Received January 17, 2019, accepted February 11, 2019, date of publication February 15, 2019, date of current version March 5, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2899663

Statistical Multiplexing Gain Analysis of Processing Resources in Centralized Radio Access Networks

ZONGSHUAI ZHANG[1,2,3], (Student Member, IEEE), LIN TIAN[1,2,3], (Senior Member, IEEE), JINGLIN SHI[1,2,3], (Member, IEEE), JINHONG YUAN[1,2,4], (Fellow, IEEE), YIQING ZHOU[1,2,3], (Senior Member, IEEE), XINYU CUI[1,2,3], LU WANG[1,2,3], AND QIAN SUN[1,2,3], (Member, IEEE)

[1]Beijing Key Laboratory of Mobile Computing and Pervasive Device, Beijing 100080, China
[2]Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China
[3]School of Computer Science and Technology, University of Chinese Academy of Sciences, Beijing 100049, China
[4]School of Electrical Engineering and Telecommunications, The University of New South Wales, Sydney, NSW 2052, Australia

Corresponding author: Lin Tian (tianlindd@ict.ac.cn)

This work was supported by the National Science and Technology Major Project of China under Grant 2018ZX03001017.

ABSTRACT The next generation of wireless networks faces the challenges of the explosion of mobile data traffic, the associated power consumption, and operation cost. The centralized radio access networks (Centralized RANs) architectures have been proposed to reduce the power consumption and the network operating cost. By integrating many distributed base stations' processing resources in a processing pool and sharing processing resources on demand, the overall required processing resources for the Centralized RAN can be reduced compared to the conventional RAN. This can be measured by the statistical multiplexing gain (SMG). However, most of the SMG analysis only considered the temporal traffic distribution which is not suitable for the current mobile networks. In this paper, we analyze the SMG of processing resources based on a temporal–spatial joint traffic distribution model, which considers the mobile data traffic distribution both in the time and space domains. Based on this model, we derive a formula for the SMG and also a closed-form approximation for that when the spatial traffic distribution is lognormal distribution. The theoretical analysis and simulation results show that the SMG increases with the service threshold ratio P_{th}, but the growth trend of SMG for different area types is not always the same. We also find that the traffic distribution parameters, such as the standard deviation of the lognormal distribution variable's natural logarithm, have a significant influence on the SMG.

INDEX TERMS Centralized RAN, processing resources pool, statistical multiplexing gain, traffic distribution.

I. INTRODUCTION

Due to a widely use of ubiquitous smart devices, the global mobile data traffic is forecasted to be increased exponentially per month and mobile network connections speed will increase by more than three-fold by 2020 [1]. To meet the high traffic demands, the number of base stations (BSs) is expected to increase significantly which results in a series of problems including higher energy consumption, higher operating cost and stronger interference in radio access networks (RAN) [2]–[4]. In order to solve these problems, centralized radio access networks (Centralized RAN) architectures have been proposed, such as C-RAN [5], Super Base Station [6], [7] and *et al*. As shown in Fig. 1, in centralized radio access networks, the capabilities of many originally distributed BS processing resources are integrated in a centralized processing pool, which is connected to remote radio heads (RRHs) via high-speed fiber-optic switching. With the centralized network architecture, power consumption and network operating cost can be reduced significantly due to the relatively simple locating of RRHs and low cost maintenance of BS processing resources [5]. Moreover, cooperative operations of different BS processing resources can be easily implemented.

One of the most important features of Centralized RAN is that the processing resources can be shared between different BSs based on the centralized processing pool. In the traditional distributed radio access networks (Distributed

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

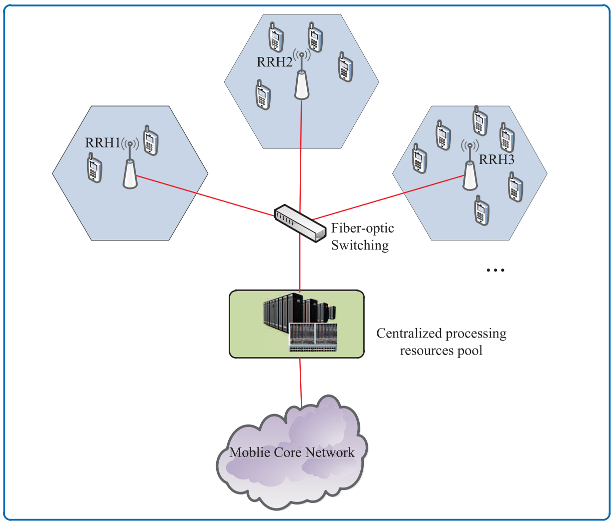


FIGURE 1. **Architecture of centralized radio access networks.**

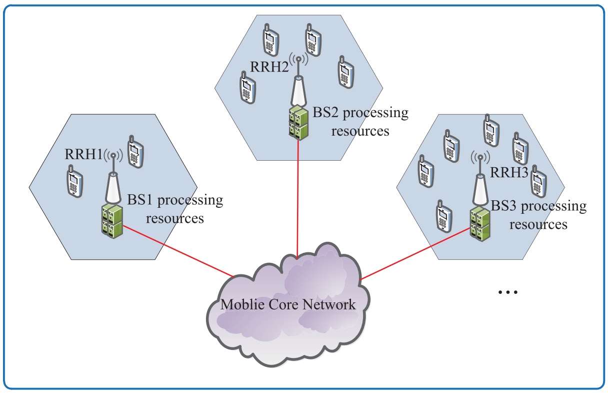


FIGURE 2. **Architecture of distributed radio access networks.**

RAN) architecture, processing resources are deployed to each BS as shown in Fig. 2. Each BS must be designed with the highest capacity to support the maximum possible traffic load in its cell. However, BSs traffic load varies throughout a day like a tide [5]. When a BS traffic load is low, most of its capacity is wasted because the physical resource of BSs cannot be shared among different BSs as they are isolated from each other. In Centralized RAN, there is no fixed mapping between the processing resources and BSs. A large amount of BS processing resources are physically grouped together and form a pool, which enable the statistical multiplexing of processing resources. Therefore, from ''exclusive owned'' processing resources in Distributed RAN to ''shared'' processing resources pool in Centralized RAN, a much higher utilization rate of processing resources can be achieved so that the related cost including power consumption, etc., can be reduced. In literature, statistical multiplexing gain (SMG) has been used to measure the performance of processing resources utilization. Although substantial statistical multiplexing gain of Centralized RAN has been observed by simulations and in real systems, a proper mathematical model and analysis are necessary to guide the practical system design.

