, the estimation precision of non-parametric algorithms is limited by the antenna aperture which cannot distinguish different paths from a very small angular range. Since 1980s, a few parametric algorithms are proposed including multiple signal classification (MUSIC), estimating signal parameters via rotational invariance techniques (ESPRIT) algorithms^[21-22]. In these years, space-alternating generalized expectation-maximization (SAGE) algorithm^[23] based on the maximum likelihood though is widely used in space-time-frequency domain parameter extraction. However, the accuracy of these algorithms is not as reliable as the parametric algorithm for the estimation results cannot be verified. In future 5G beyond systems, for the plane wave hypothesis is unsatisfied,

parametric algorithms might be fit in this case. Therefore, non-parametric algorithm may be significant when the millimeter-wave and massive MIMO technologies are used in future.

In future, channel time-variant characteristic is important in some high speed moving scenario, i. e., high-speed train and vehicular-to-vehicular communication. The channel parameters will vary quickly in such scenarios and channel parameter extraction and prediction become difficult. However, ML is very fit for the prediction problem. Combined with channel parameter estimation algorithm, we can use some data mining technologies, i. e., principal component analysis (PCA)^[24] and clustering algorithms to extract the abstract characteristics and then use Kalman filtering^[25] or some ML algorithm, i. e., neural network to match the rule of channel characteristics with time. Besides, based on the thought of sparse representation which is widely used in compressed sensing, more useful information will be gotten from a large amount of measured channel data.

1.2 Channel multipath clustering

The cluster of MPCs is defined as a group of multipath with similar parameters. There are many clustering algorithms used for MPC clustering. A kernel-power-density-based algorithm is proposed for MPC clustering, where the kernel density of the MPCs is incorporated to model the MPCs^[26]. In Ref. [27], a framework for evaluation and development of different cluster algorithm is discussed. Our team employed the Gaussian mixture model (GMM) to fit the MPCs^[18], exhibiting preferable clustering result. Using sufficient statistic characteristics of channel multipath, the GMM can get clusters corresponding to the multipath propagation characteristics. The GMM assumes that all the MPCs consist of several Gaussian distributions in varying proportions. Given a set of N channel multipath X , the log-likelihood of the Gaussian mixture model is

$$L(X; \Omega) = \sum_{i=1}^N \text{lb} \sum_{k=1}^K \pi_k p(x_i | z_i; \mu_k, \Sigma_k) \quad (1)$$

where $\Omega = \{ \pi_k, \mu_k, \Sigma_k, k=1, 2, \dots, K \}$ is the set of all the parameters, μ_k, Σ_k is the expectation and variance of k -th Gaussian distribution respectively and $\pi_k \in [0,$

$1]$ is the prior probability satisfying the constraint $\sum_{k=1}^K \pi_k = 1$. To estimate the GMM parameters, expectation maximization (EM) algorithm is employed to solve the log-likelihood function of GMM.

Fig. 2 illustrates the simulation results of GMM clustering algorithm, where the GMM clustering obtains clearly as well as compact clusters. As scattering property of the channel multipath obeys Gaussian distribution, the compact clusters can accord with the multipath scattering property.

1.3 Cluster-nuclei based channel model

In the cluster-nuclei based channel model, the MPCs are aggregated into a traditional stochastically channel model. At the same time, the scene is discerned by the computer and the environment is rebuilt by ML methods. Then, by matching the real propagation objects with the clusters, the cluster-nuclei can be easily found, which are the key factors in contacting deterministic environment and stochastic clusters. To be specific, the cluster-nuclei is defined as clusters which is aggregated by a large number of waves. There are three important features for cluster-nuclei: 1) it has a certain shape, 2) it has the mapping relation between scatters in the real propagation environment and clusters, 3) it dominates the channel impulse response generation in various scenarios and configurations. As cluster-nuclei has mapping relationship with real propagation environment, it is superior to cluster which has not physical meanings.

