

Measuring the Spatial Heterogeneity of Outdoor Users in Wireless Cellular Networks Based on Open Urban Maps

Meisam Mirahsan, Rainer Schoenen, Sebastian S. Szyszkowicz, Halim Yanikomeroglu

Abstract—Wireless cellular network planning benefits from accurate and realistic, yet relatively simple and manageable, spatial traffic models. User locations in cellular networks are often modeled as homogeneous (uniform) Poisson point processes (PPPs). However, the real user distributions are seldom purely homogeneous. Network users are usually concentrated at social attractors such as residential and office buildings, shopping malls, and bus stations. Wireless spectral efficiency depends significantly on the users’ spatial heterogeneity, and thus relevant spatial traffic generators and models are important. In future (5G) networks, for which device-to-device (D2D), millimeter-wave (mmWave), and small-cell deployments in Heterogeneous Networks (HetNets), are promising technologies, it will become more important to have spatial traffic models which can represent the broad possibilities from completely homogeneous cases (e.g., a deterministic lattice) to extremely heterogeneous cases (e.g., highly clustered scenarios).

In this paper, we study the spatial traffic heterogeneity of outdoor users in the denser areas of the city center of Paris, France. The building shape data is freely available from the OpenStreetMaps project. We measure the heterogeneity via a second-order statistic: the Coefficient of Variation (CoV) of two spatial metrics of the resulting point process: the Voronoi cell areas and the Delaunay cell edge lengths. The expected value of the CoV of these metrics allows us to study how the heterogeneity increases with the density of users. Moreover, we find that the statistical distribution of both these metrics is close to Weibull. Our results illustrate that the topology of the buildings in the city imposes a significant degree of heterogeneity on the spatial distribution of the wireless traffic.

Index Terms—Heterogeneous Spatial Traffic Distribution, Het-Nets, Wireless Cellular Networks, Open Map Data, Open-StreetMaps.

I. INTRODUCTION

The enhancement and proliferation of wireless devices with advanced communication technology and excessive processing power drives exponential growth in the wireless traffic demand. To cope with this massive demand, network operators are forced to improve network performance and provide the required wireless capacity. The statistics of the signal-to-interference-plus-noise ratio (SINR) are key to the performance analysis of wireless cellular networks. The signal strengths and interference levels depend strongly on the network geometry, i.e., the relative positions of the transmitters

and the receivers. Accordingly, in wireless cellular networks, the spatial statistics of the traffic demand, i.e., the user equipment (UE) distribution, as well as that of the service points, i.e., the base station (BS) distribution, have a direct impact on the network performance. Therefore, planning, design, and deployment of cost-effective and high-capacity wireless networks requires accurate and realistic, yet relatively simple and manageable, traffic models in both the time and space domain.

Traffic demand modeling in the time domain has been well-investigated in the literature [1, 2]. Traditionally, in voice-only networks, homogeneous Poisson process models were accurate enough to model traffic in time. After the emergence of different applications, such as video and data with variable data rate demands, the Poisson model failed to capture the traffic statistics [1]; as a result, various heterogeneous (super-Poisson) traffic models based on the hidden Markov model (HMM), Markov modulated Poisson process (MMPP) [2], and other stochastic methods have been proposed in the literature and used for performance analysis.

Similarly, in the space domain, a popular approach in modeling and analysis of the supply and demand in heterogeneous wireless cellular networks has been the use of two independent Poisson point processes (PPPs): one for the locations of the BSs, and one for those of the UEs [3–6]. This popular approach has a major shortcoming: although the PPP model may be a fitting one for the BS locations, it is less adequate for the UE locations, since it cannot be tuned to represent spatial heterogeneity (non-uniformity). There is still relatively little literature that takes into account the heterogeneous spatial distribution of the traffic demand in wireless cellular networks [7–15].

