IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 23, 2016, accepted April 7, 2016, date of publication April 12, 2016, date of current version April 29, 2016.

Digital Object Identifier 10.1109/ACCESS.2016.2552981

What Is the Best Spatial Distribution to Model Base Station Density? A Deep Dive Into Two European Mobile Networks

LUCA CHIARAVIGLIO[1], FRANCESCA CUOMO[1], (Senior Member, IEEE), MAURIZIO MAISTO[1], ANDREA GIGLI[1], JOSIP LORINCZ[2], (Member, IEEE), YIFAN ZHOU[3], ZHIFENG ZHAO[3], CHEN QI[3], AND HONGGANG ZHANG[3]

[1]Department of Information Engineering, Electronics and Telecommunications, University of Rome Sapienza, Rome 00184, Italy
[2]Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, University of Split, Split 21000, Croatia
[3]College of Information Science and Electronics Engineering, Zhejiang University, Hangzhou 310027, China

Corresponding author: L. Chiaraviglio (luca.chiaraviglio@uniroma1.it)

This work was supported by the National Basic Research Program of China (973Green) under Grant 2012CB316000.

ABSTRACT This paper studies the base station (BS) spatial distributions across different scenarios in urban, rural, and coastal zones, based on real BS deployment data sets obtained from two European countries (i.e., Italy and Croatia). Basically, this paper takes into account different representative statistical distributions to characterize the probability density function of the BS spatial density, including Poisson, generalized Pareto, Weibull, lognormal, and α-Stable. Based on a thorough comparison with real data sets, our results clearly assess that the α-Stable distribution is the most accurate one among the other candidates in urban scenarios. This finding is confirmed across different sample area sizes, operators, and cellular technologies (GSM/UMTS/LTE). On the other hand, the lognormal and Weibull distributions tend to fit better the real ones in rural and coastal scenarios. We believe that the results of this paper can be exploited to derive fruitful guidelines for BS deployment in a cellular network design, providing various network performance metrics, such as coverage probability, transmission success probability, throughput, and delay.

INDEX TERMS Base stations, spatial density, stochastic geometry, alpha-Stable distribution

I. INTRODUCTION

Cellular operators deploy the Base Stations (BSs) with the goal of providing coverage over a territory and capacity to users. As a consequence, BSs are placed following the dynamics of human sociality and people living-style, which tend to aggregate in cities. Not surprisingly, a large number of BSs is deployed in urban areas, where the BS deployment is mostly driven by the needed capacity that has to be provided to users. This trend will constantly continue in the future, due to the fact that future 5G networks (which are expressly deployed for urban environments) will require a further densification in the number of BSs in order to provide extremely high throughput to the end users [2].

In this context, studying the spatial distribution of BSs over the territory has a twofold aim. On one side, in fact, it is possible to characterize and to understand the BS deployment in current operators networks. On the other side, it is possible to provide insights into the evolution of the network as well. Specifically, the definition of theoretical models to predict the spatial distribution of BSs is gaining significant momentum [3]. In the past, one of the most popular models in cellular networks assumed that the BSs were placed at the corners of a hexagon, forming a hexagonal grid. Even though this model was useful to the industry for the implementation of frequency reuse, it failed to catch up with the actual BS deployment evolution. Additionally, current deployments are often adopting heterogeneous BSs, which can not be characterized by a regular grid. In addition to this, the Poisson Point Process (PPP) model has been frequently adopted [4], thanks to the fact that it is tractable, i.e., a closed-form can be easily obtained to express the coverage probability or the rate distribution. However, the PPP matching with a realistic BS deployment is an open issue [5]–[7]. More in depth, the BS density tends to be higher in hot spot areas (e.g., shopping centers, stadiums, airports and train stations) than the one observed in rural regions. In order to fill this gap, researchers have considered the application of heavy-tailed distributions, including: Pareto, Lognormal, Weibull and α-Stable [8]. The common feature of these statistical distributions is to have a tail not exponentially bounded. This fact may better match the actual BS deployments, which is characterized by hot spot zones.

Accordingly, several questions arise, such as: What is the best theoretical model matching the real BS spatial distribution? What is the impact of the BS implementation scenario? Does the best distribution depend on the adopted technology (like GSM, UMTS and LTE)? The answers to these questions are the goal of this paper. More in depth, we consider different case studies from two real European mobile networks, ranging from urban areas to rural and coastal ones. We then extract from them the spatial distribution of the BSs and we fit the candidate distributions to the real data. We take into account Poisson, Pareto, Lognormal, Weibull and α-Stable distributions as possible candidates. Our results, obtained across GSM/UMTS/LTE technologies, show that the Lognormal and Weibull distributions are the best candidates for rural/coastal zones, while the α-Stable fully matches the real spatial distribution in the urban scenarios. These facts stimulate further research towards to definition of tractable models tailored to specific deployment zones (i.e., urban and rural ones).

The closest work to this paper is [3], in which authors have evaluated how much the α-Stable distribution matches the real one in three Chinese cities. In contrast to them, in this work we move five significant steps further by:

- investigating what is the best distribution across different BS deployment scenarios, and not only urban ones like in [3]. Our goal is in fact to find, behind the evidences of α-Stable distribution, also its boundaries, i.e., the limits of applicability of this model.
- taking into account a big region of Italy, and extracting from it two representative scenarios. Specifically, different types of zones are considered (i.e., urban and rural ones).
- considering the country of Croatia, and extracting from it urban, rural and coastal scenarios.
- evaluating the impact of the different cellular networking technologies (i.e., GSM, UMTS, LTE) while considering all or a portion of the operators providing a service over the territory (for the Italian case).
- evaluating the impact of adding new BSs that are planned to be deployed in the future (for the Croatian case). In this way, we are able to provide fruitful indications about the evolution of BS distribution in future networks.

We believe that the results of this paper may be exploited for future works, which can provide network performance metrics like coverage and transmission success probabilities, throughput, delay, and so on. Particularly, these metrics are helpful to analyze various cellular networking scenarios including BS cooperation, uplink/downlink decoupling, BS on-off switching, small cells densification, interference cancellation, relaying, distributed massive Multiple Input Multiple Output (MIMO), full-duplex radio, and even caching on the edge nodes with backhauling.