Liu *et al.* [8], [9] assume that the user session obeys Poisson arrival, and construct a multi-dimensional Markov chain of user sessions with virtual base stations. Thereby, a statistical multiplexing gain model of virtual base station pools computing resources is established. A tractable model is proposed in [10] to analyze the statistical multiplexing gain of the fronthaul capacity in C-RAN based on the user blocking rate, which is caused by the fronthaul capacity limitation. Madhavan *et al.* [11] quantify the variation of the multiplexing gain of consolidated base stations in different traffic conditions through traffic simulation experiments. Checko *et al.* [12] present an evaluation of statistical multiplexing gain of Baseband Units (BBUs) in C-RAN based on the load of base stations throughout a day. However, the data traffic models used in these papers cannot reflect the temporal and spatial distribution characteristics of current mobile traffic. The widespread of smart mobile terminals and the rapid development of mobile internet applications have brought significant diversity in types and characteristics of mobile network services. Traditional voice-based services have evolved into various types of data services such as instant messaging, multimedia videos and so on. Different from voice services, the traffic distribution of these new mobile Internet applications have obvious of heterogeneity, burstiness, and heavy-tailed characteristics [13]. The temporal and spatial variation of the traffic distribution is more prominent in current cellular mobile networks [13]. Therefore, the SMG analysis based on the temporal traffic distribution of voice services is not suitable for current mobile networks. In this paper, we propose a model of statistical multiplexing gain of processing resources in centralized radio access networks and make detailed analysis for the SMG by considering both the temporal and spatial traffic distribution characteristics.

The main contributions of this paper are as follows:

- We give a general definition of statistical multiplexing gain of processing resources in centralized radio access networks. The SMG is expressed by the ratio between the maximum value of the temporal-spatial data traffic over time in the area and the maximum value of the average traffic value of all cells over time in the area. By considering the service threshold ratio P_{th} in practical cellular networks, we further define a specific expression of SMG as a function of the P_{th}. The SMG for practical networks is then expressed by the ratio between the maximum traffic that can be serviced with the service threshold ratio P_{th} at any time and the maximum value of the average traffic value of all cells over time in the area.
- We propose a method to compute statistical multiplexing gain of processing resources in centralized radio access networks by considering the traffic distribution both in time and space domains, i.e., temporal-spatial joint traffic distribution, instead of only considering the temporal traffic distribution in most of the current research. The SMG is analyzed based on two representative spatial

traffic distribution models which are lognormal distribution and Alpha-Stable distribution. We also derive a closed-form approximation for SMG when the spatial traffic distribution is lognormal distribution. The numerical SMG can be obtained by the closed-form approximation to provide insight on engineering practice of cellular networks.
- We provide a detailed analysis on the SMG of processing resources in centralized radio access networks. We show numerically the influence of various factors including the service threshold ratio P_{th} and distribution parameters. Both the theoretical analysis and simulation results show that the SMG increases with P_{th} for all typical area types: CDB, park, and campus. We present the numerical results for the theoretical SMG and simulated SMG based on the lognormal and Alpha-Stable spatial traffic distribution models. The gap between the simulation and theoretical results is less than 5% for all, which verifies the SMG analysis proposed in this paper.

The remaining of this paper is organized as following: In Section II, the model of statistical multiplexing gain of processing resources in centralized radio access networks is proposed. Section III focuses on analysis of the SMG of processing resources in centralized radio access networks for temporal-spatial joint traffic distribution model based on lognormal distribution. Then analysis of the SMG for temporal-spatial joint traffic distribution model based on Alpha-Stable distribution is made in Section IV. In Section V, simulations are carried out to compare with corresponding theoretical SMG values. In the end, we summarize the paper in Section VI.

II. SYSTEM MODEL

Considering that in cellular mobile networks, the number of cells in the target region is N. In the centralized radio access networks architecture, N RRUs of all cells are connected to a centralized processing pool through an optical fiber fronthaul link. For the convenience of narrative, we use T to represent the data traffic statistics period of the network, and N to represent the number of cells in spatial extent of the network. Random variable $X(i,t)(i = 1, 2, ..., N; t = 1, 2, ..., T)$ is a traffic matrix. The matrix element $x(i,t)$ represents the traffic value of cell i at time t, which denotes the traffic load of cell i from time $(t-1)\Delta t$ to $t\Delta t$. $X(i,t)$ is defined as follows:

$$X(i,t) = \begin{bmatrix} x(1,1)\, x(1,2) \ldots x(1,t) \ldots x(1,T) \\ x(2,1)\, x(2,2) \ldots x(2,t) \ldots x(2,T) \\ \ldots \\ x(i,1)\, x(i,2) \ldots x(i,t) \ \ldots x(i,T) \\ \ldots \\ x(N,1) x(N,2) \ldots x(N,t) \ldots x(N,T) \end{bmatrix} \tag{1}$$

Each row vector $X(i,:)$ of matrix $X(i,t)$ represents the traffic values of a specific base station i in the time domain, while each column vector $X(:,t)^T$ of matrix $X(i,t)$ represents the traffic values of N base stations in the space domain at specific time t. That is, the matrix $X(i,t)$ characterizes the traffic distribution in both time and space domains. Assume that in cellular mobile networks, the relationship between processing resources allocated for a base station and the data traffic to be processed by the base station is linear with coefficient ξ [14]. The processing resources required by base station i is $\xi x(i,t)$ at time t.

In a Distributed RAN architecture, each base station needs to be allocated with enough resources to process the maximum traffic value that the corresponding cell may produce at any time t. Therefore, processing resources for each base station is determined by $\max_t \xi x(i,t)$, the required processing resources in distributed radio access networks are the summation of the maximum processing resources of every base station. In order to meet the network processing capability, the total processing resources should be allocated as follows:

$$R_{DRAN} = \sum_{i=1}^{N} \max_t \xi x(i,t) \tag{2}$$

If the traffic for all the cells is similar, the required processing resources can be given by:

$$R_{DRAN} = N \max_{i,t} \xi x(i,t) \tag{3}$$

In a Centralized RAN architecture, because the processing resources are shared among different BSs, the resources should be allocated as a whole, not for separate BSs one by one as in a Distributed RAN architecture. Therefore, in order to satisfy the network processing capability at time t, processing resources should be allocated as $\sum_{i=1}^{N} \xi x(i,t)$, which is the sum of traffic values generated by all cells at time t. So, the required processing resources of the centralized radio access networks should be determined by the maximum sum of traffic values generated by all cells that may occur at any time. Therefore, the required processing resources should be:

$$R_{CRAN} = \max_t \sum_{i=1}^{N} \xi x(i,t) \tag{4}$$

The ratio between the processing resources required in distributed radio access networks and the processing resources required in centralized radio access networks $\frac{R_{DRAN}}{R_{CRAN}}$ represents the reduction of required processing resources for the centralized radio access networks. We refer to this ratio as the statistical multiplexing gain (SMG) of processing resources in centralized radio access networks, which can be expressed as follows:

$$SMG = \frac{R_{DRAN}}{R_{CRAN}} = \frac{N \max_{i,t} \xi x(i,t)}{\max_t \sum_{i=1}^{N} \xi x(i,t)} = \frac{N \max_{i,t} x(i,t)}{\max_t \sum_{i=1}^{N} x(i,t)} \tag{5}$$

