On the α-Stable Distribution of Base Stations in Cellular Networks

Yifan Zhou, Rongpeng Li, Zhifeng Zhao, Xuan Zhou, and Honggang Zhang

***Abstract*—Cellular networks are now nearly universally deployed and are under ever-growing pressure to increase the volume of data deliverable to consumers. Understanding how base stations (BSs) are spatially deployed could prominently facilitate the performance analyses of cellular networks, as well as the design of efficient networking protocols. In this letter, inspired by the clustering reality of BSs and the intrinsic heavy-tailed characteristics of human activities, we aim to re-examine the statistical pattern of BSs in cellular networks, and find the most appropriate spatial density distribution of BSs. Interestingly, by taking advantage of large amount of realistic deployment information of BSs from on-operating cellular networks, we find that the widely adopted Poisson distribution severely diverges from the practical density distribution of BSs. Instead, heavy-tailed distributions could more precisely match the practical distribution. In particular, α-stable distribution, the distribution also found in traffic pattern of broadband networks and cellular networks, is most consistent with the practical one.**

***Index Terms*—Cellular networks, base stations, spatial density distribution, Poisson point process, heavy-tailed distributions, α-stable distribution.**

I. Introduction

CELLULAR networks are becoming an inevitable data pipe for diverse mobile devices to access the intense content on the Internet. Understanding how base stations (BSs) are spatially deployed, could prominently facilitate the performance analyses of cellular networks, as well as the design of efficient networking protocols. For example, Poisson distribution is widely adopted to characterize the spatial distribution of BSs, and leads to a tractable approach to calculate the coverage probability and rate distribution in cellular networks, by taking advantage of a Poisson point process (PPP) based theory (i.e., stochastic geometry) [1], [2]. However, the modeling accuracy of Poisson distribution has been recently questioned [3]. Consequently, in order to reduce the modeling error between Poisson distributed BSs and the practical distributed ones [4], some variants of PPP have been exploited to obtain precise analysis results. On the other hand, the actual deployment of BSs in long term is highly correlated with human activities [5], [6]. Humans tend to live together, and their social behaviors would lead to traffic hotspots [6], thus causing BSs to be more tensely deployed in certain areas as clusters. Furthermore, according to an assumption named “preferential attachment” [7], Barabási *et al.* argues that many large networks grow to be heavy-tailed. Therefore, heavy-tailed distributions appear to be more suitable to precisely characterize the clusteringly distributed BSs. In a nutshell, in spite of its apparent importance, there still does not exist convincing models for the spatial distribution of BSs in cellular networks.

TABLE I
THE DATASET OF BSS AND THE RELATED CITY INFORMATION

Attributes	City A	City B	City C
No. of BSs	8826	5746	4613
City Area	16,847 km^2	9,816 km^2	9,413 km^2
Population	8.844 million	7.639 million	6.038 million
Description	Inland Provincial Captial	Coastal	Coastal

Fortunately, there have already existed substantial works towards discovering the distribution of BSs in cellular networks from various practical measurements. In the earliest stages, a two-dimensional hexagonal grid model [1] was used and implied that BSs were spatially uniformly deployed, which is obviously far from the real scenarios. In next stages, Poisson distribution [1], [2] was found to be able to roughly match the realistic BS deployment in cellular networks, while providing tractable analysis conclusions. Meanwhile, due to the ever-increasing deployment of new BSs, cell sizes in on-operating cellular networks are becoming smaller and more irregular [8]. Therefore, variants of Poisson distributions (i.e., two-tier PPP [6] and Poisson clustering process [9]) are proposed, so as to better reflect the clustering property of BSs. However, these Poisson-based distributions lack the foundation to understand the intrinsic characteristics of BSs' evolution together with the traffic dynamics motived and impacted by the human social activities.

In this letter, we aim to re-examine the statistical pattern of BSs in cellular networks, and find the most accurate spatial density distribution of BSs. By taking advantage of large amount of realistic spatial deployment information of BSs from on-operating cellular networks, we compare the practical distribution of BSs in cellular networks with various representative candidates, including Poisson distribution and some other heavy-tailed distributions (e.g., generalized Pareto distribution, log-normal distribution, Weibull distribution, and α-stable distribution [10]). Interestingly, among the exploited distributions, α-stable distribution could most precisely fit the actual deployment of legacy BSs, which is also consistent with the traffic distribution in broadband and cellular networks [11], [12]. In other words, the spatial distribution of BSs in cellular

Manuscript received June 18, 2015; revised July 30, 2015; accepted August 11, 2015. Date of publication August 14, 2015; date of current version October 8, 2015. This letter is partially supported by the National Basic Research Program of China (973Green, No. 2012CB316000), the Key Project of Chinese Ministry of Education (No. 313053), the Key Technologies R&D Program of China (No. 2012BAH75F01), and the Program for Zhejiang Leading Team of Science and Technology Innovation (No. 2013TD20). The associate editor coordinating the review of this paper and approving it for publication was X. Zhou.