2 Huge access on fractal base stations

With the explosive increase of smart mobile devices, social network traffic has witnessed unprecedented growth and imposed huge challenge on traditional content delivery paradigm. Emerging as a promising technology to offload the wireless network traffic, massive device-to-device (D2D) communications allow users in proximity to establish local links and exchange contents directly instead of obtaining data from the cellular base station (BS).

Within the massive D2D communication scenarios, besides the underlying propagation network on the

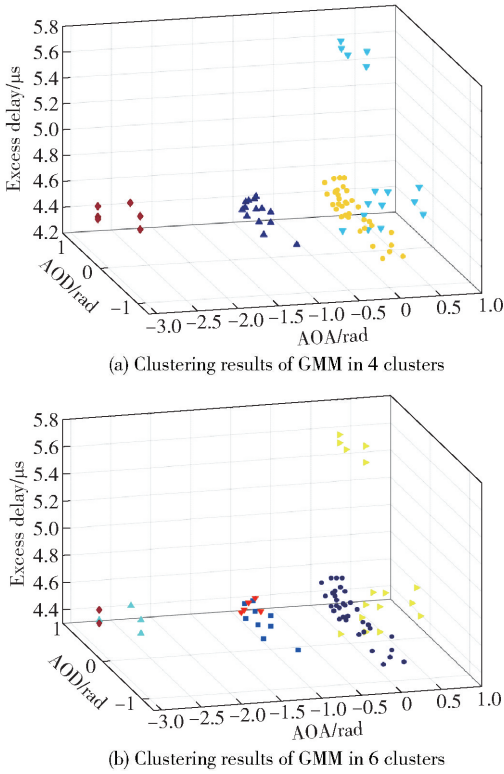


Fig. 2 Clustering results of GMM in visual aspect

physical layer, the users also forms an overlaying social network, where the communication between two users is driven by their social relationship and served by the underlying propagation network. Particularly, with the increasing awareness of security and privacy, trust has become a prerequisite for interactions between mobile users. People only communicate with trusted persons rather than geographically close ones.

As a vital property of networks, fractal phenomenon has already been discovered in many wireless networking scenarios. For example, the coverage boundary of the wireless cellular networks shows a fractal shape, and the fractal features can inspire the new design of the hand-off scheme in mobile terminals. Moreover, a large number of significant networks in the real world exhibit the fractal characteristics naturally, such as the world-wideWeb, yeast interaction, protein homology, and social networks. In addition, the concept of fractal structure has been taken advantage of in various applications, including the design of antennas for satellite down-link and up-link communications, wireless local area network (WLAN) applications, and other 5G applications.

Hereinafter, fractal organization is considered in the massive D2D social networks due to its predominant performance in terms of resilience, scalability and robustness than non-fractal organizations. Specifically, a fractal social network can recover quickly from security attacks because the breakdown of a few nodes does not cause the collapse of the whole network. Therefore, it is significantly important to study fractal D2D social networks and answer the fundamental problem like the capacity, robustness and reliability of fractal D2D social networks.

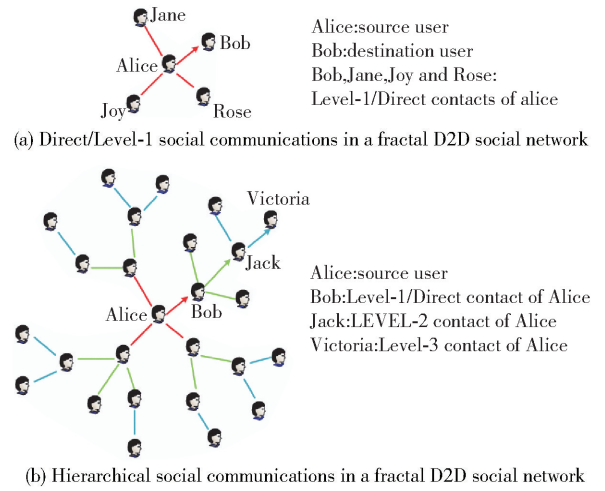
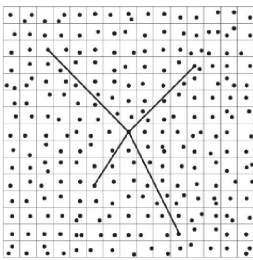