Further, the SMG can be expressed as follows:

$$SMG = \frac{\max_{i,t} x(i,t)}{\max_t \frac{\sum_{i=1}^{N} x(i,t)}{N}} \tag{6}$$

where $\max_{i,t} x(i,t)$ represents the maximum value of the temporal-spatial data traffic over time in the area. It corresponds to the maximum element in the traffic matrix $X(i,t)$. In (6), $\max_{t} \frac{\sum_{i=1}^{N} x(i,t)}{N}$ represents the maximum value of the average traffic value of all cells over time in the area. Let us use $m(t)$ to denote the average traffic value of all cells at time t in the area, expressed as:

$$m(t) = \frac{\sum_{i=1}^{N} x(i,t)}{N} \tag{7}$$

Thus, the SMG can be expressed as:

$$SMG = \frac{\max_{i,t} x(i,t)}{\max_{t} m(t)} \tag{8}$$

In practical cellular networks, the processing resources of a BS are usually configured to satisfy a very high percentage of required resources, rather than always satisfy the processing requirements of the BS maximum traffic at any time because of the deployment cost. In other words, there is a service threshold ratio P_{th}, specified by practical systems. For a given t, let us denote the service traffic threshold corresponding to the service threshold ratio P_{th} by $x_{th}(t, P_{th})$. Here, $x_{th}(t, P_{th})$ is the maximum traffic that can be serviced with the service threshold ratio P_{th} at time t. Then we have:

$$P_{th} = Pr(x(i,t) \leq x_{th}(t, P_{th})) \tag{9}$$

where P_{th} can be set according to the actual network scenario and engineering experience, e.g. 95%, 97%, 99%, etc.

Therefore, for a given P_{th}, the SMG can be expressed as:

$$SMG(P_{th}) = \frac{\max_{t} x_{th}(t, P_{th})}{\max_{t} m(t)} \tag{10}$$

Due to the relative stability of the total number of mobile users in the entire area and the uneven distribution of mobile users in spatial locations, the traffic of a single cell has strong non-uniformity both in time and space domain [15]. Obtaining the temporal-spatial joint distribution of traffic in cellular mobile networks is critical for the analysis of SMG. There are two representative spatial traffic distribution, i.e., lognormal distribution [15]–[17] and Alpha-Stable distribution [19], for current cellular mobile networks. Therefore, Section III will give an analysis of SMG based on lognormal distribution, and Section IV will give an analysis of SMG based on Alpha-Stable distribution.

III. ANALYSIS OF STATISTICAL MULTIPLEXING GAIN BASED ON LOGNORMAL DISTRIBUTION

In order to analyze the statistical multiplexing gain, we need to know the temporal-spatial distribution of traffic and the average traffic value of all cells over time in the area. The temporal traffic distribution model describes the traffic characteristics of cellular networks over time. Generally, the temporal traffic distribution model exhibits strong periodicity, in which the busy time and idle time are clear during a day. For example, in the working hours, a large number of users move from the residential area to the office area, and during the non-working hours, a large number of users return to the residential area from the office area. As these users move, the traffic load of the cellular networks also shows a phenomenon of migration, known as tidal effect in the network over time. Wang *et al.* [15] analyzed the traffic value of base stations in time domain, and found that mobile users have repetitive behavior during the time period in one day. A sine signal superposition model was proposed to fit the real traffic distribution in time domain for cellular networks. The model is given by [15]:

$$m(t) = a_0 + \sum_{k=1}^{n} a_k \sin(w_k t + \varphi_k) \tag{11}$$

where $m(t)$ is the average traffic value of all cells over time in the area, a_0 is a constant, w_k is the frequency components of traffic variation, n is the number of frequency components, a_k and φ_k are the amplitudes and phases, respectively, of the corresponding traffic variation.

In general, traffic is usually not uniformly distributed in real networks because of the convergence of user social behavior. The spatial distribution of traffic in cellular networks is geographically dependent. Users in different areas show different behavior characteristics, resulting in uneven distribution of traffic in the spatial domain. The work in [15]–[17] illustrated that the spatial inhomogeneity of traffic in cellular networks can be described by a lognormal distribution. The traffic density, which is regarded as real traffic demand of users, can be approximated by a lognormal distribution [17]. Therefore, statistical multiplexing gain will be analyzed based on this representative spatial traffic distribution model.

A lognormal distribution is a continuous probability distribution of a random variable whose logarithm is normally distributed. The probability density function (PDF) of the lognormal distribution is:

$$f_X(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left\{-\frac{(\ln x - \mu)^2}{2\sigma^2}\right\}, \quad (x > 0) \tag{12}$$

where μ and σ are the mean and standard deviation of the logarithm of the variable, respectively. The cumulative distribution function (CDF) of the lognormal distribution is:

$$F_X(x) = \Phi\left\{-\frac{(\ln x) - \mu}{\sigma}\right\} \tag{13}$$

Based on the lognormal model for the spatial traffic distribution, the parameter μ of lognormal distribution at time t can be computed with the following expression [15]:

$$\mu(t) = \ln(m(t)) - \frac{1}{2}\sigma^2 \tag{14}$$

Formula (14) is based on the relationship between characteristic parameters (mean and variance) of a lognormal random variable and characteristic parameters (mean and standard deviation) of the associated normal distribution, expressed as:

$$\mu = \ln(m^2 / \sqrt{\upsilon + m^2}) \tag{15}$$

$$\sigma = \sqrt{\ln(\upsilon/m^2 + 1)} \tag{16}$$

where m and υ are the mean and variance of a lognormal random variable, respectively, μ and σ are the mean and standard deviation of the associated normal distribution, respectively.

According to the temporal traffic distribution model formula given by (11), a temporal-spatial traffic distribution modeling approach is established by using lognormal distribution with parameters $\mu(t)$ and σ as follows:

$$\begin{aligned} & f_X\,(x(i,t); \mu, \sigma) \\ &= f_X\left(x(i,t); \ln(m(t)) - \frac{1}{2}\sigma^2, \sigma\right) \\ &= \frac{1}{x(i,t)\sigma\sqrt{2\pi}} \exp\left\{-\frac{\left(\ln x(i,t) - \ln(m(t)) + \frac{1}{2}\sigma^2\right)^2}{2\sigma^2}\right\} \end{aligned} \tag{17}$$

where $x(i, t)$ represents the traffic value of cell i at time t in the area.