Y. Zhou, Z. Zhao, and H. Zhang are with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China (e-mail: zhouyftt@zju.edu.cn; zhaozf@zju.edu.cn; honggangzhang@zju.edu.cn).

R. Li and X. Zhou are with the Huawei Technologies Company, Ltd., Shanghai 201652, China (e-mail: lirongpeng@zju.edu.cn; zhouxuan@zju.edu.cn).

Digital Object Identifier 10.1109/LCOMM.2015.2468718

(a) (b) (c)

Fig. 1. An illustration of the deployment of base stations in three typical cities with geographical landforms. (a) City A; (b) City B; (c) City C.

networks reflects the basic characteristics of traffic demands from users, and could partially exhibit the nature of human activities. We believe that this new finding could contribute to the understanding of the evolution of cellular networks as well as the relevant society development.

The letter is organized as follows. In Section II, we firstly describe the details of the utilized practical datasets, and then present some necessary background of heavy-tailed distributions. In Section III, we provide fitting results for the spatial distribution of BSs. Finally, we conclude this paper with a summary and future research direction in Section IV.

II. Background

A. Data Description

In order to reach credible results, we collect a massive amount of practical data of BSs information from China Mobile in a well-developed eastern province of China. The collected dataset, containing over 47 000 BSs of GSM cellular networks and serving over 40 million subscribers, encompasses all BS-related records like location information (i.e. longitude, latitude, etc.) and BS type (i.e. macrocell or microcell).

Based on the coverage area and location information, we divide the dataset into disjoint subsets. Accordingly, we can classify the dataset as subsets of urban areas and rural areas, by matching the geographical landforms with local maps. In this letter, for simplicity of representation, we primarily take account of urban areas, and try to select one most precise spatial distribution for BS deployment from various well-known candidate models. Specifically, we choose three typical cities capable of fully reflecting the BS deployment phenomena of metropolis city, big city and medium city, respectively. In Table I, we have summarized the detailed information of these selected areas. Meanwhile, we plot the BS deployment with the geographical landforms in Fig. 1, which demonstrates that most BSs are densely clustered while some others are more sparsely deployed.

B. Mathematical Background

Heavy-tailed distributions could be widely applied to explain a number of natural phenomena, including the Internet topology [13]. Mathematically, heavy-tailed distributions are probability

TABLE II
The List of Candidate Distributions and Estimated Parameters in Fig. 2

Distribution	PDF	Estimated Parameters
Generalized Pareto (GP)	$\frac{1}{b}(1+\frac{a}{b}x)^{-(1+\frac{1}{a})}$	a=0.0488, b=3.3502
Weibull	$pqx^{q-1}e^{-px^q}$	p=0.7285, q=0.8279
Log-normal	$\frac{1}{\sqrt{2\pi}nx}e^{-\frac{(\ln x-m)^2}{2n^2}}$	m=-0.1835, n=1.0483
α-Stable	Closed form not always exists. Characteristic Function in Eq. (1).	α=0.6207, β=1.0000 σ=0.2053, μ=0.0658
Poisson	$\frac{\lambda^k}{k!}e^{-\lambda}$	λ=1.6759

distributions whose tails are not exponentially bounded. In other words, they have heavier tails than the exponential distribution.

There exist many statistical distributions proving to be heavy-tailed. Among them, generalized Pareto (GP) distribution, Weibull distribution, and log-normal distribution belong to one-tailed ones with the probability density function (PDF) in closed-forms (see Table II). Another famous heavy-tailed distribution is α-stable distribution, who manifests itself in the capability to characterize the distribution of normalized sums of a relatively large number of independent identically distributed random variables [10]. However, α-stable distribution, with few exceptions, lacks a closed-form expression of the PDF, and is generally specified by its characteristic function.