The rest of the paper is organized as follows. Sec. II reviews the related works. The datasets description as well as a brief overview of the adopted distributions are reported in Sec. III. Results are presented in Sec. IV. Additionally, Sec. V discusses our work. Finally, conclusions and future works are drawn in Sec. VI.

II. RELATED WORKS

The spatial distribution of BSs has been always a fundamental issue in wireless communications, due to the need of frequency reuse and interference characterization. Step over from hexagonal model, in recent years, PPP has been widely adopted as valuable choice to model the BS locations in cellular networks [4], [9]–[11]. As a baseline role, PPP model can provide tractable results for performance evaluation in both one-tier and multi-tier networking scenarios [12]. However, it may be not accurate enough to model BS locations. Specifically, researchers still hardly reach a consensus on PPP's performance to model the real deployment. For example, in [4] and [13] the authors observe inconsistent coverage probability performance of the PPP model when modeling the real BS locations data from different cities of the world. Given the conflicting results above, it is still worthwhile to conduct a deeper investigation.

On the other hand, the actual deployment of BSs in long term is highly correlated with human sociality and traffic demand distribution [14], [15]. Human beings tend to live together, and they aggregate near hot spots. As a consequence, their social behaviors would lead to traffic peaks, thus causing BSs to be more densely deployed in certain areas as clusters. Furthermore, according to an assumption named ''preferential attachment'', Barabasi and Albert in [16] argue that many large networks grow to be heavy-tailed. Therefore, heavy-tailed distributions appear to be more suitable to precisely characterize the clustered nature of BSs.

As mentioned before, the authors in [3] have investigated several heavy-tailed distributions to examine the spatial density of BSs in cellular networks. Besides the inaccuracy verification of Poisson distribution, they claim that α-Stable distribution can precisely characterize the spatial density of BSs in three cities of China. Although the relevant modeling results present very sound performance in different kinds of cities, there is definitely the need for more real data and in-depth investigation to verify its global applicability and conformity across the world. Moreover, it is of mandatory importance to consider the impact of different operators and of different cellular technologies.

III. DATASET AND DISTRIBUTION DESCRIPTIONS

A. ITALIAN DATASET DESCRIPTION

We initially focus on the Emilia-Romagna region of Italy, which is covered by four different cellular operators (referred as A, B, C and D in the following). Table 1 reports the

TABLE 1. Main features of the italian data set.

Operator ID	Number of BSs	Subscribers [millions]	Avg. BS Density $[1/km^2]$
A	1570	2.20	0.07
B	1388	2.36	0.07
C	1179	1.58	0.05
D	833	0.70	0.04

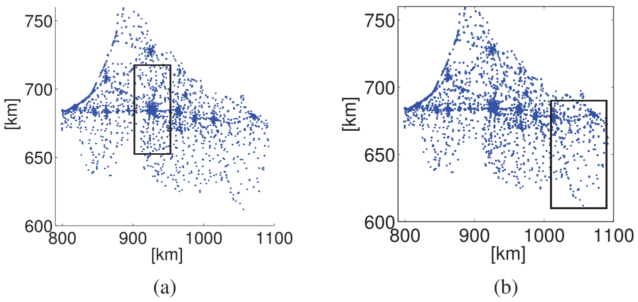


FIGURE 1. Italian Dataset: BS positions and considered scenarios (inside the rectangular boxes). (a) IT urban scenario. (b) IT rural scenario.

main features of the considered dataset. The total number of deployed BSs considering the whole set of operators is more than 4900 BSs. Focusing then on each operator, the number of deployed BSs is similar for operator *A* and operator *B*, while it is slightly lower for operators *C* and *D*. More in depth, operator *D* reuses part of the cellular infrastructure of the two largest operators to guarantee coverage in the zones not covered by its own BSs. As for the morphological characteristics of the area, the whole region spans over 22000 km^2, which includes rural areas, town areas and one metropolitan area. This is also reflected in the number of subscribers, which is larger than 6.5 millions in total, with the largest number of subscribers living in the metropolitan area. Finally, the average BS density (i.e., the total number of deployed BSs for each operator over the total region), is always lower than one, due to the fact that in rural areas less BSs are deployed compared to urban ones. However, the density is larger for operator *A* and *B*, and slightly lower for the other operators.

In this work, we are interested in the BS spatial distribution over different areas. In particular, we have first selected the metropolitan area in the region, as illustrated in Fig. 1(a). This area spans over 3300 km^2. In this subregion, more than 1000 BSs are deployed, which correspond to more than 20% of the total number of BSs in the whole scenario. The resulting average BS density is equal to 0.3 km^{-2}. Moreover, the total number of subscribers is more than 3.0 millions, thus representing more than 45% of the whole number of subscribers. Additionally, we have also selected a rural area, mainly composed of countryside and small towns, as illustrated in Fig. 1(b). For each zone the GPS coordinates of the BSs are available. In order to obtain the spatial positions of BSs we have applied the Universal Transverse Mercator (UTM) conformal projection [17]. In this way, the GPS coordinates are mapped to a plane.

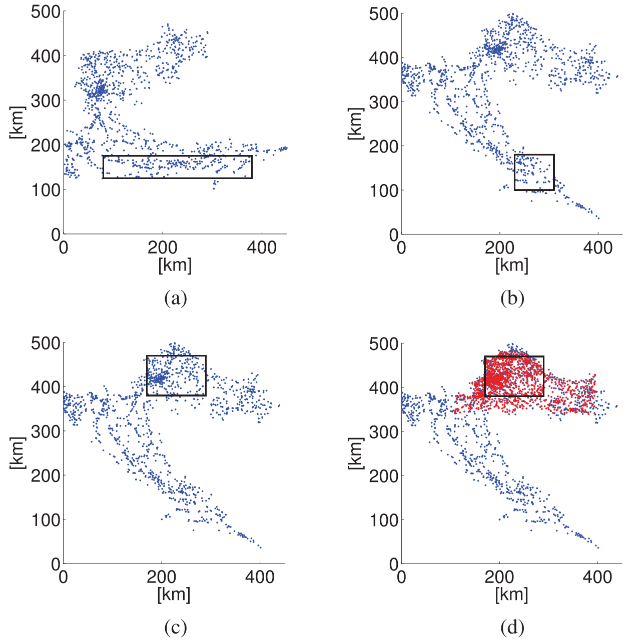


FIGURE 2. Croatian Dataset: BS positions and considered scenarios (inside the rectangular boxes). (a) CRO coastal scenario. (b) CRO rural scenario. (c) CRO urban scenario. (d) CRO urban scenario including future planned BSs.

