A Reality Check of Base Station Spatial Distribution in Mobile Networks

Luca Chiaraviglio,[1] Francesca Cuomo,[1] Andrea Gigli,[1] Maurizio Maisto,[1]
Yifan Zhou,[2] Zhifeng Zhao,[2] Honggang Zhang[2]
1) DIET Department, University of Rome, Sapienza, Rome, Italy, (e-mail: luca.chiaraviglio@uniroma1.it)
2) College of Information Science and Electronics Engineering, Zhejiang University, Hangzhou, China

***Abstract*—We consider the spatial distribution of Base Stations (BSs) in different scenarios of a European country, with the goal of finding the best theoretical distribution fitting the real data. We focus on a set of possible distributions, including: Poisson, Weibull, generalized Pareto, Lognormal and α-Stable. Our results show that the Lognormal is the best distribution in a rural scenario, while the α-Stable is the most realistic one in an urban case. This finding is confirmed for different sample areas, and for most of cellular technologies (i.e., 2G and 3G).**

I. Introduction

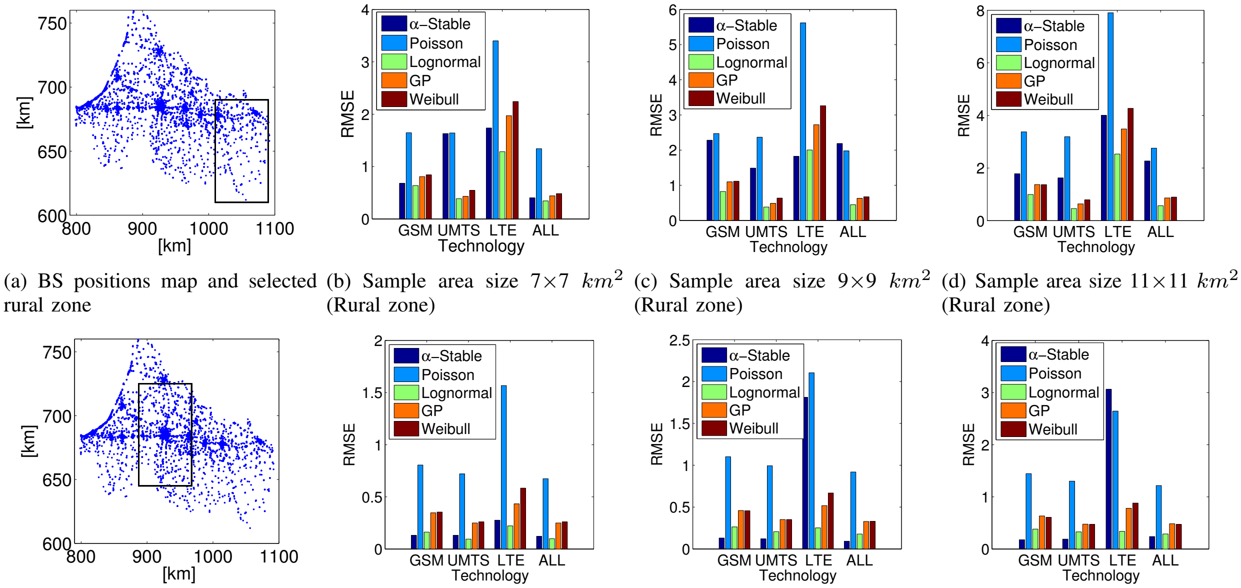
Cellular operators deploy the Base Stations (BSs) with the goal of providing coverage over a territory and capacity to users. As a consequence, BSs are placed following the dynamics of human sociality and living-style people, which tend to aggregate in cities. Not surprisingly, a large number of BSs is deployed in urban areas, where the BS deployment is always 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 [1].

Studying the Spatial Distribution (SD) of BSs over the territory has a double scope. On one side, in fact, it is possible to characterize and to understand the BS deployment in current networks. On the other side, it is possible to provide insights about the evolution of the network as well. Specifically, the definition of theoretical models to predict the SD of BSs is gaining significant importance [2]. In the past, one of the most popular models assumed that the BSs were placed at the corners of an hexagon, forming a hexagonal grid. Even though this model was useful to the industry for the implementation of frequency reuse, it failed to characterize the actual BS deployment. 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 [3], thanks to the fact that it is tractable, i.e., a closed form can be easily obtained for expressing the coverage probability or the rate distribution. However, its actual matching with a realistic BS deployment is an open issue [4], [5]. 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 distribution, including: Pareto, Lognormal, Weibull and α-Stable [2]. The common feature of these 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.

In this context, several questions arise, like: What it the best theoretical model that matches the real SD? What is the impact of the scenario? Is the best distribution changing across different technologies (like 2G, 3G and 4G)? The answers to these questions is the goal of this paper. More in depth, we consider two realistic case studies from a European country (namely one urban and one rural). We then compute the SD of the BSs and we fit the candidate distributions to the real data. We take into account Poisson, Pareto, Lognormal, Weibull and α-Stable as possible candidates. Our results, obtained across 2G/3G/4G technologies, show that the Lognormal distribution is the best candidate for rural environments, while the α-Stable precisely matches the real SD in the urban scenario. These facts stimulate further research towards to definition of tractable models tailored to specific zones (i.e., urban and rural ones).