Fig. 3 Social communications in a fractal D2D social network

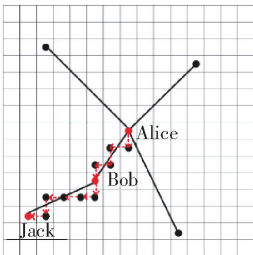
Fig. 3 (a) illustrates the direct/level-1 social communications in a fractal D2D social network. As we can see, four users, namely Bob, Jane, Joy and Rose, are directly connected with Alice and are regarded as the direct, or level-1 contacts of Alice. If Alice chooses to communicate with Bob among her four direct contacts, then Alice and Bob are known as the source user and the destination user, respectively. Usually, a user has more than one direct contacts, and the degree k refers to the number of his/her level-1 contacts. In the case of level-1 social communications, the degree distribution and the joint probability distribution are the aforementioned $P(k)$ and $P(k_1, k_2)$, respectively. The direct/level-1 contact does not imply there physically exist some direct links. Instead, the pair of users for direct social contact might have to rely on some relaying nodes in the underlying physical propagation network.

In addition to the direct case, the social communications in fractal D2D social networks can actually be hierarchical as depicted in Fig. 3(b). If Alice wants to get in touch with Victoria who she does not trust, the data packets have to be transmitted through the inter-users Bob and Jack. That is to say, a source user can communicate with one of his/her level- L ($L = 1, 2, \dots, L_{\max}$) contacts through $L - 1$ inter-users to make sure that every transmission is carried out between two users with mutual trust, and L_{\max} refers to the maximum social relationship level. For instance, in the case of level-2 social communications, Jack is indirectly connected with the source user Alice through one inter-user Bob, so Jack is one of the level-2 contacts of Alice, and he can be selected as the destination user among all the level-2 contacts to communicate with Alice. Similarly, Victoria is referred to as one of the level-3 contacts of Alice, and so on.

In order to clarify the performance of the above fractal D2D social networks with both direct and hierarchical communications clearly and orderly, it is assumed that all of the n users are uniformly distributed in a unit area square. Also the fractal D2D social network is treated as a static network because the users barely move during one transmission frame.



(a) An illustrative part of the overlaying fractal D2D network with social interconnections



(b) The underlying physical propagation network serves to forward data for a transmission via multi-hop routing between any pair of social contacts

Fig. 4 Fractal D2D network with social interconnections

All the potential users form an underlying D2D propagation network on the physical layer, as well as an overlaying fractal social network from the viewpoint of social connections. An illustrative part of the overlaying fractal D2D social network is shown in Fig. 4(a), and the connection between two users stands for the relationship of mutual trust. It is noteworthy that the topological fractal social network is formed by the D2D social connections of all the involved users following the aforementioned degree distributions $P(k)$ and $P(k_1, k_2)$, which is not contradictory with the general assumption of physically uniformly distributed users.

As depicted in Fig. 4(b), the underlying D2D physical propagation network has to be distinguished from the overlaying fractal social network, where the propagation network serves the social communications and forwards data for a transmission via multi-hop routing between any pair of social contacts. For example, when Alice wants to communicate with Jack, she has to get in touch with Jack through Bob. However, Alice and Bob cannot exchange data directly even though they are socially connected because they are not physically close enough to exchange contents locally. In order to transmit a packet from Alice to Bob, a few other nodes in the underlying D2D propagation network have to serve as relay nodes, mentioned as the red dotted path in Fig. 4(b), so does the transmission from Bob to Jack. It has been explained that the relay nodes will never cause traffic bottleneck, so the underlying propagation network will not change the capacity of the overlaying social network.