The cumulative distribution function(CDF) of $x(i, t)$ is:

$$\begin{aligned} F_X(x(i,t)) &= \int_{-\infty}^{X} f_X\,(x(i,t))\,dX \\ &= \frac{1}{2}(1 + erf(\frac{\ln x(i,t) - \ln(m(t)) + \frac{1}{2}\sigma^2}{\sqrt{2}\sigma})) \end{aligned} \tag{18}$$

where erf is the error function, which is defined as:

$$erf\,(x) = \frac{1}{\sqrt{\pi}} \int_{-x}^{x} e^{-t^2} dt = \frac{2}{\sqrt{\pi}} \int_{0}^{x} e^{-t^2} dt \tag{19}$$

According to (9) and (18), the service threshold ratio P_{th} can be expressed as:

$$\begin{aligned} P_{th} &= F_X(x_{th}(t, P_{th})) \\ &= \frac{1}{2}(1 + erf(\frac{\ln x_{th}(t, P_{th}) - \ln(m(t)) + \frac{1}{2}\sigma^2}{\sqrt{2}\sigma})) \end{aligned} \tag{20}$$

Then, we have:

$$ln\frac{x_{th}(t, P_{th})}{m(t)} = \sqrt{2}\sigma erf^{-1}(2P_{th} - 1) - \frac{1}{2}\sigma^2 \tag{21}$$

Thus, we can get the expression of $x_{th}(t, P_{th})$ as:

$$x_{th}(t, P_{th}) = m(t)\exp(\sqrt{2}\sigma erf^{-1}(2P_{th} - 1) - \frac{1}{2}\sigma^2) \tag{22}$$

where $erf^{-1}(2P_{th} - 1)$ returns the value of the inverse error function for $2P_{th} - 1$.

Therefore, considering the threshold ratio P_{th}, based on (10) and (22), the SMG of processing resources in centralized radio access networks is further expressed as:

$$\begin{aligned} SMG(P_{th}) &= \frac{\max\limits_t x_{th}(t, P_{th})}{\max\limits_t m(t)} \\ &= \frac{\max\limits_t m(t)\exp(\sqrt{2}\sigma erf^{-1}(2P_{th} - 1) - \frac{1}{2}\sigma^2)}{\max\limits_t m\,(t)} \\ &= \exp(\sqrt{2}\sigma erf^{-1}(2P_{th} - 1) - \frac{1}{2}\sigma^2) \end{aligned} \tag{23}$$

For the inverse error function, an approximation is given by [18]:

$$\begin{aligned} erf^{-1}(x) \approx \Bigg[&- \frac{100}{7\pi} - \frac{\ln(1 - x^2)}{2} \\ &+ \sqrt{\left(\frac{100}{7\pi} + \frac{\ln(1 - x^2)}{2}\right)^2 - \frac{50}{7}\ln(1 - x^2)}\Bigg]^{1/2} \end{aligned} \tag{24}$$

the relative precision of this approximation is better than $4 \cdot 10^{-3}$, uniformly for all real x in the interval (0, 1).

Therefore, according to (23) and (24), the SMG can be expressed as (25), shown at the bottom of this page. It can be seen that the SMG of processing resources in centralized radio access networks is related to the service threshhold ratio P_{th} and the parameter σ of lognormal distribution. A reasonable service threshold ratio is usually set as $P_{th} = 95\% \sim 100\%$ for the practical deployments of cellular networks. The corresponding numerical SMG can be obtained by closed-form approximation (25) to provide insight on engineering practice of cellular networks.

IV. ANALYSIS OF STATISTICAL MULTIPLEXING GAIN BASED ON ALPHA-STABLE DISTRIBUTION

The work in [19] and [20] showed that the Alpha-Stable distribution model is suitable to characterize the spatial inhomogeneity of traffic in cellular networks. The Alpha-Stable model is able to capture the appropriate level of burstiness of different types of traffic by selecting the proper parameters [21]. It could most precisely fit the actual traffic spatial

$$\begin{aligned} SMG(P_{th}) \approx \exp\Bigg(\sqrt{2}\sigma\Bigg[&- \frac{100}{7\pi} - \frac{\ln(1 - (2P_{th} - 1)^2)}{2} \\ &+ \sqrt{\left(\frac{100}{7\pi} + \frac{\ln(1 - (2P_{th} - 1)^2)}{2}\right)^2 - \frac{50}{7}\ln(1 - (2P_{th} - 1)^2)}\Bigg]^{1/2} - \frac{1}{2}\sigma^2\Bigg) \end{aligned} \tag{25}$$

distribution in cellular networks, reflecting the basic characteristics of traffic demands from users, such as self-similarity and long-range dependence, and partially exhibit the nature of human activities [22]. With the property of burstiness and heavy-tailed distribution, Alpha-Stable model manifests itself in the capability to characterize the distribution of normalized sums of a relatively large number of independent identically distributed random variables [19].

In the previous section, we use random variable $x(i, t)$ to represent the traffic value of cell i at time t. In this section, we consider Alpha-Stable distribution to characterize the traffic model in cellular networks. Since this model does not specify the traffic in each cell, rather it specifies the traffic model for each spatial sampling area. We need to redefine the traffic matrix and corresponding SMG in this section. Considering that in cellular mobile networks, the target region is divided into M sampling areas with the same size, and the coverage area of a sampling area is A. Let us define the traffic matrix as $X_a(j, t)(j = 1, 2, ..., M; t = 1, 2, ..., T)$. The element $x_a(j, t)$ represents the traffic value of the spatial sampling area j at time t, which denotes the traffic load of the sampling area j from time $(t-1)\Delta t$ to $t\Delta t$. In this scenario, similar to the discussion in Section II, the SMG can be expressed as:

$$SMG = \frac{\max\limits_{j,t} x_a(j,t)}{\max\limits_{t} \frac{\sum_{j=1}^{M} x_a(j,t)}{M}} \tag{26}$$

where $\max\limits_{j,t} x_a(j, t)$ represents the maximum value of the temporal-spatial data traffic over time in the region. It corresponds to the maximum element in the traffic matrix $X_a(j, t)$. Let us use $m(t)$ to denote the average traffic value of all cells at time t in the area, and n_o is the average number of cells in a sampling area. Thus, we have:

$$\frac{\sum_{j=1}^{M} x_a(j,t)}{M} = m(t)n_o \tag{27}$$

So, the SMG can be expressed as:

$$SMG = \frac{\max\limits_{j,t} x_a(j,t)}{\max\limits_{t} m(t)n_o} \tag{28}$$

Since the expression of PDF $f_X(x)$ is unknown in closed-form for most stable distributions, Alpha-Stable distribution is generally specified by its characteristic function. A random variable X is said to obey the Alpha-Stable distribution: $X \sim S(\alpha, \beta, \gamma, \delta)$, if there are parameters $\alpha \in (0, 2]$, $\beta \in [-1, 1]$, $\gamma \in [0, +\infty)$, $\delta \in R$, such that its characteristic function is of the following form [23]:

$$\Phi(\omega) = \begin{cases} \exp\{-\gamma^\alpha|\omega|^\alpha(1 - i\beta sign(\omega)\tan(\frac{\pi\alpha}{2})) + i\delta\omega\}, \\ \qquad\qquad\qquad\qquad\qquad\qquad \alpha \neq 1 \\ \exp\{-\gamma|\omega|(1 + i\beta sign(\omega)\ln|\omega|) + i\delta\omega\}, \\ \qquad\qquad\qquad\qquad\qquad\qquad \alpha = 1 \end{cases} \tag{29}$$

where α is the characteristic exponent and indicates the index of stability, β is identified as the skewness parameter, α and β together determine the shape of the model, γ and δ are scale and shift parameters respectively. An Alpha-Stable distribution can be described by only four parameters. The PDF of the Alpha-Stable distribution is the inverse Fourier transform of the characteristic function, shown as:

$$f_X(x) = F^{-1}[\Phi(w)] = \frac{1}{2\pi}\int_{-\infty}^{\infty} \Phi(w)e^{iwx}dw \tag{30}$$