B. CROATIAN DATASET DESCRIPTION

In addition to the Italian dataset, we have considered the set of BSs sites having freestanding masts from the country of Croatia. In particular, more than 2600 BSs are deployed in an area of around 56000 km^2. The database is composed of the BSs sites owned by the telecom operators currently active in Croatia, serving in total more than 4.6 millions of users. The morphological characteristics of the country include one large metropolitan area around the capital Zagreb, different rural zones, and one coastal zone including most of tourist attractions. In addition to the BSs sites actually deployed in the network, the positions of planned BSs sites to be installed in the future is also provided, considering a vast region in the north of the country.

In our analysis, we have selected four representative scenarios, namely: i) one coastal region, ii) one rural zone, iii) the metropolitan region of Zagreb, and iv) the same metropolitan region including also future planned BSs. Fig. 2 shows the selected areas (i.e., the ones inside the rectangles). Moreover, Table 2 reports the characteristics of each scenario,

TABLE 2. Considered scenarios.

Scenario	Number of BSs	Area $[km^2]$	Avg. BS Density $[1/km^2]$
IT Urban	1002	51×65	0.3
IT Rural	538	80×80	0.08
CRO Urban	337	120×90	0.03
CRO Urban (with planned BSs)	1037	120×90	0.09
CRO Rural	101	80×80	0.015
CRO Coastal	234	300×50	0.015

in terms of: number of considered BSs, size of the area, and average BS density. Interestingly, the BS density in the Croatian scenarios is in general lower than that in the Italian ones, e.g. the Italian urban scenario has a density of 0.3 $1/km^2$, while the Croatian urban equals 0.09 $1/km^2$. This is mainly due to the two reasons: i) the LTE network is not so widespread in Croatia like it is in the Italian region, resulting in a lower number of deployed BSs, ii) the morphological characteristics are different in the two areas. In particular, the considered region of Italy is composed of a large plain and an area with hills. The Croatian country instead includes one large plain in the north of the country, different mountains in the middle, and a coastal zone with a large number of islands in the south west. For all these reasons, we believe that the Croatian country is also representative and useful to be analyzed in addition to the Italian dataset.

TABLE 3. List of candidate distributions.

Distribution	Probability Density Function
Generalized Pareto (GP)	ax^{-b}
Weibull	$abx^{b-1}e^{-ax^b}$
Lognormal	$\frac{1}{\sqrt{2\pi}bx}e^{-\frac{(\ln x-a)^2}{2b^2}}$
α-Stable	Closed form not always exists. Characteristic Function presented in Eq. (1).
Poisson	$\frac{\lambda^k}{k!}e^{-\lambda}$

C. CANDIDATE DISTRIBUTIONS

We now focus on the theoretical statistical distributions that can model the spatial distributions of BSs over the territory. Specifically, many phenomena in telecommunication industry can be characterized by heavy-tailed distributions, including the teletraffic statistics [18] and Internet topology [19]. Meanwhile, there exist many statistical distributions proving to be heavy-tailed. Among them, Generalized Pareto (GP) distribution, Weibull distribution, and Lognormal distribution belong to the ones with Probability Density Function (PDF) in closed-forms (see Table 3). Another famous heavy-tailed distribution is the α-Stable one, which manifests itself in the capability to characterize the distribution of normalized sums of a relatively large number of independent identically distributed random variables [8]. However, the α-Stable distribution, with a few exceptions, lacks a closed-form expression of the PDF, and it 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:

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

with Φ is given by

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

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.

IV. CASE-STUDIES RESULTS

Given the BS positions in each scenario, we then compute the empirical spatial distribution of the BS density. Initially, we sample each scenario with a small area of fixed size. We then randomly select 10000 squares of fixed area size. For each square, we compute the number of BSs falling into it. This number, divided by the area size, represents the BS density. From the BS densities, we derive the PDF. This spatial distribution is then used as reference one vs. the possible candidates (i.e., Poisson, GP, Weibull, Lognormal and α-Stable). For each candidate distribution, we estimate the unknown parameters by applying the Maximum Likelihood Estimation (MLE) criterion. For estimating the parameters of the α-Stable distribution, we adopted a similar procedure like the one reported in [3], due to the fact that the closed-form for this distribution does not always exist.

We initially focus on the urban area of the Italian scenario. As a showcase, we compute the PDF of BS density with a sample area of size 10×10 km^2. Moreover, we have taken into account the BSs from all the operators in order to maximize the number of BSs under consideration. Fig. 3 reports the empirical PDF (i.e., the real one) with the fitting of various candidate distributions. Interestingly, the best fitting is obtained with the α-Stable distribution, while the other ones perform consistently worse. To give more insight, we have provided the log-log plot of the distributions in Fig. 4.

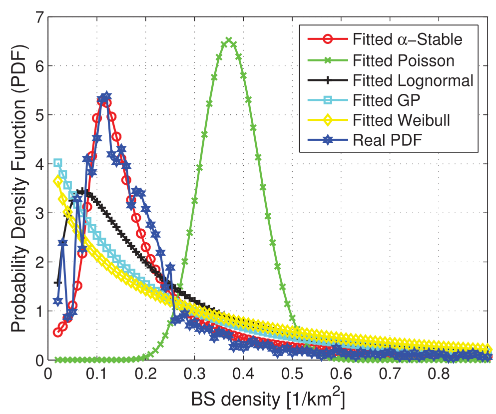


FIGURE 3. Italian urban scenario: probability density function of the BS density with all operators and sample squared area with size 10 km^2.