In particular, we have investigated the maximum capacity of fractal D2D social networks with both direct and hierarchical communications^[28]. Under the condition of direct social communications, it has been proved that if the source user communicates with one of his/her direct contacts randomly, the maximum capacity corresponds to the classical well-known result $\Theta\left(\frac{1}{\sqrt{n \ln n}}\right)$ achieved by Kumar. On the other hand, if the two users with distance d communicate with each other according to the probability in proportion to $d^{-\beta}$, the maximum capacity is

$$\lambda_{\max} = \begin{cases} \Theta\left(\frac{1}{\sqrt{n \ln n}}\right), & 0 \leq \beta \leq 2 \\ \Theta\left(\frac{1}{\sqrt{n^{3-\beta} \ln n^{\beta-1}}}\right), & 2 < \beta < 3 \\ \Theta\left(\frac{1}{\ln n}\right), & \beta \geq 3 \end{cases}$$

While taking social communications of all levels into account, for both uniform and power-law destination selection cases, it is discovered that the hierarchical social communications further decreases the respective maximum capacity in a proportion related to the number of users n , and the corresponding reduction factor varies by different values of the correlation exponent ε of the fractal D2D social networks:

$$\lambda_{\max}^{(H)} = \begin{cases} \Theta\left(\lambda_{\max} \frac{1}{\ln n}\right), & 2 < \varepsilon < 3 \\ \Theta(\lambda_{\max} n^{-1}), & \varepsilon = 3 \end{cases}$$

Surely, there are still some issues remain to be solved in the future studies. For instance, why the condition $\varepsilon = 3$ is the boundary to determine whether or not the fractal network is extensible. Moreover, why is there a leap in the reduction coefficient of hierarchical social communications when $\varepsilon = 3$. We leave all these open issues in the future works.

3 UDN resource minimization via user mobility recognition

The UDNs has been identified as an appealing solution to address the huge service demands in future 5G and beyond. Heterogeneous, overlapping and efficient deployment of UDNs with a large number of access points will be the important coverage features. But how to enable a wide range of mobility support is a great challenge, for which, dividing the UDN into subnets to meet different user groups with certain user demands might be one possible solution. The basic motivation of our work is that, by wireless big data analysis, the user mobility behavior and the user demand can be better understood, thus we may design better subnets of UDN using minimum radio resource. One simple example of this radio resource design may be the optimization of frequency reuse. In this section, a novel optimization design for UDN by the user grouping

is given based on the mobility and the subnet parameters are jointly adjusted to meet the service demand using minimum resource cost.

3.1 Problem formulation

Consider one circle area whose radius is R , covered by UDN with fixed wireless radio resources B , where such resources are defined as frequency bandwidth in this work. The UDN can be divided into G subnets with different coverage and bandwidth for maximizing the overall capacity, and the transmission capacity of g -th subnet is $C_g(t)$, $g = 1, 2, \dots, G$, so the total transmission capacities of UDN is defined by $C_{\text{UDN}}(t)$ as:

$$C_{\text{UDN}}(t) = \sum_{g=1}^G C_g(t) \quad (2)$$

Note that, g -th subnet is assigned with $B_g(t)$ bandwidth and subnet coverage radius $R_g(t)$, where $\sum_{g=1}^G B_g(t) \leq B$ and $R_g(t) \leq R$. Thus the area capacity of g -th subnet, $C_g(t) = B_g(t) \sigma_g(R_g(t))$, where $\sigma_g(\cdot)$ denotes the area capacity per unit bandwidth (also called as the area spectral efficiency^[29]).

Assume M users are located in this area, and the total service demands from these users at time t is denoted as $S_{\text{total}}(t) = \sum_{m=1}^M s_m(t)$, $M \gg 1$, where $s_m(t)$ is a time-dependent function which denotes the service request of m -th user.