The CDF of the Alpha-Stable distribution is:

$$F_X(x) = \int_{-\infty}^{x} f_X(y)dy \tag{31}$$

The parameter δ of Alpha-Stable distribution indicates the shift of the PDF and it equals to the mean of the variable when $\alpha \in (0, 1)$. The PDF $f_X(x)$ for some negative interval is non-zero and then it makes no sense for the traffic of cellular networks when $\delta < 0$. Therefore, we consider the non-negative interval of the variable X and normalize the PDF for such an interval with practical meaning by defining a Truncated Alpha-Stable distribution: $X \sim \overline{S}(\alpha, \beta, \gamma, \delta)$. The PDF of the Truncated Alpha-Stable distribution is:

$$\overline{f_X}(x) = \begin{cases} \frac{f_X(x)}{\int_{y\geq 0} f_X(y)dy}, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{32}$$

The CDF of the Truncated Alpha-Stable distribution is:

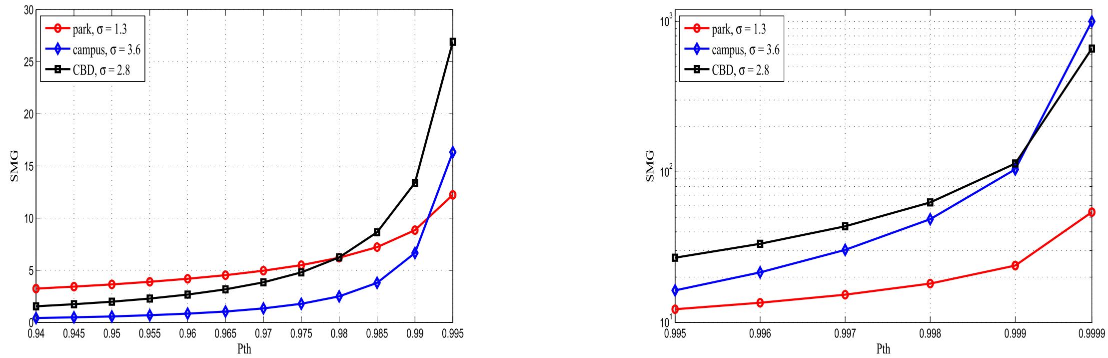
$$\overline{F_X}(x) = \int_{0}^{x} \overline{f_X}(y)dy \tag{33}$$

In an Alpha-Stable model for the spatial traffic distribution in cellular networks, for a sampling area $j(j = 1, 2, ..., M)$ at time $t(t = 1, 2, ..., T)$. Let $\lambda_{BS}(j, t)$ denote the density of base stations and $\lambda_{TR}(j, t)$ denote the density of spatial traffic for sampling area j at time t, both $\lambda_{TR}(j, t)$ and $\lambda_{BS}(j, t)$ follow the Alpha-Stable distribution. In addition, a linear regression can characterize $\lambda_{BS}(j, t)$ and $\lambda_{TR}(j, t)$, which can be stated as [19]:

$$\lambda_{BS}(j, t) = k\lambda_{TR}(j, t) \tag{34}$$

where k is a linear coefficient representing the number of base stations required per unit spatial traffic. The parameters of the Alpha-Stable distribution of λ_{BS} are fixed values for a same area [17]. Thus, the temporal-spatial traffic model can be established by using the Truncated Alpha-Stable distribution as follows:

$$x_a(j, t) = Am(t)\lambda_{BS}(j, t) \tag{35}$$



(a) (b)

FIGURE 3. Theoretical SMG as function of P_{th} (Lognormal distribution). (a) $P_{th} = 0.94 - 0.995$. (b) $P_{th} = 0.995 - 0.9999$.

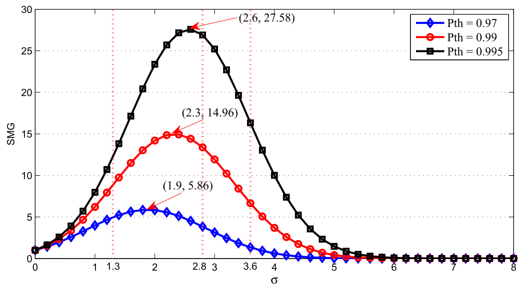


FIGURE 4. Theoretical SMG as function of σ (Lognormal distribution).

According to the linear nature of the Alpha-Stable distribution [24], if $X \sim \overline{S}(\alpha, \beta, \gamma, \delta)$, we have:

$$bX \sim \begin{cases} \overline{S}(\alpha, sign(b)\beta, |b|\,\gamma, b\delta), & \alpha \neq 1 \\ \overline{S}(1, sign(b)\beta, |b|\,\gamma, b\delta - \frac{2}{\pi}b(\ln|b|)\beta\gamma), & \alpha = 1 \end{cases} \tag{36}$$

Thus, for $\lambda_{BS}(j,t) \sim \overline{S}(\alpha_0, \beta_0, \gamma_0, \delta_0)$, we have $x_a(j,t) \sim \overline{S}(\alpha(j,t), \beta(j,t), \gamma(j,t), \delta(j,t))$:

$$\begin{aligned} x_a(j,t) &= Am(t)\lambda_{BS}(j,t) \\ &\sim \begin{cases} \overline{S}(\alpha_0, sign(Am(t))\beta_0, |Am(t)|\,\gamma_0, Am(t)\delta_0), \\ \qquad\qquad \alpha_0 \neq 1 \\ \overline{S}(1, sign(Am(t))\beta_0, |Am(t)|\,\gamma_0, Am(t)\delta_0 \\ -\frac{2}{\pi}Am(t)(\ln|Am(t)|)\beta_0\gamma_0), \quad \alpha_0 = 1 \end{cases} \end{aligned} \tag{37}$$

Similar to the discussions in Section III, for a given service threshold ratio P_{th}, the SMG of processing resources in centralized radio access networks can be expressed as:

$$\begin{aligned} SMG(P_{th}) &= \frac{\max_t x_{ath}(t, P_{th})}{\max_t m(t)n_o} \\ &= \frac{\max_t \overline{F}_{X_a}^{-1}(t, P_{th})}{\max_t m(t)\, n_o} \end{aligned} \tag{38}$$

Based on formula (30) to (33), we have:

$$\overline{F}_{X_a}^{-1}(t, P_{th}) = F_{X_a}^{-1}(t, P_{th} + (1 - P_{th})F_{X_a}(0,t)) \tag{39}$$

Therefore, the SMG can be expressed as:

$$SMG(P_{th}) = \frac{\max_t F_{X_a}^{-1}(t, P_{th} + (1 - P_{th})F_{X_a}(0,t))}{\max_t m(t)n_o} \tag{40}$$

It can be seen that the SMG of processing resources in centralized radio access networks is related to the threshhold ratio P_{th}, the maximum value of average traffic $m(t)$ of all cells in the area, the average number of cells n_0 in a sampling area, and the Truncated Alpha-Stable distribution parameters $(\alpha_0, \beta_0, \gamma_0, \delta_0)$ of the density of base stations. Since the expression of PDF $f_X(x)$ is unknown in closed-form for Alpha-Stable distributions, we will evaluate it numerically in the following section.