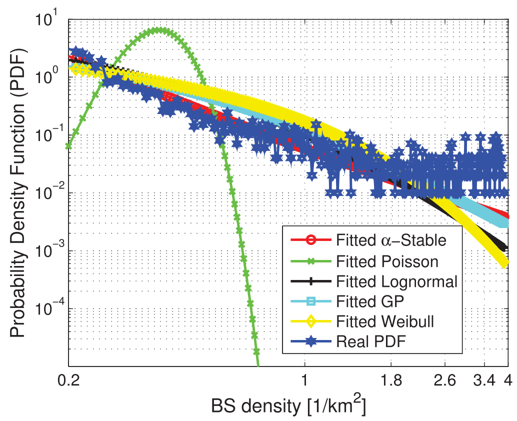


FIGURE 4. Italian urban scenario: Log-Log plot of the probability density function of the BS density with all operators and sample squared area with size 10 km^2.

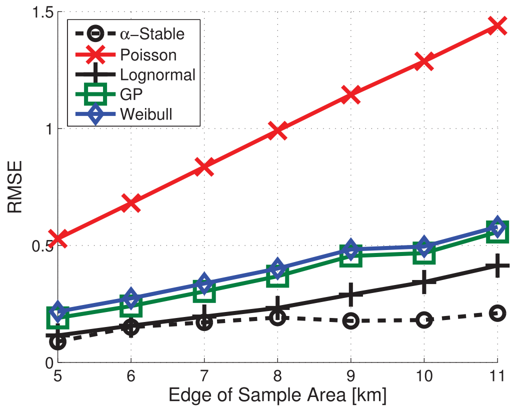


FIGURE 5. Italian urban scenario: RMSE vs. size of the sample squared area.

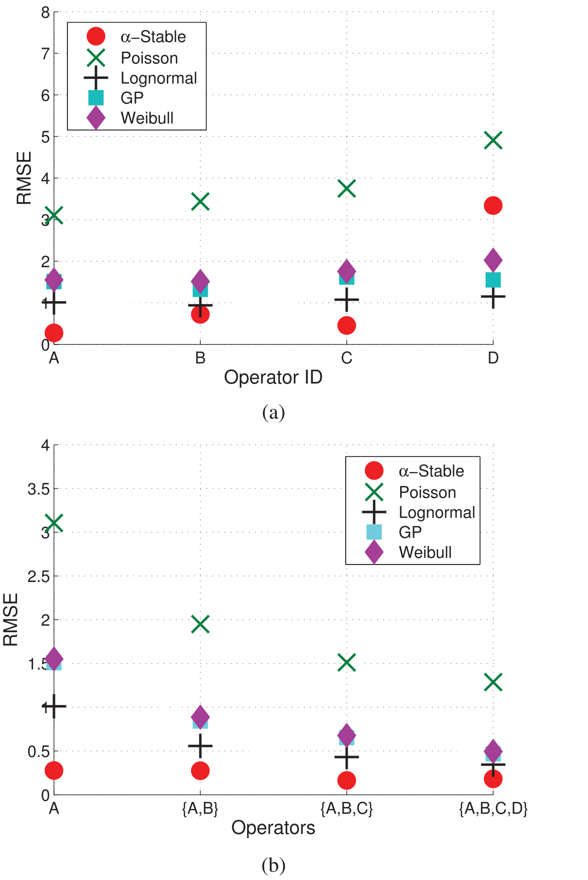


FIGURE 6. Italian urban scenario: RMSE for single and multiple operators. (a) Single operators. (b) Multiple operators.

We can clearly see from the figure that the α-Stable distribution matches the real PDF. For very high densities (i.e., higher than 2.6 km^{-2}) the real PDF tends to be constant as the density increases, while all the fitting curves tend to decrease. This is due to the fact that the number of samples with this characteristic (i.e., a very high density) is too inadequate to be captured, therefore the fitting tends to be less accurate compared to the zones with lower densities (i.e., the left-hand side of the figure).

In the following, we have computed the Root Mean Square Error (RMSE) of the different fittings against the empirical PDF. This metric is useful to capture the fitting accuracy of the considered distribution. In this case, for modeling generalization purpose, we have also considered the variation of the sample area between 5×5 km^2 and 11×11 km^2 in the scenario. Recall that for each sample area size we randomly select 10000 samples in the scenario. Fig. 5 illustrates the obtained results. Obviously, the α-Stable is the best fitting for all the considered sample areas, with a RMSE always lower than 0.3. On the other hand, the Poisson distribution exhibits a RMSE always larger than 0.5, thus confirming our intuition that it is not suitable to capture the spatial density distribution of real BSs. Moreover, the RMSE of the other distributions (i.e., Lognormal, generalized Pareto, Weibull) is always between the one of the Poisson distribution and the α-Stable fitting. From this figure, we can observe that the RMSE tends to increase when the sample area size is increased: this is due to the fact that the number of samples tends to be decreased, and each sample tends to have a similar density compared to the others.

Furthermore, we have investigated the impact of single operators. Unless otherwise specified, we set a sample area size equal to 10×10 km^2. Fig. 6(a) reports the RMSE values for each single operator. Recall that A and B exhibit the largest number of BSs, while operator D tends to exploit the BSs of the other operators to provide user coverage. Surely, the α-Stable is the best fitting for operators A, B and C. On the contrary, for operator D the α-Stable RMSE is lower than the Poisson distribution but higher than the other ones. This is due to the fact that this operator does not spread its own BSs in the same way like the other ones, resulting in a different density distribution. Moreover, we can see that the RMSE tends to increase from left to right (i.e., towards operators with less BSs).