Wireless spectral efficiency can be significantly affected by the UEs’ spatial heterogeneity [8, 9], and thus relevant spatial traffic generators and models are important. Future (5G) wireless networks will likely be based on technologies where this heterogeneity is likely to significantly affect performance. Device-to-device (D2D) communications, where UEs can communicate in direct links, are sensitive to the distances between neighbouring UEs [16], and neighbouring links may cause or suffer severe interference because of the short distances involved [17]. Communications at millimeter wave (mmWave) frequencies, which are proposed for 5G networks [18–21], are expected to be very sensitive to the shapes and layouts of buildings. For example, because of the

M. Mirahsan, R. Schoenen, S. Szyszkowicz, and H. Yanikomeroglu are with the Department of Systems and Computer Engineering, Carleton University, Ottawa, ON, Canada (e-mail: {mirahsan, sz, rs, halim}@sce.carleton.ca).

This work is supported in part by Huawei Canada Co., Ltd., and in part by the Ontario Ministry of Economic Development and Innovations ORF-RE (Ontario Research Fund - Research Excellence) program.

low expected wave penetration through building walls [21], it makes sense to analyze indoor and outdoor environments separately; furthermore, since indoor UEs' traffic may be off-loaded to Wi-Fi and femto-cells, the study of the distribution of the outdoor UEs becomes very interesting.

The main contribution of this paper is the study of the heterogeneity of the spatial distribution of outdoor users based on building data taken from the OpenStreetMap (OSM) project [22], using the city center of Paris, France, as an example. The authors in [10] note that it is generally difficult to obtain real-world location data of wireless devices; the freely-available building data from OSM provides a large source of urban maps that can give an approximation of the real spatial user distributions: the buildings impose a 'thinning' on the PPP (which results in inhomogeneity[10]), while the large amount of available map data ensure statistically significant results. The resulting point process is investigated via random tessellation metrics and approach developed in [8], which uses the second-order statistic 'the Coefficient of Variation (CoV)' of two scalar traffic metrics in the space domain: that of the 'Voronoi cell area' and that of the 'Delaunay cell edge length'. We also plot the sample cumulative distribution function (CDF) of each of these two traffic metrics, and find they fit the Weibull distribution well. A good statistical understanding of these spatial traffic metrics can be valuable for designing spatial traffic models which are tunable and realistic.

The remainder of this paper is organized as follows: In Section II, the methodology of the extraction and processing of the maps is described. The two scalar traffic metrics that are used to capture spatial traffic characteristics are described in Section III. In Section IV, the statistics of the traffic distribution based on the open map data are investigated and the results are presented. Section V concludes the paper.

II. MAP EXTRACTION AND PROCESSING

Maps of the city center of Paris, France are obtained from the OSM project [22], which is a volunteer project whose purpose is to recreate vector maps of the whole earth for open access. These maps contain, among others, the 'footprints' of the various buildings, i.e., the 2D polygons that describe the shape of each building as it stands on the ground. OSM data has previously been used for studying the wireless channel in a European historical urban center, but only for one specific city block [23]. In this study, we perform a statistical study of an entire historical city center cut up into smaller square tiles.

Central Paris has several interesting characteristics that make it a good choice for study, among which:

- It is an important urban center, with a large resident, business, and tourist populations, and therefore deserves to be studied for wireless system planning.
- Both its historical and current importance make it a city of great interest to the volunteer community: thus we can expect the building data to both reliable and complete, or at least close enough for statistical purposes.