V. SIMULATION RESULTS

In this section, we present the numerical results of the theoretical SMG values and simulated SMG values based on the temporal-spatial joint traffic distribution models. For the temporal-spatial joint traffic distribution model based on the lognormal distribution, we select three typical area types: central business district (CBD), park, and campus. Each area type has a typical standard deviation of its traffic distribution. The lognormal distribution parameter σ of them has been obtained in [15]. For the temporal-spatial joint traffic distribution model based on Alpha-Stable distribution, in order to utilize the temporal traffic distribution of the whole area given in [15], we consider the spatial traffic distribution in the whole area with parameters specified in Table 1.

For lognormal distribution, the theoretical values of SMG is shown as Fig. 3. It can be seen from the Fig. 3 that the SMG increases with P_{th} for all typical area types: CBD, park, and campus. This is because the larger P_{th}, the more processing resources needed to be configured to support high traffic load of cells in DRAN architecture. In centralized RAN, processing resources can be shared among different BSs. When a BS is in a high traffic load, more processing

TABLE 1. **Values of parameters.**

Parameter	Value
N	1000
T	24h
Δt	1h
σ(CBD)	2.8
σ(park)	1.3
σ(campus)	3.6
P_{th}	(0.94, 0.95, 0.96, 0.97, 0.98, 0.99, 0.999, 0.9999)
$m(t)$(CBD)	$75.72+47.52*\sin(\frac{\pi}{12}t-2.56)+16.71*\sin(\frac{\pi}{6}t+1.45)$
$m(t)$(park)	$351.06+222.7*\sin(\frac{\pi}{12}t+3.11)+96.24*\sin(\frac{\pi}{6}t+2.36)$
$m(t)$(campus)	$323.04+148.3*\sin(\frac{\pi}{12}t+2.98)+109.4*\sin(\frac{\pi}{6}t+2.15)+38.43*\sin(\frac{\pi}{4}t+1)$
$m(t)$(whole area)	$173.29+89.83*\sin(\frac{\pi}{12}t+3.08)+52.6*\sin(\frac{\pi}{6}t+2.08)+16.68*\sin(\frac{\pi}{4}t+1.13)$
A	$1km^2$
n_o	15
$\alpha_o, \beta_o, \gamma_o, \delta_o$	(1.79, 1.00, 6.33, 15.37)

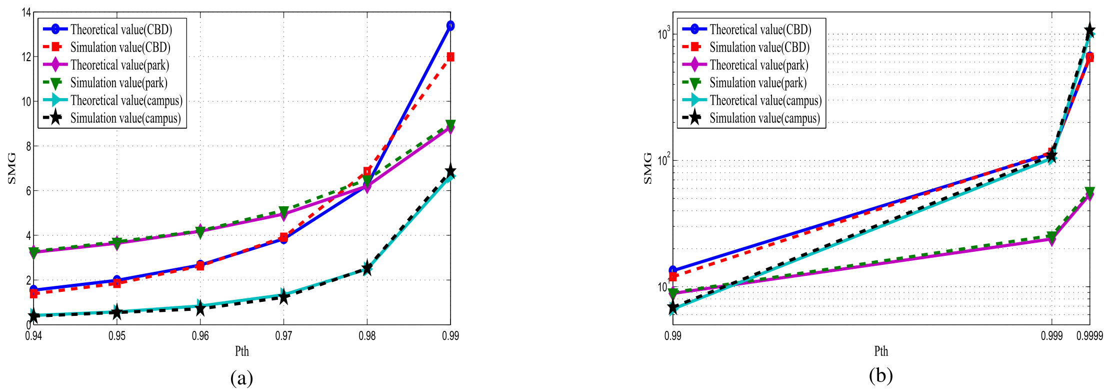


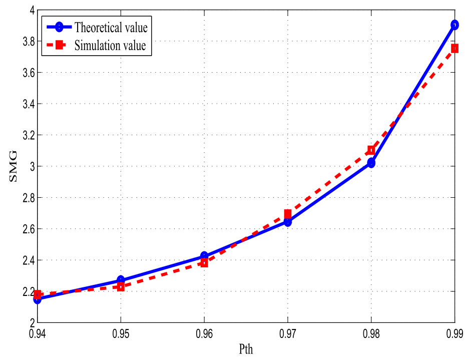
FIGURE 5. **Simulated versus theoretical SMG as function of P_{th} (Lognormal distribution). (a) $P_{th} = 0.94 - 0.99$. (b) $P_{th} = 0.99 - 0.9999$.**

resources are allocated to it. When its traffic load is low, the redundant resources could be dynamically re-assigned to other BSs with a higher traffic load. Therefore, P_{th} does not affect the resource configuration in centralized RAN significantly. Moreover, when P_{th} is fixed, the SMG is different in CBD, park and campus. The reason is that the standard deviations of their traffic distribution are different, which will be discussed in detail in Fig. 4.

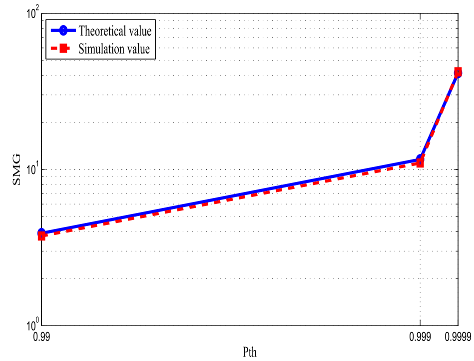
Fig. 4 shows the relationship between SMG and σ when the P_{th} is fixed as 0.97, 0.99, 0.995, respectively. It can be seen that the SMG gets the maximum value (5.86, 14.96, 27.58) when $\sigma = 1.9, 2.3, 2.6$. The SMG increases with σ when σ is less than 1.9, 2.3, 2.6, respectively, and SMG decreases with σ when σ is larger than 1.9, 2.3, 2.6, respectively.

Fig. 5 illustrates the SMG simulation and theoretical results over different P_{th} for three typical areas: CBD, park and campus. The gap between the simulation and theoretical values for three typical areas is less than 5%, which verifies the theory analysis of SMG in above sections.