To give more insight, Fig. 6(b) provides the results when multiple operators are considered to compute the BS density. The set $\{A, B, C, D\}$ is equivalent to the case in which all

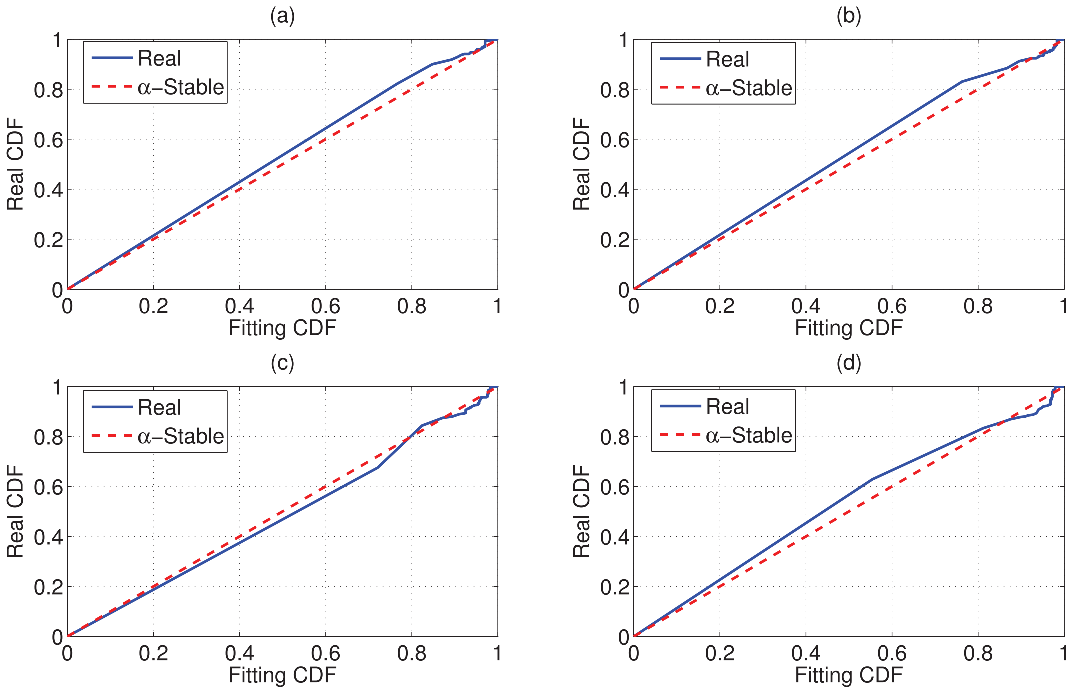


FIGURE 7. Italian urban scenario: p-p plot for different sample area sizes (all operators).

operators are considered. Interestingly, the α-Stable fitting tends to be almost constant, while the RMSEs of the other distributions tend to increase when the number of operators is decreased. In particular, operator A, which is also the largest one, has deployed its BSs in the scenario in order to always provide coverage to users with its own BSs. On the contrary, operator D tends to lease the infrastructure from the other operators. The case with the single operator A matches better a complete BSs deployment in the targeted region, resulting in a low RMSE with the α-Stable model.

In the following part we have taken into account the impact of various cellular networking technologies on the BS density. Together with each BS position, in fact, our dataset includes information about the technology, which can be GSM, UMTS, LTE, or not specified. Each BS entry in the BS database includes a list of the supported technologies. Specifically, by manually checking in the BS database, we have found that the UMTS service is always provided in the considered region, except for the BSs for which the technology is not specified. At the same time, when the LTE service is provided, also GSM and UMTS services are available. Therefore, we have considered the following categories: GSM/UMTS, GSM/UMTS/LTE, or the entire dataset (i.e., including the BSs for which the technology is not specified). For each category, we have then computed the empirical PDF as well as the distribution fitting. Table 4 describes the obtained RMSE values. Once again, these results confirm that the α-Stable fitting reaches the highest accuracy in this scenario, while all the other distributions have a RMSE at least more than doubled.

TABLE 4. Italian urban scenario: RMSE values vs. BS technology.

Technology	α-Stable	Poisson	Lognorm	GP	Weibull
GSM UMTS	0.2077	1.7125	0.4171	0.6898	0.7232
GSM UMTS LTE	0.2124	1.5746	0.4452	0.6122	0.6457
ALL	0.1811	1.2872	0.3437	0.4671	0.4957

In order to better evaluate how much the fitted α-Stable distribution is close to the real one, we have computed the probability-probability (P-P) plot of the fitted α-Stable distribution against the real one. Fig. 7 reports the obtained results for different sample area sizes. The range of each plot is a unit square $[0, 1] \times [0, 1]$, due to the fact that each Cumulative Distribution Function (CDF) is defined between 0 and 1. The figures also illustrate a diagonal line of the square, representing the P-P plot of the fitted α-Stable distribution with itself. Surely, we can notice that the P-P plot of the α-Stable distribution is very close to the real one, for all the values of the different sample area sizes. Thus, we can conclude that, based on the P-P plot expression for this scenario, the α-Stable is the best fitting model matching the real distribution of BSs.

In the following, we have moved our attention to the Italian rural scenario. Differently, from the previous case, in this scenario there are no big towns, and the BS distribution over the territory is rather sparse. In order to evaluate the behavior of the different distributions, we have computed the RMSE

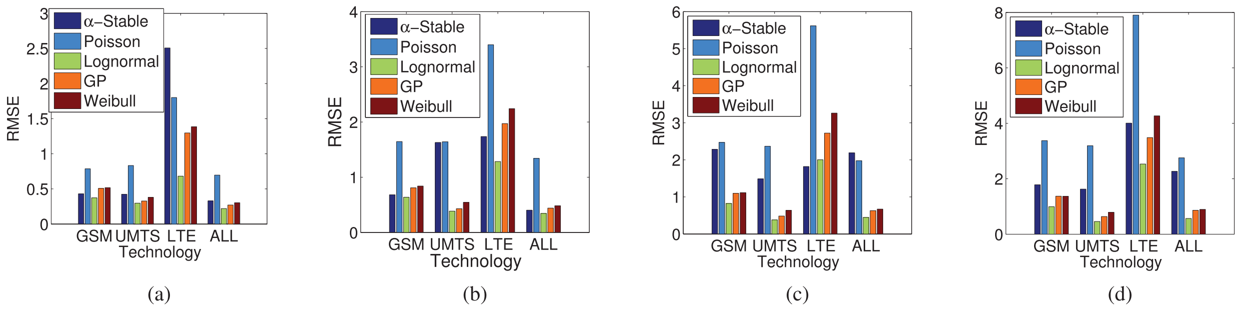


FIGURE 8. Italian rural scenario results. (a) Sample area size 5×5 km^2. (b) Sample area size 7×7 km^2. (c) Sample area size 9×9 km^2. (d) Sample area size 11×11 km^2.