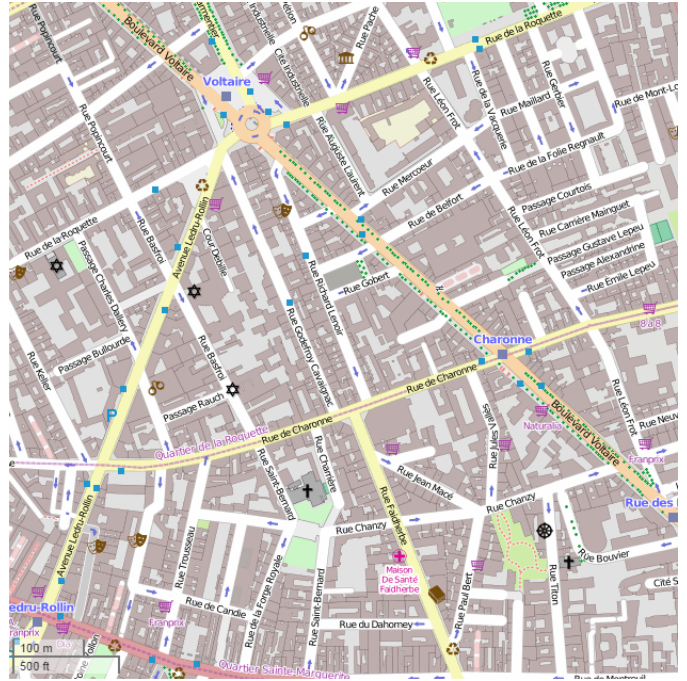


Fig. 1. An example of a city area as visualized on the OSM website.

- It is mostly topographically flat. The building heights are also quite regular, most measuring around 20m (7 floors). The 2D representation of the city is thus sufficient for most purposes.

The data are extracted as follows:

- 1) A master file containing the whole city is downloaded.
- 2) City tiles of size $1\text{km} \times 1\text{km}$ are extracted from the master file. We consider an area of $10\text{km} \times 8\text{km}$, with consecutive tiles shifted 500m both horizontally and vertically (thus the same area may be considered up to 4 times): this results in 19×15 total tiles.
- 3) For each tile, the building shapes and types are extracted. Some unnecessary building types (e.g., bridges) are removed. The building node coordinates are converted from spherical earth coordinates to local rectangular coordinates, flattened at the center of the $1\text{km} \times 1\text{km}$ tile (this results in a horizontal flattening error of a few centimeters in each tile corner).
- 4) The building polygons are then merged into 'islands', of which the interiors ('courtyards') are filled (they are considered as indoor).
- 5) We only keep tiles that have a high building coverage. In this study, a tile is kept if its total 'island' coverage area is above 50 percent for the whole $1\text{km} \times 1\text{km}$ tile, as well as over 40 percent for each of the four $500\text{m} \times 500\text{m}$ corner sub-tiles. The purpose of this filtering is to focus our study on high-building-density areas, which are of most interest, and avoid areas that contain many parks, rivers, canals, and other areas of low building and user density. This results in 83 out of 285 tiles being kept.

An example city tile, visualized on the OSM website, and

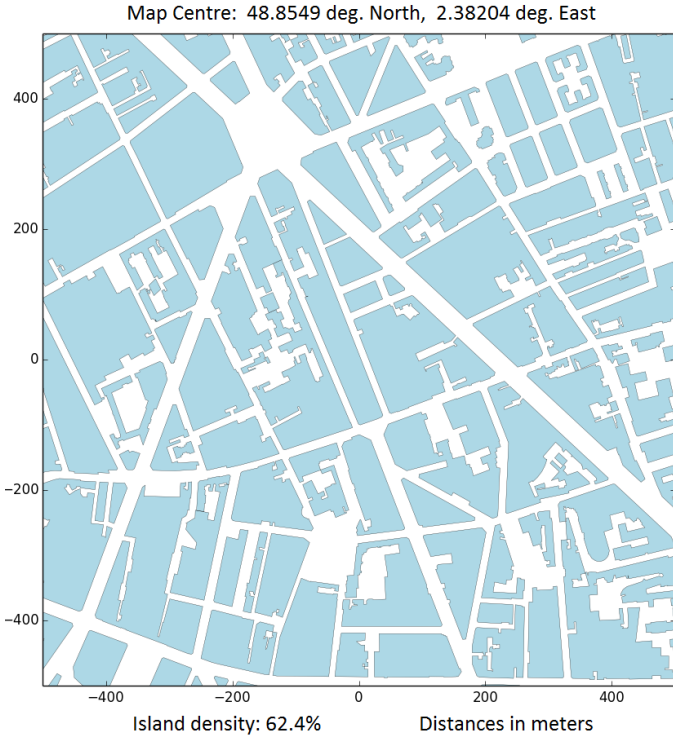


Fig. 2. The city tile from Fig. 1 after processing, showing indoor (colored) and outdoor (white) locations.

then processed by the algorithm above, is shown in Figs 1 and 2, respectively.