Definition 1: A random variable X is said to obey the α-stable distribution if there are parameters $0 < \alpha \leq 2$, $\sigma \geq 0$, $-1 \leq \beta \leq 1$, and $\mu \in \mathcal{R}$ such that its characteristic function is of the following form:

$$\begin{aligned}\phi(\omega) &= E(\exp j\omega X)\\ &= \exp\left\{-\sigma^\alpha|\omega|^\alpha\left(1-j\beta\left(\mathrm{sgn}(\omega)\right)\Phi\right)+j\mu\omega\right\},\end{aligned} \tag{1}$$

with Φ is given by

$$\Phi = \begin{cases}\tan\frac{\pi\alpha}{2}, & \alpha\neq 1;\\ -\frac{2}{\pi}\ln|\omega|, & \alpha = 1.\end{cases} \tag{2}$$

Here, the function $E(\cdot)$ represents the expectation operation with respect to a random variable. α is called the characteristic exponent and indicates the index of stability, while β is identified as the skewness parameter. α and β together determine the shape of the models. Moreover, σ and μ are called scale and shift parameters, respectively. Specifically, if $\alpha = 2$, α-stable distribution reduces to Gaussian distribution.

Usually, it's challenging to prove whether a dataset follows a specific distribution, especially for α-stable distribution without a closed-form expression for its PDF. Therefore, when a dataset is said to satisfy a specific distribution, it usually means the dataset is consistent with this hypothetical distribution and its corresponding properties. In other words, the validation needs to firstly estimate the unknown parameters from a given dataset, and then check the fitting error between the real distribution of the dataset and the estimated one [14].

III. The Statistical Pattern of Base Stations With Large-Scale Identification

In this section, we present the fitting results to the real data. [15] shows that 10% of BSs experience roughly about 50–60% of the aggregate traffic load, which implies that the spatial traffic dynamics in cellular networks exhibit heavy-tailed pattern with densely clustering characteristic. Accordingly, in order to fulfill this nonuniform traffic demand, BSs in urban cities tend to be deployed in clusters as well. Intuitively, the BS spatial density distribution would be heavy-tailed just like the spatial traffic dynamics. Therefore, in order to characterize this realistic phenomenon, besides the traditional Poisson distribution, we choose several popular heavy-tailed candidates in Table II.

Afterwards, based on the large amount of BS location data, we sample one certain city randomly with a fixed sample area size. Then, we compute the spatial density for different 10 000 sample areas and obtain the empirical density distribution, by counting and sorting the number of BSs in each sample area. Next, we estimated the unknown parameters in candidate distributions (except α-stable distribution) using maximum likelihood estimation (MLE) methodology. For α-stable distribution, we estimate the relevant parameters using quantile methods [16], correspondingly build the model to generate the corresponding random variable, and finally compare its induced PDF with the exact (empirical) one.

In the first place, we refer to City B as an example, and compute the PDF of BS density under the sample area size 4×4 km^2. After fitting the corresponding PDF to distributions (and giving the estimated parameters) in Table II, we provide the comparison between the empirical BS density distribution with candidate ones in Figs. 2 and 3(b). As depicted in Fig. 3(b), the statistical pattern of BSs obviously exhibits heavy-tailed characteristics. Besides, among all candidate distributions, α-stable distribution most precisely match the empirical PDF. On the other hand, we provide the numerical comparison in Table III, in terms of root mean square error (RMSE). Indeed, the RMSE results in Table III show α-stable distribution has the minimum RMSE value (0.0279) while Poisson distribution has the maximum one (0.2537), and once again strengthen this aforementioned conclusion.

Meanwhile, for verifying the general accuracy of candidate distributions, we change the sample area sizes to 3×3 and 5×5 km^2 respectively, and plot the related results in Fig. 3(a) and (c). Obviously, compared with other candidate distributions, α-stable distribution still provides the most accurate fitting results for the BS density distribution in City B.

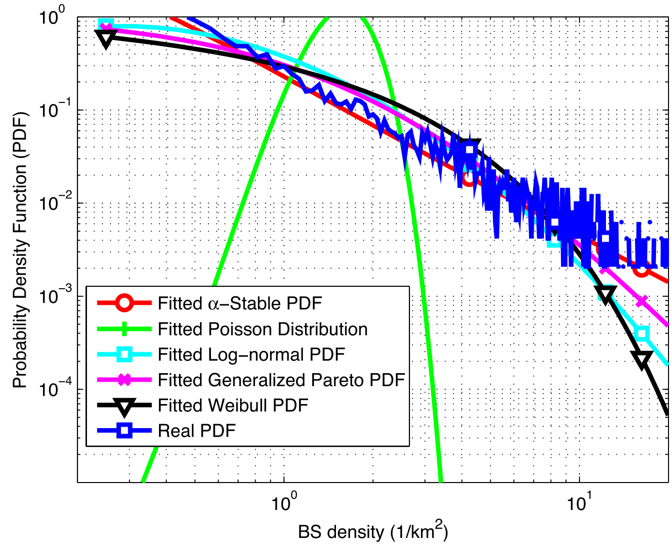

Fig. 2. The log-log comparison between practical BS density distribution in City B with candidate ones, when sample area size equals 4×4 km^2.