So the total transmission capacities of UDN $C_{\text{UDN}}(t)$ must meet the total service demands, that is

$$C_{\text{UDN}}(t) = \sum_{g=1}^G C_g(t) \geq S_{\text{total}}(t) = \sum_{m=1}^M s_m(t) \quad (3)$$

$v_m(t)$ is defined as the velocity function of user m , and the speed interval of user velocity, accepted by g -th subnet is denoted as $V_g = [v_g, v_{g+1})$. So the users can be assigned to different subnet according to their moving speed. Thus M users can be clustered first by their mobility nature, or moving speed for the simplest case in this work, using data mining method of wireless big data, and then assigned to G subnets. The users being assigned to g -th subnet is denoted as U_g . Different types of subnets, with different coverage radius in our case, like as the typical macro cell network, the

micro cell network, the small cell structure, and the pico/femto cells, might be deployed in given areas to support different user demands.

So, the total M users can be divided into G user groups or subnet, where each group or subnet is defined with a user moving speed range, so that the speed of those users in this group are in this speed range. Note that, this is optimization problem which is to obtain the maximal area capacity through properly dividing the speed intervals. Thus, the problem is formulated as follows:

$$\begin{aligned}
 & \min_{G, [v_g(t), v_{g+1}(t)]_{g=1}^G} \sum_{g=1}^G B_g(t) \\
 \text{s. t. } & \sum_{g=1}^G B_g(t) \leq B \\
 & \bigcup_{g=1}^G U_g = M \\
 & U_{g_1} \cap U_{g_2} = \emptyset, g_1 \neq g_2 \\
 & 1 \leq g \leq G
 \end{aligned} \quad (4)$$

3.2 Optimization methods

From the problem formulation above, we can prove that the total service demands can be achieved using the least resource consumption, according to the area capacity efficiency maximized which is corresponding to the relative grouped demands. For more details, the readers can refer to the literature^[30].

Then we summarized the optimization methods as following three steps:

(i) Divide the service requirements of many users into G service demands grouped by user moving speeds based on wireless big data.

(ii) Analyze the appropriate variables and the achievable capacity efficiency of the selected subnets to meet the necessary conditions.

(iii) Based on above (i) and (ii), the resource consumption of g -th subnet, B_g , is made minimum to satisfy the constraint conditions.

Therefore, the optimization design of UDNs combines the user groups and the subnets capacity efficiency based on WBD to achieve the huge service demands by the least resource cost.

3.3 Numerical results

We present a simple but straight forward example

to demonstrate the problem. In this example, total of 200 users are considered, each having service demand ranging from 0 Mbit/s to 40 Mbit/s and the moving speed ranging from 0 to 140 km/h. The detailed example setting is illustrated in Fig. 5, where the total service demands of all users is 30.608 Gbit/s. The solid line indicates the real data, and the dotted line indicates the fitting result.

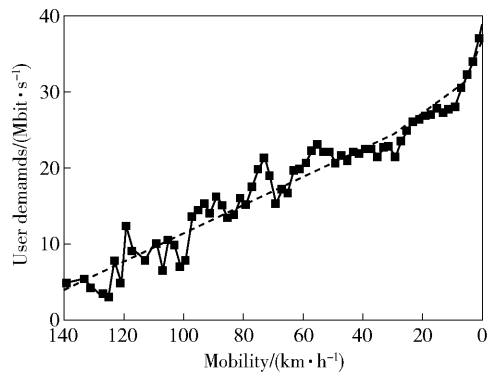


Fig. 5 Service demands and moving speed setting of users

Fig. 6 presents the numerical results, where at most 4 subnets are considered. The total service demands of all users is 30.608 Gbps in 1 km² area. In addition, we consider three moving speed interval separation methods: (A) Evenly split, (B) Doubled dropping and (C) Optimal interval. Thus, the total bandwidth consumption for three different strategies considering 1–4 subnets is summarized.