III. TRAFFIC METRICS AND STATISTICS

In order to measure the inhomogeneity of the point process of the users located in the outdoor areas of the city tiles, we use two metrics developed in [8] based on the point pattern's Voronoi tessellation, as well as its dual: the Delaunay triangulation.

Given a set of points $P = \{p_1, p_2, \dots, p_n\}$ in two dimensional space \mathbb{R}^2 , the Voronoi tessellation $T = \{c_{p_1}, c_{p_2}, \dots, c_{p_n}\}$ is the set of cells such that every location $y \in c_{p_i}$, is closer to p_i than any other point in P . This can be expressed formally as

$$c_{p_i} = \{y \in \mathbb{R}^2 : |y - p_i| \leq |y - p_j| \text{ for } i, j \in 1, \dots, n\}. \quad (1)$$

The Voronoi tessellation in \mathbb{R}^2 has the property that each of its vertices is given by the intersection of exactly three Voronoi cells. The corresponding three points define a Delaunay cell. Figure 3 illustrates an example pattern of points with its Voronoi and Delaunay tessellations.

Various statistical quantities can now be considered for measuring a point pattern. The 'Voronoi cell area' (V) is the first natural choice. Various statistics of the Voronoi cells for a Poisson point process (Poisson-Voronoi Tessellation) are well investigated in the literature [24]. The second metric is the 'Delaunay cell edge length' (D), whose various statistics are investigated in [25].

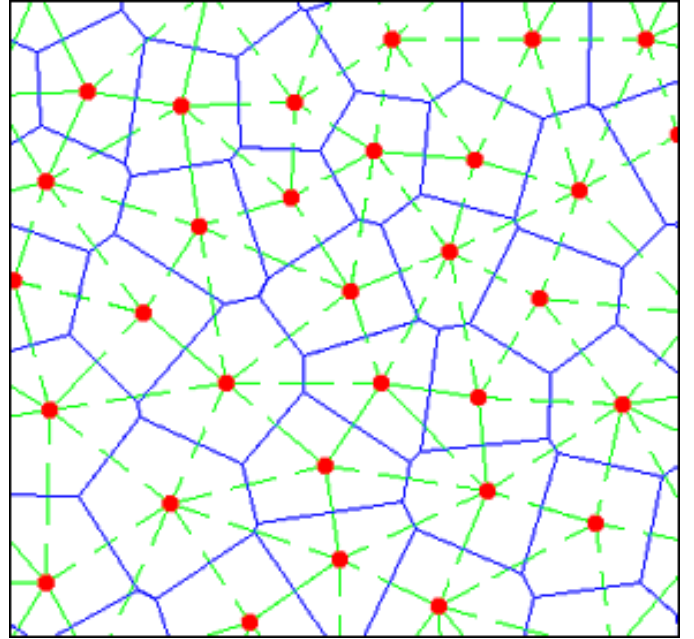


Fig. 3. The Voronoi tessellation (solid lines) and the Delaunay tessellations (dashed lines) of a sample set of points.

Here we use the CoV (i.e., the standard deviation normalized by the mean) of the traffic metrics V and D to capture the heterogeneity of spatial traffic. The advantage of the CoV is that it captures the amount of heterogeneity via only one scalar value. It is shown in [24] and [25] that the expected value of the CoV of V and D are, respectively, 0.530 and 0.492 for a homogeneous PPP. We therefore use the normalized metrics

$$C_V = \frac{\sigma_V}{0.530\mu_V} \quad (2)$$

and

$$C_D = \frac{\sigma_D}{0.492\mu_D}, \quad (3)$$

such that $\mathbb{E}(C_V) = \mathbb{E}(C_D) = 1$ for a homogeneous PPP.