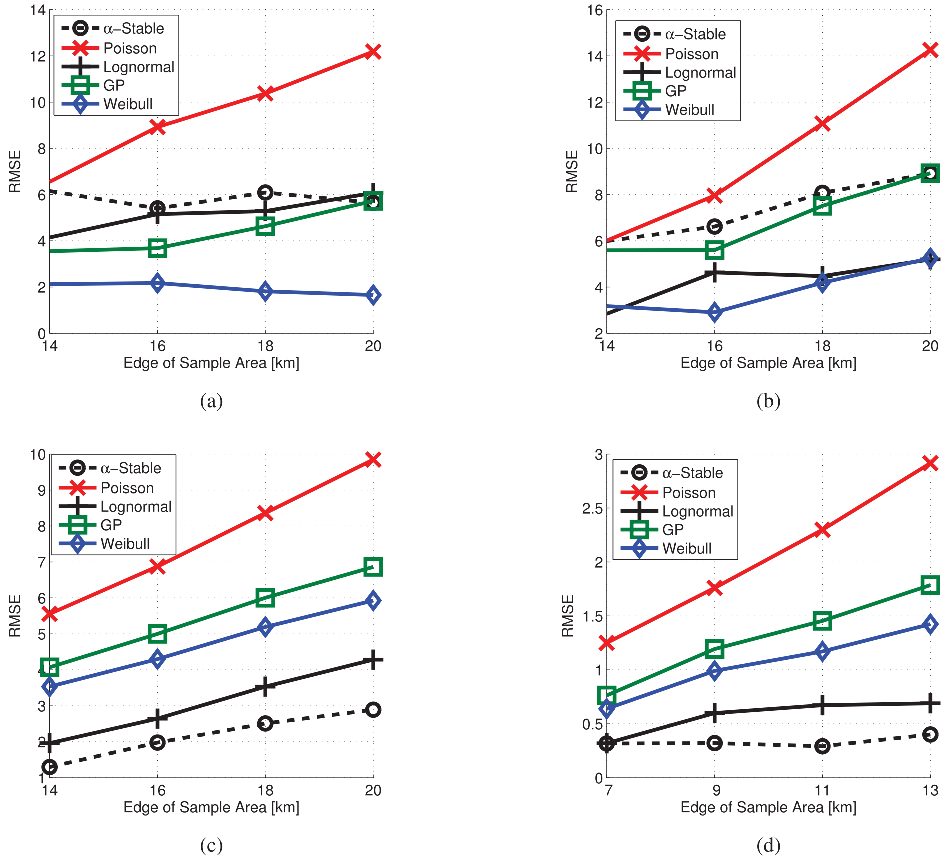


FIGURE 9. Croatian scenarios: RMSE of the distributions for different sample area sizes. (a) Coastal. (b) Rural. (c) Urban. (d) Urban with Future Planned BSs.

for different sample area sizes, and for different technologies, as reported in Fig. 8. As expected, the Poisson distribution does not adhere to the empirical distribution, resulting in the highest RMSE. The other distributions tend to have a lower RMSE. Among them, the best candidate is the Lognormal distribution in most of the cases. On the contrary, the α-Stable distribution exhibits a higher RMSE than the Lognormal one (but lower than the Poisson one). This fact is confirmed across the different technologies, and for the different area sizes. Therefore, the α-Stable distribution is useful to capture the BS spatial behavior in urban maps. When the BS distribution becomes more sparse than in a city, like in this case, the best candidates are other types of distributions (i.e., like the Lognormal one in this case).

We have investigated in the next step the Croatian scenarios. Fig. 9 reports the obtained results in terms of RMSE for the different distributions. In this case, we have also varied the size of the sample area. Particularly, since the BSs are rather sparse in the rural, coastal and urban scenarios (without planned BSs), we have adopted a larger sample area size than the Italian cases (i.e., ranging between 14×14 km^2 and 20×20 km^2). On the contrary, we have adopted a sample area comparable with the Italian case for the urban scenario with future planned BS, since the BS density is quite similar in these two cases (see also Table 1). Focusing on the obtained results, the best fitting for the coastal case is the Weibull distribution (reported in Fig. 9(a)), while the Lognormal one tends to achieve comparable RMSE values when the rural case in considered (see Fig. 9(b)). However, when the urban scenarios are considered (Fig. 9(c) and Fig. 9(d)) the distribution achieving the lowest RMSE is the α-Stable. This fact further confirms our intuition that the α-Stable model matches the BS spatial distribution in urban areas. Moreover, our results also imply that the definition of a universal model, covering all kinds of urban, coastal, and rural scenarios, is still an open issue.

V. DISCUSSION

Our results have indicated that the α-Stable model is able to characterize the BS distribution in urban scenarios, while the Lognormal and Weibull distributions are the best candidates in rural and coastal zones. In addition to these facts, we have found that the Poisson distribution tends to be less accurate. However, the Poisson distribution has the main advantage of being tractable, e.g., it is possible to compute different system performance metrics in a closed-form. Moreover, recent works by [20] and then later by [6] show that random shadowing and fading models render networks to appear Poisson to the user, even if the BSs do not form a PPP. In contrast to these works, here we are more focused on highlighting the spatial distribution of BSs. Our results show that the pure PPP is not matched in the considered BSs deployment, thus paving a way for future models which should be sufficiently accurate and also tractable.