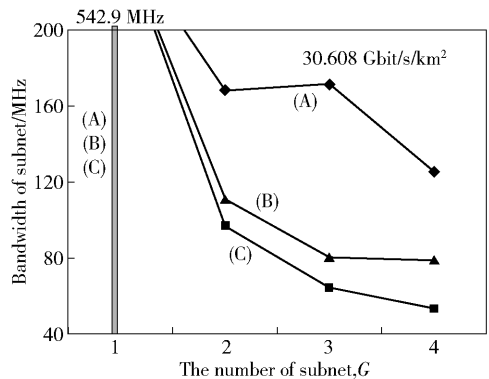


Fig. 6 Service demands and bandwidth cost for different strategies

For the $G = 1$ (only macro cell) case, the resource cost is $B_{\text{total}} = 542.9$ MHz to support the high moving speed. For the $G = 4$ case, where 4 subnets are used, the minimum frequency bandwidth using pro-

posed optimization method is reduced to $B_{\text{total}} = 53.0$ MHz, which is a 10x reduction than coverage cost using only one subnet.

4 Conclusion

This paper discusses several emerging technologies enabled by wireless big data for beyond 5G systems, which may be of essential importance for future mobile communications. To the best of our knowledge, the state of art research progress in wireless channel modeling, network topology recognition, and one novel network topology design are introduced. We note that there are other important topics and interesting works, but for the limitation of space, those are not covered in this paper. We hope this paper can open a new dimension to other researchers, to pursue further results in this area.

参考文献:

- [1] Qian Lijun, Zhu Jinkang, Zhang Sihai. Survey of wireless big data[J]. Journal of Communications and Information Networks, 2017, 2(1): 1-18.
- [2] Dörner S, Cammerer S, Hoydis J, et al. Deep learning based communication over the air[J]. IEEE Journal of Selected Topics in Signal Processing, 2018, 12(1): 132-143.
- [3] Klaine P V, Imran M A, Onireti O, et al. A survey of machine learning techniques applied to self organizing cellular networks[J]. IEEE Communications Surveys and Tutorials, 2017, 19(4): 2392-2431.
- [4] Samuel N, Diskin T, Wiesel A. Deep MIMO detection [C]//18th IEEE International Workshop on Signal Processing Advances for Wireless Communications (SPAWC). New York: IEEE Press, 2017.
- [5] O'Shea T J, Erpek T, Clancy T C. Deep learning based MIMO communications[EB/OL]. (2017-7-25) [2018-10-19]. arXiv preprint arXiv: 1707. 07980.
- [6] Dong Fang, Liu Junbiao, He Liang, et al. Channel estimation based on extreme learning machine for high speed environments[M]//Cao Jiuwen, Mao Kezhi, Wu J, et al. Proceedings of ELM-2015(Volume 1). Theory, algorithms and applications (I). Berlin: Springer, 2016: 159-167.
- [7] Zhang Jianhua. The interdisciplinary research of big data and wireless channel: a cluster-nuclei based channel model[J]. China Communications, 2016, 13 (Sup2): 14-26.
- [8] Nachmani E, Marciano E, Lugosch L, et al. Deep learning methods for improved decoding of linear codes[J]. IEEE Journal of Selected Topics in Signal Processing, 2018, 12(1): 119-131.
- [9] Liang Fei, Shen Cong, Wu Feng. An iterative BP-CNN architecture for channel decoding[J]. IEEE Journal of Selected Topics in Signal Processing, 2018, 12(1): 144-159.
- [10] Wen C K, Shih W T, Jin S. Deep learning for massive MIMO CSI feedback[J]. IEEE Wireless Communications Letters, 2018, 7(5): 748-751.
- [11] He Hengtao, Wen C K, Jin S, et al. Deep learning-based channel estimation for beamspace mmWave massive MIMO systems[J]. IEEE Wireless Communications Letters, 2018, 7(5): 852-855.
- [12] Zhang Jianhua, Zhang Yuxiang, Yu Yawei, et al. 3-D MIMO: how much does it meet our expectations observed from channel measurements? [J]. IEEE Journal on Selected Areas in Communications, 2017, 35 (8): 1887-1903.
- [13] Xu Fengli, Li Yong, Chen Min, et al. Mobile cellular big data: linking cyberspace and the physical world with social ecology[J]. IEEE Network, 2016, 30(3): 6-12.
- [14] Ma Ge, Wang Zhi, Zhang Miao, et al. Understanding performance of edge content caching for mobile video streaming[J]. IEEE Journal on Selected Areas in Communications, 2017, 35(5): 1076-1089.
- [15] Paschos G S, Iosifidis G, Tao M X, et al. The role of caching in future communication systems and networks [J]. IEEE Journal on Selected Areas in Communications, 2018, 36(6): 1111-1125.
- [16] Wang Rui, Peng Xi, Zhang Jun, et al. Mobility-aware caching for content-centric wireless networks: modeling and methodology[J]. IEEE Communications Magazine, 2016, 54(8): 77-83.
- [17] Deng Tao, Ahani G, Fan Pingzhi, et al. Cost-optimal caching for D2D networks with user mobility: modeling, analysis, and computational approaches [J]. IEEE Transactions on Wireless Communications, 2018, 17 (5): 3082-3094.
- [18] Li Yupeng, Zhang Jianhua, Ma Zhanyu, et al. Clustering analysis in the wireless propagation channel with a variational Gaussian mixture model[J]. IEEE Transactions on Big Data (Early Access), 2018.
- [19] Van Trees H L. Optimum array processing. Part IV of