IV. TRAFFIC GENERATION AND SIMULATION RESULTS

On each of the 83 high building-density $1000\text{m} \times 1000\text{m}$ maps obtained from the procedure described in Section II, we distribute points (network users) according to a homogeneous PPP with density λ , and only keep the points which fall in outdoor locations. As a result, the density of points in all outdoor areas is constant across all the maps. Figure 4 illustrates users distributed in the outdoor areas of the map in Figure 2, and the resulting Voronoi tessellation. Figure 5 shows the CoV values of the Voronoi cell areas and the Delaunay cell edge lengths of the resulting spatial traffic patterns averaged over all 83 maps. We observe that low user densities have CoVs close to 1: the same as that of a homogeneous PPP, while high user densities result in increased CoVs, as the spatial heterogeneity becomes more pronounced.

For a more accurate statistical description of the outdoor network users, we investigate the distributions of the two

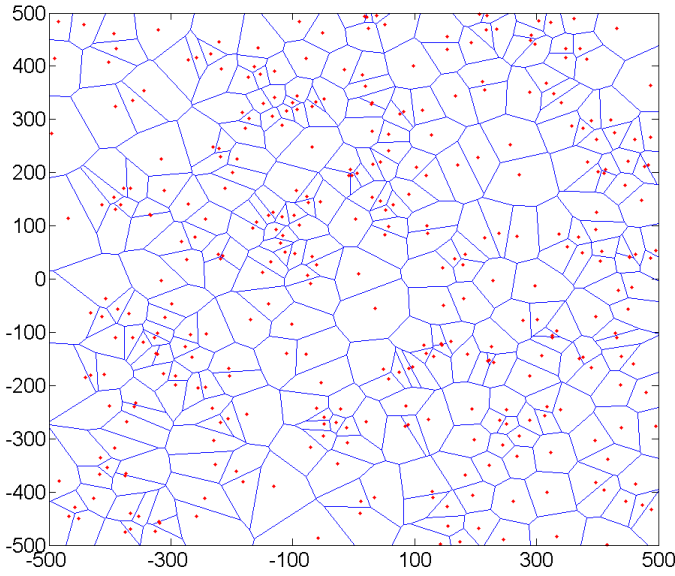


Fig. 4. The Voronoi tessellation of a sample distribution of points (network users) (with density $\lambda = 1000$ users per km^2 of outdoor area) in outdoor locations of the city tile from Fig. 1.

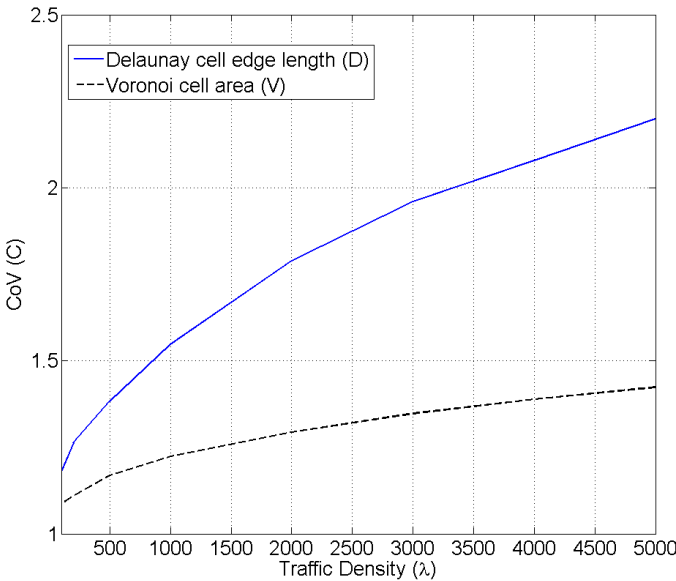


Fig. 5. The CoV values of the 'Voronoi cell area' and the 'Delaunay cell edge length' metrics for building-dense areas of central Paris as averaged over 83 1 km^2 tiles. The traffic density is in terms of the number of users per km^2 of outdoor area.