II. Input Data and Reference Distributions

We consider a region of Italy, which spans over 22,000 km^2. In this region, the population is more than 6.5 millions in total, and more than 5,000 BSs are deployed in order to serve the users. We then select two distinct zones in the region: the first one is a urban zone of 80×80 km^2, centered around the main city of the region, while the second one is a rural zone of 80×80 km^2, without big cities. For each zone we have given the GPS coordinates of the BSs. In order to obtain the spatial positions of BSs we have first applied the Universal Transverse Mercator (UTM) conformal projection [6] in order to map the GPS coordinates to a plane. Fig. 1(a) and Fig. 1(e) report the BS positions and the selected zones. Given the BS positions, we then compute the empirical SD of the BS density. Initially, we sample each zone (urban and rural) 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 obtain the Probability Density Distribution (PDF) of the BS densities. This SD is then used as reference one against the possible candidates (i.e., Poisson, Generalized Pareto (GP), Weibull,



(a) BS positions map and selected rural zone (b) Sample area size 7×7 km^2 (Rural zone) (c) Sample area size 9×9 km^2 (Rural zone) (d) Sample area size 11×11 km^2 (Rural zone)

(e) BS positions map and selected urban zone (f) Sample area size 7×7 km^2 (Urban zone) (g) Sample area size 9×9 km^2 (Urban zone) (h) Sample area size 11×11km^2 (Urban zone)

Fig. 1. Results from the rural zone and the urban one.

Lognormal and α-Stable). For each candidate distribution, we estimate the unknown parameters by applying the Maximum Likelihood Estimation (MLE). For estimating the parameters of the α-Stable distribution, we adopted a similar procedure like the one reported in [2], due to the fact that the closed form for this distribution does not always exist.

III. Results

We first consider the rural zone, reported in Fig. 1(a). We then sample the zone with the following sample sizes: 7×7 km^2, 9×9 km^2, 11×11 km^2. Moreover, we consider different set of cellular networks, including: only 2G BSs (GSM), only 3G ones (UMTS), only 4G enhanced NodeBs (LTE), or all of them. We then compute the Root Mean Square Error (RMSE) of the different fittings against the empirical SD. This metric is used as a term of comparison for the different distributions.[1] Fig. 1(b)-1(d) report the obtained results for the different cellular technologies and for the different sample area sizes. 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. This fact is confirmed across the different technologies, and for the different area sizes. In the following, we have repeated our analysis on the urban zone, which is highlighted in Fig. 1(e). Fig. 1(f)-1(h) report the obtained results. Interestingly, the α-Stable distribution tends to achieve the lowest RMSE in most of the cases. This is especially true when the sample size is increased, i.e., the probability of under-sampling the area is lower. Focusing on the different cellular technologies, the α-Stable experiences an high RMSE for the LTE technology for sample areas 9×9 km^2 and 11×11 km^2, respectively. This may be due to the fact that the 4G technology is not yet fully deployed in the territory, resulting in an empirical SD which may be hardly captured by the α-Stable model. On the contrary, when more mature technologies are considered (like 2G or 3G), the α-Stable model is close to the real SD, resulting in the lowest RMSE.

[1] We leave the investigation of more complex metrics to compare the distributions as future work.

IV. Conclusions

We have analyzed the SD for a set of BSs in a urban zone and a rural one of a European country. Our results show that the Lognormal distribution is the best candidate to fit the real SD of BSs in rural zones, while the α-Stable is the best one in urban ones. This fact suggests that urban and rural zones should be considered separately for modeling the SD of BSs. As next step, we plan to deploy universal and tractable models for both urban and rural environments, as well to consider the impact of different tiers (e.g., Macro BSs and Micro BSs).

References

[1] N. Bhushan *et al.*, “Network densification: the dominant theme for wireless evolution into 5G,” *IEEE Commun. Magazine*, vol. 52, no. 2, pp. 82–89, 2014.

[2] Y. Zhou *et al.*, “On the α-stable distribution of base stations in cellular networks,” *IEEE Commun. Letters*, vol. 19, pp. 1750–1753, Aug. 2015.

[3] J. Andrews *et al.*, “A tractable approach to coverage and rate in cellular networks,” *IEEE Trans. On Wireless Commun.*, vol. 59, pp. 3122–3134, Nov. 2011.

[4] B. Blaszczyszyn *et al.*, “Wireless networks appear poissonian due to strong shadowing,” *IEEE Trans. On Wireless Commun.*, vol. 14, no. 8, pp. 4379–4390, 2015.

[5] Y. Zhou *et al.*, “Two-tier spatial modeling of base stations in cellular networks,” in *Proc. IEEE PIMRC 2014*, (Washington, DC, USA), Sept. 2014.

[6] E. Grafarend, “The optimal universal transverse mercator projection,” in *Geodetic Theory Today*, pp. 51–51, Springer, 1995.