- detection, estimation, and modulation theory[M]. New York: John Wiley & Sons, 2002.
- [20] Capon J. High-resolution frequency-wavenumber spectrum analysis[J]. Proceedings of the IEEE, 1969, 57(8): 1408-1418.
- [21] Schmidt R O. Multiple emitter location and signal parameter estimation[J]. IEEE Transactions on Antennas and Propagation, 1986, 34(3): 276-280.
- [22] Roy R, Kailath T. ESPRIT-estimation of signal parameters via rotational invariance techniques [J]. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1989, 37(7): 984-995.
- [23] Fleury B H, Jourdan P, Stucki A. High-resolution channel parameter estimation for MIMO applications using the SAGE algorithm[C]//2002 International Zurich Seminar on Broadband Communications Access-Transmission- Networking. New York: IEEE Press, 2002.
- [24] Ma Xiaochuan, Zhang Jianhua, Zhang Yuxiang, et al. Data scheme-based wireless channel modeling method: motivation, principle and performance [J]. Journal of Communications and Information Networks, 2017, 2(3): 41-51.
- [25] Heidari A, Khandani A K, McAvoy D. Adaptive modeling and long-range prediction of mobile fading channels [J]. IET Communications, 2010, 4(1): 39-50.
- [26] Schneider C, Gedschold J, Kaske M, et al. Estimation and characterization of multipath clusters in urban scenarios[C]//12th European Conference on Antennas and Propagation (EUCAP). London: IET Press, 2018.
- [27] NGMN Alliance. NGMN 5G white paper v1. 0 [EB/OL]. (2015-02-17). (2018-10-19). https://www.ngmn.org/fileadmin/ngmn/content/images/news/ngmn_news/NGMN_5G_White_Paper_V1_0.pdf.
- [28] Chen Ying, Li Rongpeng, Zhao Zhifeng, et al. On the capacity of wireless networks with fractal and hierarchical social communications [EB/OL]. (2018-8-11). (2018-10-19) arXiv preprint arXiv: 1708. 04585, 2017.
- [29] Alouini M S, Goldsmith A J. Area spectral efficiency of cellular mobile radio systems[J]. IEEE Transactions on Vehicular Technology, 1999, 48(4): 1047-1066.
- [30] Zhu Jinkang, Zhao Ming, Zhou Shengli. An optimization design of ultra dense networks balancing mobility and densification[J]. IEEE Access, 2018, 6: 32339-32348.