traffic metrics, and compare them with several well-known distribution functions. Figures 6 and 7 show the CDFs and the fitness of different distributions for, respectively, the 'Voronoi cell areas' and the 'Delaunay cell edge lengths' of the traffic point patterns for $\lambda = 1000$ users per km^2 of outdoor area. To evaluate the goodness-of-fit of the simulated data to various distributions, we compare their log-likelihood (LL) values [26]. The Weibull distribution shows the highest fitness. Its

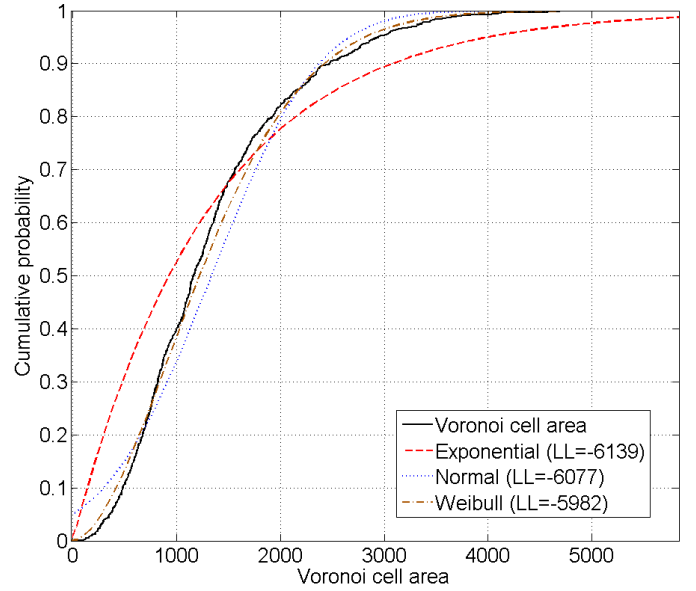


Fig. 6. The CDF of the 'Voronoi cell areas' of the traffic patterns (with a density of 1000 users per km^2 of outdoor area) and the log-likelihood value of various distribution functions. The cell areas are in meters squared.

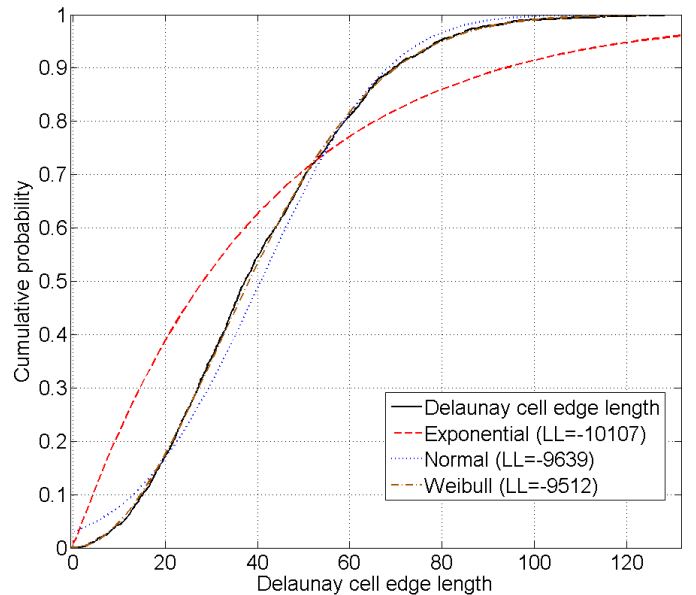


Fig. 7. The CDF of the 'Delaunay cell edge lengths' of the traffic patterns (with a density of 1000 users per km^2 of outdoor area) and the log-likelihood value of various distribution functions. The cell edge lengths are in meters.