Focusing on other point process models, Determinantal Point Processes (DPPs) [21], [22] have been fitted to cities by [23]. The main advantage of DPPs is the possibility to model the repulsiveness among macro BSs, i.e., the fact that BSs are not installed very close to each other. Moreover, DPPs are more tractable than pure PPPs. We plan to compare DPPs and the α-Stable model as future work.

Finally, in [24] BSs are first divided into different subsets according to geographical factor (e.g., urban or rural) and functional type (e.g., macrocells or microcells). Then, as a second step, a detailed spatial analysis is performed on each subset. Results show that the PPP is rather inaccurate for modeling the BS locations. Moreover, Gibbs Point Processes [25] as well as Neyman-Scott Point Processes [26] are taken into account and compared in view of large-scale modeling test. The obtained results highlight the clustering nature of BSs deployment. Moreover, the need of characterizing the different BSs tiers (e.g., macro BSs and micro BSs) clearly emerges. Therefore, in the future, we will also investigate the impact of applying the α-Stable model to different BSs tiers.

VI. CONCLUSIONS AND FUTURE WORK

We have studied the BS spatial distributions across different scenarios obtained from Italy and Croatia, considering urban, coastal, and rural zones. We have compared the real distribution against different candidate ones. Our results show that the best distribution matching the real one is the α-Stable model in urban scenarios. This fact is confirmed across different sample area sizes, operators, and technologies. On the contrary, the Lognormal and Weibull distributions tend to fit better the real one in coastal and rural scenarios. We believe that this work can be used to derive fruitful guidelines for the BS deployment. As next step, we will complement these findings with a detailed study of spatial and temporal variations of user traffic. Moreover, we will extend our analysis to other countries (such as in Asia and in North America). Finally, we will study the possibility of deriving a universal model covering rural, urban and coastal zones.

ACKNOWLEDGEMENT

This paper was presented at the A reality check of base station spatial distribution in mobile networks, in IEEE Infocom 2016 (Poster), San Francisco, USA, April. 2016.

REFERENCES

[1] L. Chiaraviglio *et al.*, "A reality check of base station spatial distribution in mobile networks," in *Proc. IEEE INFOCOM*, San Francisco, CA, USA, Apr. 2016, pp. 1–2.

[2] N. Bhushan *et al.*, "Network densification: The dominant theme for wireless evolution into 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 82–89, Feb. 2014.

[3] 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.

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

[5] D. B. Taylor, H. S. Dhillon, T. D. Novlan, and J. G. Andrews, "Pairwise interaction processes for modeling cellular network topology," in *Proc. IEEE GLOBECOM*, Anaheim, CA, USA, Dec. 2012, pp. 4524–4529.

[6] B. Błaszczyszyn, M. K. Karray, and H. P. Keeler, "Wireless networks appear Poissonian due to strong shadowing," *IEEE Trans. Wireless Commun.*, vol. 14, no. 8, pp. 4379–4390, Aug. 2015.

[7] Y. Zhou, Z. Zhao, Q. Ying, R. Li, X. Zhou, and H. Zhang, "Two-tier spatial modeling of base stations in cellular networks," in *Proc. IEEE PIMRC*, Washington, DC, USA, Sep. 2014, pp. 1570–1574.

[8] G. Samoradnitsky and M. S. Taqqu, *Stable Non-Gaussian Random Processes: Stochastic Models With Infinite Variance*. New York, NY, USA: Chapman & Hall, Jun. 1994.

[9] M. Haenggi, J. G. 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.

[10] J. G. Andrews, R. K. Ganti, M. Haenggi, N. Jindal, and S. Weber, "A primer on spatial modeling and analysis in wireless networks," *IEEE Commun. Mag.*, vol. 48, no. 11, pp. 156–163, Nov. 2010.

[11] H. S. Dhillon, R. K. Ganti, F. Baccelli, and J. G. Andrews, "Modeling and analysis of K-tier downlink heterogeneous cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 3, pp. 550–560, Apr. 2012.

[12] W. C. Cheung, T. Q. S. Quek, and M. Kountouris, "Throughput optimization, spectrum allocation, and access control in two-tier femtocell networks," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 3, pp. 561–574, Apr. 2012.

[13] C.-H. Lee, C.-Y. Shih, and Y.-S. Chen, "Stochastic geometry based models for modeling cellular networks in urban areas," *Wireless Netw.*, vol. 19, no. 6, pp. 1063–1072, Aug. 2013.

[14] X. Zhou, Z. Zhao, R. Li, Y. Zhou, J. Palicot, and H. Zhang, "Human mobility patterns in cellular networks," *IEEE Commun. Lett.*, vol. 17, no. 10, pp. 1877–1880, Oct. 2013.

[15] 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.

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

[17] E. Grafarend, "The optimal universal transverse mercator projection," in *Geodetic Theory Today*. Berlin, Germany: Springer, 1995, p. 51.
[18] S. I. Resnick, "Heavy tail modeling and teletraffic data: Special invited paper," *Ann. Statist.*, vol. 25, no. 5, pp. 1805–1869, 1997.
[19] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On power-law relationships of the Internet topology," in *Proc. ACM SIGCOMM*, Cambridge, MA, USA, 1999, pp. 251–262.
[20] H. P. Keeler, N. Ross, and A. Xia. (2014). "When do wireless network signals appear poisson?" [Online]. Available: http://arxiv.org/abs/1411.3757
[21] 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.
[22] H. ElSawy, E. Hossain, and M. Haenggi, "Stochastic geometry for modeling, analysis, and design of multi-tier and cognitive cellular wireless networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 996–1019, Jul. 2013.
[23] Y. Li, F. Baccelli, H. S. Dhillon, and J. G. Andrews, "Fitting determinantal point processes to macro base station deployments," in *Proc. IEEE GLOBECOM*, Austin, TX, USA, Dec. 2014, pp. 3641–3646.
[24] Y. Zhou *et al.*, "Large-scale spatial distribution identification of base stations in cellular networks," *IEEE Access*, vol. 3, pp. 2987–2999, Dec. 2015.
[25] B. D. Ripley and F. P. Kelly, "Markov point processes," *J. London Math. Soc.*, vol. 2, no. 1, pp. 188–192, 1977.
[26] D. R. Cox and V. Isham, *Point Processes*, vol. 12. Boca Raton, FL, USA: CRC Press, 1980.