For Alpha-Stable distribution, Fig. 6 shows theoretical values of SMG and the simulated temporal-spatial joint traffic distribution of 1000 base stations over time of one day in the target area. Fig. 6 illustrates simulated SMG values and theoretical SMG values over different P_{th} values. The relative error between the simulated SMG values and the theoretical SMG values is less than 4%, which verifies the correctness of analysis. In addition, it can be seen from the Fig. 6 that the SMG increases with P_{th}. The reason is consistent with the SMG based on the lognormal distribution.



(a)



(b)

FIGURE 6. Simulated versus theoretical SMG as function of P_{th} (Alpha-Stable distribution). (a) $P_{th} = 0.94 - 0.99$. (b) $P_{th} = 0.99 - 0.9999$.

However, when the P_{th} is 0.9999, the SMG is much less than the SMG based on the lognormal distribution. It shows that different spatial traffic distributions have a great influence on SMG.

VI. CONCLUSIONS

In this paper, we investigated the SMG of processing resources in centralized radio access networks by considering the temporal and spatial distribution of data traffic in current cellular networks. We also proposed a model of SMG and made detailed analysis for it by using the model of statistical multiplexing gain based on the temporal-spatial joint traffic distribution model. We found that the SMG is related to the service threshold ratio P_{th} and certain spatial traffic distribution parameters, for example, σ of lognormal spatial traffic distribution, $(\alpha, \beta, \gamma, \delta)$ of Alpha-Stable spatial traffic distribution. Moreover, the SMG is related to the maximum average traffic value $m(t)$ of all cells in the target area when the spatial traffic distribution is Alpha-Stable distribution. In general, the SMG increases with P_{th}. We compared simulated SMG values with corresponding theoretical SMG values for three typical area types of park, campus and CBD over different values of the service threshold ratio P_{th}, respectively. The gap between the simulated SMG values and the theoretical SMG values is less than 5% for all. Overall, the simulation results prove the correctness of our theoretical analysis.

REFERENCES

[1] "Cisco visual networking index: Global mobile data traffic forecast update, 2015–2020," Cisco, San Jose, CA, USA, White Paper 1454457600805266, Feb. 2016.

[2] L. Liu, Y. Zhou, V. Garcia, L. Tian, and J. L. Shi, "Load aware joint CoMP clustering and inter-cell resource scheduling in heterogeneous ultra dense cellular networks," *IEEE Trans. Veh. Technol.*, vol. 67, no. 3, pp. 2741–2755, Mar. 2018.

[3] Y. Zhou, H. Liu, Z. Pan, L. Tian, and J. Shi, "Energy-efficient two-stage cooperative multicast: Effect of user density," *IEEE Trans. Veh. Technol.*, vol. 65, no. 9, pp. 7297–7307, Sep. 2016.

[4] L. Liu, Y. Zhou, L. Tian, and J. Shi, "CPC-based backward-compatible network access for LTE cognitive radio cellular networks," *IEEE Commun. Mag.*, vol. 53, no. 7, pp. 93–99, Jul. 2015.

[5] "Toward 5G C-RAN: Requirements, Architecture and Challenges," China Mobile Research Institute, White Paper v1.0, Nov. 2016.

[6] G. Zhai, L. Tian, Y. Zhou, and J. Shi, "Load diversity based optimal processing resource allocation for super base stations in centralized radio access networks," *Sci. China Inf. Sci.*, vol. 57, no. 4, pp. 1–12, Apr. 2014.

[7] M. Qian, Y. Wang, Y. Zhou, L. Tian, and J. Shi, "A super base station based centralized network architecture for 5G mobile communication systems," *Digit. Commun. Netw.*, vol. 1, no. 2, pp. 152–159, 2014.

[8] J. Liu, S. Zhou, J. Gong, Z. Niu, and S. Xu, "On the statistical multiplexing gain of virtual base station pools," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Austin, TX, USA, Dec. 2014, pp. 2283–2288.

[9] J. Liu, S. Zhou, J. Gong, Z. Niu, and S. Xu, "Statistical multiplexing gain analysis of heterogeneous virtual base station pools in cloud radio access networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 8, pp. 5681–5694, Aug. 2016.

[10] L. Wang and S. Zhou, "On the fronthaul statistical multiplexing gain," *IEEE Commun. Lett.*, vol. 21, no. 5, pp. 1099–1102, May 2017.

[11] M. Madhavan, P. Gupta, and M. Chetlur, "Quantifying multiplexing gains in a Wireless Network Cloud," in *Proc. IEEE Conf. Commun. (ICC)*, Ottawa, Cnanda, Jun. 2012, pp. 3212–3216.

[12] A. Checko, H. Holm, and H. Christiansen, "Optimizing small cell deployment by the use of C-RANs," in *Proc. 20th Eur. Wireless Conf. (EW)*, 2014, pp. 1–6.

[13] U. Paul, A. P. Subramanian, M. M. Buddhikot, and S. R. Das, "Understanding traffic dynamics in cellular data networks," in *Proc. IEEE INFOCOM*, Shanghai, China, Apr. 2011, pp. 882–890.

[14] Z. Zhang, L. Tian, Y. Zhou, L. Liu, B. Sun, and J. Shi, "Energy efficient dynamic computing resource allocation in centralized radio access networks," in *Proc. IEEE Conf. Commun. (ICC)*, Kansas City, MO, USA, May 2018, pp. 1–6.

[15] S. Wang, X. Zhang, J. Zhang, J. Feng, W. Wang, and K. Xin, "An approach for spatial-temporal traffic modeling in mobile cellular networks," in *Proc. 27th Int. Teletraffic Congr.*, Ghent, Belgium, 2015, pp. 203–209.

[16] D. Lee, S. Zhou, and Z. Niu, "Spatial modeling of scalable spatially-correlated log-normal distributed traffic inhomogeneity and energy-efficient network planning," in *Proc. IEEE WCNC*, Apr. 2013, pp. 1285–1290.

[17] D. Lee, S. Zhou, X. Zhong, Z. Niu, X. Zhou, and H. Zhang, "Spatial modeling of the traffic density in cellular networks," *IEEE Wireless Commun.*, vol. 21, no. 1, pp. 80–88, Feb. 2014.

[18] S. Winitzki. (2008). Accessed: Jun. 2010. *A Handy Approximation for the Error Function and Its Inverse*. [Online]. Available: http://homepages.physik.uni-muenchen.de/Winitzki/erf-approx.pdf

[19] M. Li, Z. Zhao, Y. Zhou, X. Chen, and H. Zhang, "On the dependence between base stations deployment and traffic spatial distribution in cellular networks," in *Proc. 23rd Int. Conf. Telecommun. (ICT)*, Thessaloniki, Greece, 2016, pp. 1–5.

[20] R. Li, Z. Zhao, J. Zheng, C. Mei, Y. Cai, and H. Zhang, "The learning and prediction of application-level traffic data in cellular networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 3899–3912, Jun. 2016.

[21] J. R. Gallardo, D. Makrakis, and L. Orozco-Barbosa, "Use of alpha-stable self-similar stochastic processes for modeling traffic in broadband networks," *Proc. SPIE*, vol. 3530, pp. 1–6, Nov. 1998.