CDF is given by:

$$\begin{cases} 1 - e^{-(x/\alpha)^k} & \text{if } x \geq 0, \\ 0 & \text{if } x < 0, \end{cases} \quad (4)$$

where, α is the scale parameter, and k is the shape parameter. The parameters of the Weibull distributions that best fit the simulated data, for both V and D , are summarized in Table I. These distributions give a more accurate statistical description of the neighbourhood of each point.

TABLE I

The parameters of the fitted Weibull distributions for both spatial metrics, when the user density is 1000 users per km² of outdoor area.

Metric	Scale param.	Shape param.
Voronoi cell area	1594.64	1.75071
Delaunay cell length	46.9586	2.00505

V. CONCLUSION

In this paper, we studied the spatial heterogeneity of network users as distributed on the outdoor areas of maps of dense parts of central Paris, France; our approach can be reproduced on any major urban area for which building data is openly available. The availability of a large amount of maps ensures a large sample for statistical analysis.

We studied the CoVs and the CDFs of two spatial traffic metrics: the ‘Voronoi cell area’ and the ‘Delaunay cell edge length’ of the outdoor user patterns. Our results show that, even without considering population distribution information, the thinning induced by the layout of the buildings can lead to significant heterogeneity of the traffic in the space domain as the density of users increases. The CDFs of the spatial metrics give a more accurate statistical description of the neighbourhood of each point.

This work can be extended in a number of directions. Adding information on population density and social attractors would increase the accuracy of the model, with an expected increase in the heterogeneity as well. Moreover, combining the spatial distribution with user rate demands in the time domain would also increase the overall heterogeneity, as well as the accuracy of the model.

ACKNOWLEDGMENT

The authors would like to thank Andrés Lou for his collaboration and Dr. Gamini Senarath and Dr. Ngoc Dao (Huawei Technologies Canada Co., LTD.) for their support.

REFERENCES

- [1] V. Paxson and S. Floyd, “Wide area traffic: the failure of Poisson modeling,” *IEEE/ACM Transactions on Networking (ToN)*, vol. 3, no. 3, pp. 226–244, June 1995.
- [2] W. Fischer and K. Meier-Hellstern, “The Markov-Modulated Poisson Process (MMPP) cookbook,” *Performance Evaluation*, vol. 18, no. 2, pp. 149–171, 1993.
- [3] J. G. Andrews, R. K. Ganti, M. Haenggi, N. Jindal, and S. Weber, “A primer on spatial modeling and analysis in wireless networks,” *IEEE Communications Magazine*, vol. 48, no. 11, pp. 156–163, November 2010.
- [4] J. G. Andrews, H. Claussen, M. Dohler, S. Rangan, and M. C. Reed, “Femtocells: Past, present, and future,” *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 3, pp. 497–508, 2012.
- [5] H. S. Dhillon, R. K. Ganti, F. Baccelli, and J. G. Andrews, “Modeling and analysis of k-tier downlink heterogeneous cellular networks,” *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 3, pp. 550–560, 2012.
- [6] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi, “A survey on 3GPP heterogeneous networks,” *IEEE Wireless Communications*, vol. 18, no. 3, pp. 10–21, June 2011.