LUCA CHIARAVIGLIO received the Ph.D. degree from the Politecnico di Torino (Italy). He is an Assistant Professor with the Department of Information, Electronics and Telecommunications, Sapienza University of Rome, Italy, an Adjunct Assistant Professor with the University of Cassino and Lazio Meridionale, Italy, and a Research Associate with CNIT, Italy. He has been a post-doctoral with the Politecnico di Torino, and INRIA Sophia Antipolis (France), and a Research Fellow with CNIT. Moreover, he was a Visiting Scholar with Boston University (USA) from 2008 to 2009, and a Visiting Assistant Professor with the Auckland University of Technology (New Zealand) in 2015. He has coauthored over 75 papers in top international journals and conferences, and he participates on the Program Committees of several conferences, including the IEEE GLOBECOM, the IEEE ICC, and the IEEE Greencomm. Moreover, he serves as an Associate Technical Editor of the *IEEE Communications Magazine*. His main research interests are in the fields of sustainable networking, 5G networks, network measurement, and mobile networks deployment.

FRANCESCA CUOMO (SM'12) received the Laurea (*magna cum laude*) degree in electrical and electronic engineering and the Ph.D. degree in information and communications engineering from the University of Rome Sapienza, in 1993 and 1998, respectively. Since 2005, she has been an Associate Professor with the University of Rome Sapienza and teaches courses in Telecommunication Networks. She has advised numerous master's students in computer engineering, and has been the Advisor of nine Ph.D. students in networking. Her current research interests focus on vehicular networks and sensor networks, cognitive radio networks, 4g and 5g systems, and energy saving in the Internet and in the wireless systems. She has participated in several National and European projects on wireless network systems. She has been on the Editorial Board of the *Elsevier Computer Networks* journal and now is a member of the Editorial Board of the *Ad-Hoc Networks* (Elsevier) journal. She has served on several technical program committees. She has authored over 100 peer-reviewed papers in prominent international journals and conferences.

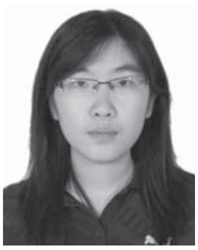
MAURIZIO MAISTO received the bachelor's degree in electronics engineering from the Sapienza University of Rome, Italy, in 2014, where he is currently pursuing the master's degree in artificial intelligence and robotics.

ANDREA GIGLI received the bachelor's degree in engineering in computer science from the Sapienza University of Rome, Italy, in 2014, where he is currently pursuing the M.S. degree in artificial intelligence and robotics.

JOSIP LORINCZ (M'07) received the M.S. and Ph.D. degrees in electrical engineering and computer science from the University of Split, FESB, Croatia, in 2002 and 2010, respectively. He was a Visiting Researcher with the Advanced Network Technologies Laboratory, Politecnico di Milano, Italy. He is the Co-Chair of the Symposium on Green Networking and Computing, organized in the frame of SoftCOM conference. He also serves as the Technical Program Committee Member for many international scientific conferences and reviewer for scientific journals and project proposals. He was awarded as Outstanding Young Researcher by the Croatian Academy of Engineering in 2013. His current research interests include energy-efficient wireless and wired networks, optimization in telecommunications, and advanced design of heterogeneous communication networks. He is a member of ACM, and the first President of the Croatian ACM Chapter. Since 2004, he has owned Cisco CCNA, CCAI, and BCMSN certificates.

YIFAN ZHOU received the B.S. degree in communication engineering from Zhejiang University, China, in 2012, where he is currently pursuing the Ph.D. degree with the College of Information Science and Electrical Engineering. His research interests include spatial modeling of cellular networks and SDN utilization in wireless networks.

ZHIFENG ZHAO is currently a Full Associate Professor with the College of Information Science and Electronic Engineering, Zhejiang University, China. He received the bachelor's degree in computer science and the master's and Ph.D. degrees in communication and information systems from the PLA University of Science and Technology, Nanjing, China, in 1996, 1999, and 2002, respectively. From 2002 to 2004, he acted as a Post-Doctoral Researcher with Zhejiang University, China, where his research was focused on multimedia next-generation networks and soft-switch technology for energy efficiency. From 2005 to 2006, he acted as a Senior Researcher with the PLA University of Science and Technology, where he performed research and development on advanced energy-efficient wireless router, ad-hoc network simulator, and cognitive mesh networking test-bed. His current research area includes software defined network, wireless multi-hop networks (ad hoc, mesh, and WSN), wireless multimedia network, cognitive networks, and green communications. He was the Symposium Co-Chair of ChinaCom 2009 and 2010. He was the Technical Program Committee Co-Chair of the Tenth IEEE International Symposium on Communication and Information Technology 2010.



CHEN QI received the B.S. degree in communication engineering from Zhejiang University, Hangzhou, China, in 2014, where she is currently pursuing the Ph.D. degree with the College of Information Science and Electrical Engineering. Her research interests include data mining of wireless networks, and green communications.

HONGGANG ZHANG is a Full Professor with the College of Information Science and Electronic Engineering, Zhejiang University, China. He was the International Chair Professor of Excellence for Université Européenne de Bretagne and Supélec, France (2012-2014). He is an Honorary Visiting Professor with the University of York, U.K. He is currently active in the research on cognitive green communications and was the leading Guest Editor of the *IEEE Communications Magazine* special issues on Green Communications. He is taking the role of the Associate Editor-in-Chief *China Communications* and the Series Editor of the *IEEE Communications Magazine* for its Green Communications and Computing Networks Series. He served as the Chair of the Technical Committee on Cognitive Networks of the IEEE Communications Society from 2011 to 2012. He was the co-author and an Editor of two books entitled *Cognitive Communications-Distributed Artificial Intelligence, Regulatory Policy and Economics, Implementation* (John Wiley & Sons) and *Green Communications: Theoretical Fundamentals, Algorithms and Applications* (CRC Press), respectively.