In order to examine the geographical impact on the fitting results, we further analyze the density distribution of BSs in City A and City C using a sample area size of 4×4 km^2. Due to the factor of geographical irregularity, there is a noticeable gap between the α-stable distribution and the empirical PDF of City A and C in comparison with City B. Nevertheless, as shown in Table III and Fig. 4, it can be observed that, α-stable distribution could match the practical one in both cities, with RMSE values equaling 0.0177 and 0.0451 respectively and being less than those of other candidate distributions. Moreover, the same conclusions concerned with sample area sizes of 3×3 and 5×5 km^2, could be also testified in Table III.

Based on the extensive analyses above, we could confidently reach the following remark.

Remark: The spatial pattern of deployed BSs exhibits strong heavy-tailed characteristics. Based on the large-scale identification, α-stable distribution manifests itself as the most precise one. On the contrary, the popular Poisson distribution is an inappropriate model for the BS density distribution, in terms of the root mean square error.

IV. Conclusions and Future Works

In this letter, based on the practical BS deployment information of one on-operating cellular networks, we carried out a thorough investigation over the statistical pattern of BS density. Our study showed that the distribution of BS density exhibits strong heavy-tailed characteristics. Furthermore, we found that the widely adopted Poisson distribution severely diverges from the realistic distribution. Instead, α-stable distribution, the distribution also found in the traffic dynamics of broadband networks and cellular networks, most precisely match the practical one. Moreover, our study could contribute to the understanding of evolution trend of BS deployment, as well as the impact of human social activities in long term.

Currently, the lack of closed-form for α-stable distribution makes it difficult to reach tractable solutions and might hinder its applications in networking performance (e.g., coverage, rate,

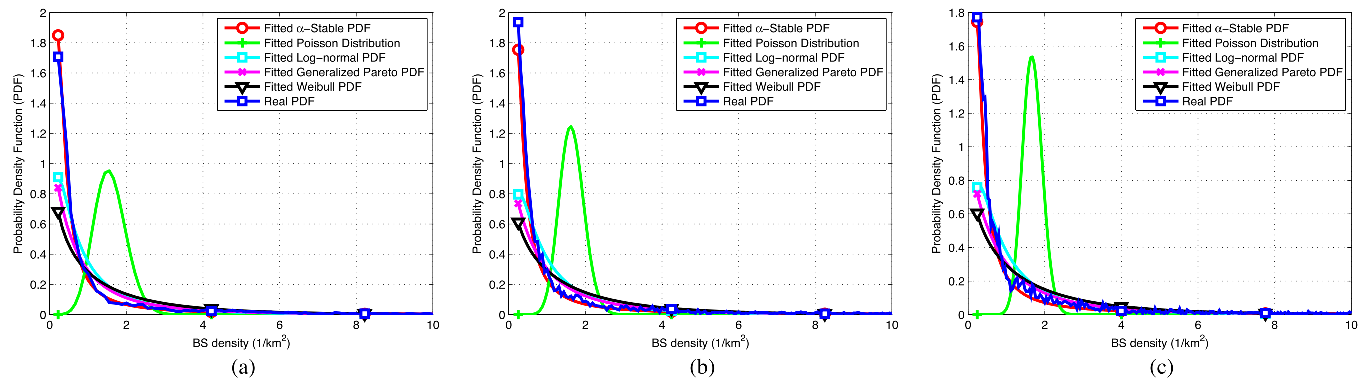

Fig. 3. The results after fitting BS density distribution in City B to candidate distributions, when sample area sizes vary. (a) 3×3 km^2; (b) 4×4 km^2; (c) 5×5 km^2.

TABLE III
RMSE VALUES AFTER FITTING CANDIDATE DISTRIBUTIONS TO EMPIRICAL ONE IN THREE CITIES

City	Sample Area Size (km^2)	α-Stable	Poisson	Log-normal	GP	Weibull
A	3×3	0.0105	0.1214	0.0207	0.0274	0.0361
	4×4	0.0177	0.1465	0.0269	0.0339	0.0418
	5×5	0.0286	0.1702	0.0293	0.0357	0.0432
B	3×3	0.0207	0.2088	0.0658	0.0770	0.0924
	4×4	0.0279	0.2537	0.0905	0.1017	0.1151
	5×5	0.0300	0.2913	0.0971	0.1085	0.1217
C	3×3	0.0373	0.2332	0.0513	0.0755	0.0910
	4×4	0.0451	0.2918	0.0697	0.0948	0.1076
	5×5	0.0487	0.3405	0.0705	0.0960	0.1064