- [7] M. Mirahsan, R. Schoenen, and H. Yanikomeroglu, “HetHet-Nets: Heterogeneous traffic distribution in heterogeneous wireless cellular networks,” to appear in *IEEE Journal on Selected Areas in Communications, Special Issue on Recent Advances in Heterogeneous Cellular Networks*, June 2015.
- [8] M. Mirahsan, R. Schoenen, Z. Wang, H. Yanikomeroglu, and M. St-Hilaire, “Unified and non-parameterized statistical modeling of temporal and spatial traffic heterogeneity in wireless cellular networks,” in *Proc. IEEE ICC’14 - Workshop on 5G Technologies (ICC’14 WS - 5G)*, Sydney, Australia, June 2014.
- [9] M. Mirahsan, R. Schoenen, and H. Yanikomeroglu, “Statistical modeling of spatial traffic distribution with adjustable heterogeneity and bs-correlation in wireless cellular networks,” in *Proc. IEEE Global Communications Conference (Globecom)*, Austin, TX, USA, December 2014.
- [10] C. Bettstetter, M. Gyarmati, and U. Schilcher, “An inhomogeneous spatial node distribution and its stochastic properties,” in *Proc. ACM Symposium on Modeling, Analysis, and Simulation of Wireless and Mobile Systems*, New York, USA, 2007, pp. 400–404.
- [11] C. Qvarfordt and P. Legg, “Evaluation of LTE HetNet deployments with realistic traffic models,” in *Proc. IEEE International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, Barcelona, Spain, September 2012, pp. 307–311.
- [12] “3GPP TR 36.814: Further advancements for E-UTRA physical layer,” 3GPP, Tech. Rep., 2010.
- [13] H. Dhillon, R. Ganti, and J. Andrews, “Modeling non-uniform UE distributions in downlink cellular networks,” *IEEE Wireless Communications Letters*, vol. 2, pp. 339–342, June 2013.
- [14] M. Taranez and M. Rupp, “Performance of femtocell access point deployments in user hot-spot scenarios,” in *Proc. IEEE Telecommunication Networks and Applications Conference (ATNAC)*, Brisbane, Australia, November 2012.
- [15] D. Lee, S. Zhou, X. Zhong, Z. Niu, X. Zhou, and H. Zhang, “Spatial modeling of the traffic density in cellular networks,” *IEEE Journal on Wireless Communications*, vol. 21, no. 1, pp. 80–88, February 2014.
- [16] S. Hakola, T. Chen, J. Lehtomaki, and T. Koskela, “Device-to-device (D2D) communication in cellular network - performance analysis of optimum and practical communication mode selection,” in *IEEE Wireless Communications and Networking Conference*, Apr. 2010, pp. 1–6.
- [17] N. Bhushan, J. Li, D. Malladi, R. Gilmore, D. Brenner, A. Damnjanovic, R. Sukhvasi, C. Patel, and S. Geirhofer, “Network densification: the dominant theme for wireless evolution into 5G,” *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 82–89, Feb. 2014.
- [18] S. Singh, M. N. Kulkarni, A. Ghosh, and J. G. Andrews, “Tractable model for rate in self-backhauled millimeter wave cellular networks,” *arXiv preprint arXiv:1407.5537*, 2014.
- [19] M. R. Akdeniz, Y. Liu, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, “Millimeter wave channel modeling and cellular capacity evaluation,” *arXiv preprint arXiv:1312.4921*, 2013.
- [20] J. G. Andrews, S. Buzzi, W. Choi, S. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, “What will 5G be?” *arXiv preprint arXiv:1405.2957*, 2014.
- [21] T. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. Wong, J. Schulz, M. Samimi, and F. Gutierrez, “Millimeter wave mobile communications for 5G cellular: It will work!” *Access, IEEE*, vol. 1, pp. 335–349, 2013.
- [22] “OpenStreetMap,” <http://www.openstreetmap.org/>.
- [23] J. Nuckelt, D. Rose, T. Jansen, and T. Kurner, “On the use of OpenStreetMap data for V2X channel modeling in urban scenarios,” in *Antennas and Propagation (EuCAP), 2013 7th European Conference on*, April 2013, pp. 3984–3988.
- [24] E. Gilbert, “Random subdivisions of space into crystals,” *The Annals of Mathematical Statistics*, vol. 33, no. 3, pp. 958–972, 1962.
- [25] L. Muehe, “Distributional properties of the three-dimensional Poisson Delaunay cell,” *Journal of Statistical Physics*, vol. 84, no. 1-2, pp. 147–167, 1996.
- [26] J. A. Nelder and R. Baker, *Generalized linear models*. Wiley Online Library, 1972.