[22] Y. Zhou, R. Li, Z. Zhao, X. Zhou, and H. Zhang, "On the α-stable distribution of base stations in cellular networks," *IEEE Commun. Lett.*, vol. 19, no. 10, pp. 1750–1753, Oct. 2015.

[23] X. Ge, G. Zhu, and Y. Zhu, "On the testing for alpha-stable distributions of network traffic," in *J. Electron. (China)*, vol. 20, no. 4, pp. 309–312, 2003.

[24] M. Shao and C. L. Nikias, "Signal processing with fractional lower order moments: Stable processes and their applications," *Proc. IEEE*, vol. 81, no. 7, pp. 986–1010, Jul. 1993.

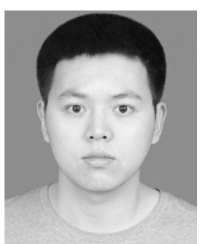
ZONGSHUAI ZHANG (S'13) received the B.E. degree in communication engineering from Tongji University, Shanghai, China, in 2011, and the M.E. degree in electronics and communication engineering from the University of Chinese Academy of Sciences, Beijing, China, in 2015, where he is currently pursuing the Ph.D. degree with the Institute of Computing Technology, Chinese Academy of Sciences. His research interests include resource management in the centralized radio access networks, and the convergence of communication and computing.

LIN TIAN (M'07–SM'17) received the B.S. and M.S. degrees in computer science from Beihang University, Beijing, China, in 2002 and 2005, respectively, and the Ph.D. degree in computer architecture from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, in 2012, where she is currently a Professor with the Wireless Communication Technology Research Center. She has published more than 50 research papers and three books or book chapters in the area of mobile communication and computing. She is also an Inventor of more than 40 Chinese patents and pending applications. She received the Best Paper Award from the ICC 2018, the ISCIT 2016, and the WCNC 2013. She was the Symposium Co-Chair of the ChinaCom 2013 and the Publication Chair of the ChinaCom 2012. She has also served as a Reviewer for a number of referred journals and international conferences.

JINGLIN SHI is currently the Director of the Wireless Communication Technology Research Center, Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, and a Visiting Professor with the Beijing University of Posts and Telecommunications, Beijing; The University of Sydney, Sydney, NSW, Australia; the University of Wollongong, Wollongong, NSW; and Macquarie University, Sydney. As a Team Leader, he successfully led the development of TD-SCDMA, WiMAX, and LTE protocol stack systems. He is currently responsible for several national projects in broadband wireless communications, including TDD-LTE baseband chip design and research on radio resource management techniques toward IMT-A. He has published two books and more than 100 papers in telecommunications journals and conference proceedings, and holds more than 30 granted patents. His research interests include wireless communications system architecture and management, wireless signal processing theory, and wireless communications baseband processor design. He was a member of the technical program committees of the IEEE WCNC, the IEEE ICC, the IEEE AusWireless 2006, the 7th IEEE ISCIT 2007, the ChinaCom 2007, and the ChinaCom 2009. He has also served on the organizing and technical committees of numerous national and international conferences. He was the General Co-Chair of the ChinaCom 2012.

JINHONG YUAN (M'02–SM'11–F'16) received the B.E. and Ph.D. degrees in electronics engineering from the Beijing Institute of Technology, Beijing, China, in 1991 and 1997, respectively. From 1997 to 1999, he was a Research Fellow with the School of Electrical Engineering, The University of Sydney, Sydney, Australia. In 2000, he joined the School of Electrical Engineering and Telecommunications, The University of New South Wales, Sydney, where he is currently a Professor and the Head of the Telecommunication Group. He has published two books, five book chapters, over 300 papers in telecommunications journals and conference proceedings, and 50 industrial reports. He has co-authored the papers, which received four best paper awards and the Best Poster Award, including the Best Paper Award from the IEEE International Conference on Communications, Kansas City, USA, in 2018, the Best Paper Award from the IEEE Wireless Communications and Networking Conference, Cancun, Mexico, in 2011, and the Best Paper Award from the IEEE International Symposium on Wireless Communications Systems, Trondheim, Norway, in 2007. He is a Co-Inventor of one patent on multi-in multi-out systems and two patents on low-density parity-check codes. His current research interests include error control coding and information theory, communication theory, and wireless communications. He is currently serving as an Associate Editor for the IEEE Transactions on Wireless Communications. He served as the IEEE NSW Chapter Chair for the Joint Communications/Signal Processions/Ocean Engineering Chapter, during 2011–2014, and as an Associate Editor for the IEEE Transactions on Communications, during 2012–2017.

YIQING ZHOU (S'03–M'05–SM'10) received the B.S. degree in communication and information engineering and the M.S. degree in signal and information processing from Southeast University, China, in 1997 and 2000, respectively, and the Ph.D. degree in electrical and electronic engineering from The University of Hong Kong, Hong Kong, in 2004. She is currently a Professor with the Wireless Communication Research Center, Institute of Computing Technology, Chinese Academy of Sciences. She has published over 100 papers and four books or book chapters in the areas of wireless mobile communications. She received the Best Paper Award from the ICC 2018, the ISCIT 2016, the IEEE PIMRC 2015, the ICCS 2014, and the WCNC 2013, and the 2014 Top 15 Editor Award from the IEEE Transactions on Vehicular Technology (TVT). She is the Associate/Guest Editor of the IEEE TVT, the IEEE JSAC (Special issue on Broadband Wireless Communication for High Speed Vehicles and Virtual Multi-In Multi-Out), WCMC, ETT, and JCST. She is also the TPC Co-Chair of the ChinaCom 2012, the Symposia Co-Chair of the IEEE ICC 2015, the Symposium Co-Chair of the GLOBECOM 2016 and the ICC 2014, the Tutorial Co-Chair of the ICCC 2014 and the WCNC 2013, and the Workshop Co-Chair of the SmartGridComm 2012 and the GlobeCom 2011.



XINYU CUI received the B.Sc. degree from the North University of China, Taiyuan, China, in 2018. He is currently pursuing the master's degree with the Institute of Computing Technology, Chinese Academy of Sciences, University of Chinese Academy of Sciences, Beijing, China. His research interests include the area of wireless networks, caching, and recommender systems.

LU WANG received the B.S. degree in communication engineering from Yunnan University, Kunming, China, in 2017. She is currently pursuing the master's degree with the Institute of Computing Technology, Chinese Academy of Sciences, University of Chinese Academy of Sciences, Beijing, China. Her research interests focuse on resource management in the centralized radio access networks, and massive multiple-input multiple-output wireless communication.

QIAN SUN (M'17) received the B.Sc. and M.Sc. degrees from Beijing Jiaotong University, Beijing, China, in 2012 and 2015, respectively. She is currently an Assistant Professor with the Wireless Communication Research Center, Institute of Computing Technology, Chinese Academy of Sciences, Beijing. Her research interests include the convergence of communications and computing and radio resource management in wireless communication.