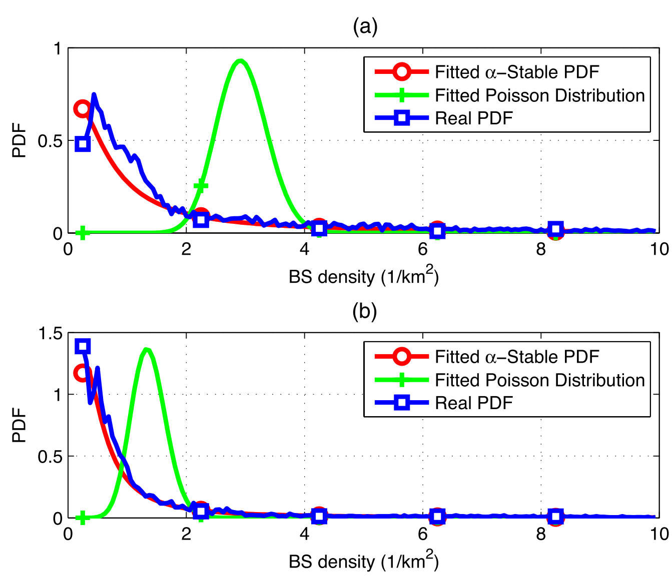

Fig. 4. The comparison between BS density distribution and α-stable distribution in City A and City C, when sample area size equals 4×4 km^2.

etc.) analyses. Therefore, we are dedicated to handle the related meaningful yet more challenging issues over applications of α-stable distribution in the future.

REFERENCES

[1] J. Andrews, F. Baccelli, and R. Ganti, "A tractable approach to coverage and rate in cellular networks," *IEEE Trans. Wireless Commun.*, vol. 59, no. 11, pp. 3122–3134, Nov. 2011.

[2] M. Haenggi, J. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 7, pp. 1029–1046, Sep. 2009.

[3] Y. Zhou *et al.*, "Two-tier spatial modeling of base stations in cellular networks," in *Proc. IEEE PIMRC*, Washington, DC, USA, Sep. 2014, pp. 1–5.

[4] N. Deng, W. Zhou, and M. Haenggi, "The ginibre point process as a model for wireless networks with repulsion," *IEEE Trans. Wireless Commun.*, vol. 14, no. 1, pp. 107–121, Jan. 2015.

[5] X. Zhou *et al.*, "Human mobility patterns in cellular networks," *IEEE Commun. Lett.*, vol. 17, no. 10, pp. 1877–1880, Oct. 2013.

[6] J. Zhang *et al.*, "Base stations from current mobile cellular networks: Measurement, spatial modeling and analysis," in *Proc. IEEE WCNC*, Shanghai, China, Apr. 2013, pp. 1–5.

[7] A.-L. Barabsi and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, Oct. 1999.

[8] J. Andrews, H. Claussen, M. Dohler, S. Rangan, and M. Reed, "Femtocells: Past, present, and future," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 3, pp. 497–508, Apr. 2012.

[9] Y. J. Chun, M. O. Hasna, and A. Ghrayeb, "Modeling heterogeneous cellular networks interference using poisson cluster processes," *IEEE J. Sel. Areas Commun.*, to be published.

[10] G. Samorodnitsky, *Stable Non-Gaussian Random Processes: Stochastic Models With Infinite Variance*. New York, NY, USA: Chapman & Hall, Jun. 1994. [Online]. Available: http://www.amazon.com/Stable-Non-Gaussian-Random-Processes-Stochastic/dp/0412051710

[11] M. Crovella and A. Bestavros, "Self-similarity in World Wide Web traffic: Evidence and possible causes," *IEEE/ACM Trans. Netw.*, vol. 5, no. 6, pp. 835–846, Dec. 1997.

[12] J. R. Gallardo, D. Makrakis, and L. Orozco-Barbosa, "Use of alpha-stable self-similar stochastic processes for modeling traffic in broadband networks," in *Proc. SPIE Conf.*, Boston, MA, USA, Nov. 1998, vol. 3530, pp. 1–16.

[13] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On power-law relationships of the Internet topology," in *Proc. ACM SIGCOMM*, 1999, pp. 251–262.

[14] X. Ge, G. Zhu, and Y. Zhu, "On the testing for alpha-stable distributions of network traffic," *Comput. Commun.*, vol. 27, no. 5, pp. 447–457, Mar. 2004.

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

[16] J. H. McCulloch, "Simple consistent estimators of stable distribution parameters," *Commun. Stat. Simul.*, vol. 15, no. 4, pp. 1109–1136, Jan. 1986.