

# Statistical Modeling of Spatial Traffic Distribution with Adjustable Heterogeneity and BS-Correlation in Wireless Cellular Networks

Meisam Mirahsan, Rainer Schoenen, and Halim Yanikomeroglu  
Department of Systems and Computer Engineering  
Carleton University, Ottawa, ON, Canada  
e-mail: {mirahsan, rs, halim}@sce.carleton.ca

**Abstract**—Future generation (5G and beyond) cellular networks have to deal not only with an extreme traffic demand increase, but also an extreme level of heterogeneity in the distribution of that demand in both space and time. Traffic modeling in the time domain has been investigated well in the literature. In the space domain, however, there is a lack of statistical models for the heterogeneous User Equipment (UE) distribution beyond the classical Poisson Point Process (PPP) model. In this paper, we introduce a methodology for the generation and analysis of spatial traffic which allows statistical adjustments. Only two parameters, namely, Coefficient of Variation (CoV) and Correlation Coefficient, are adjusted to control the UE distribution heterogeneity and correlation with Base Stations (BSs). The methodology is applied to cellular networks to show the impact of heterogeneous network geometry on network performance.

**Index Terms**—Statistical Modeling, Spatial Traffic Distribution, Stochastic Geometry, Point Process, Cellular Network.

## I. INTRODUCTION

Due to the rapid proliferation of a broad range of wireless devices such as smart phones and tablets with powerful processing capabilities and ambitious data rate expectations, the number of users relying on the wireless cellular infrastructure for Internet connectivity, as well as the traffic demand per user, are increasing dramatically while wireless resources remain limited. Getting the maximum performance gain out of these limited resources is the main focus of wireless cellular research. Adjustable and realistic network traffic models in the space domain (i.e., UE distribution in the network) and the time domain (i.e., data rate demand for each UE over time) play a significant role in characterizing and analyzing the performance of a network as they provide traffic patterns to test network behavior under various conditions.

Traffic modeling in the time domain has been investigated well in the literature [1–6]. In the traditional voice-only networks, the homogeneous Poisson models were sufficiently accurate to model traffic in time. After the emergence of different applications such as video and data with variable

data rate demands, Poisson modeling has become insufficient to capture traffic statistics [1], as such various heterogeneous (Super-Poisson) traffic models based on Hidden Markov Models (HMM) and Markov Modulated Poisson Process (MMPP) [2], as well as other methods, have been proposed in the literature and used for performance analysis. The basic idea behind these models is to generate traffic with variable rate over time which is controlled by a Markovian process.

Traffic Generators (TGs) like MMPP and HMM receive a number of Input Parameters (TGIPs) to generate traffic with various statistical properties. For instance, MMPP receives the number of Markov chain states, state transition matrix and traffic rate at each state as its TGIPs. Therefore, a main track of research in time domain is to fit TGIPs to generate traffic patterns with similar statistical properties to a given real traffic trace [6–8]. This problem is called parameter fitting.

In the space domain, on the other hand, the unrealistic homogeneous Poisson modeling (e.g., in IMT-Advanced evaluation guidelines [9]) is still commonly used by researchers for UE distribution and performance analysis of wireless cellular networks. There are only few works in the literature which consider adjustable heterogeneous or BS-correlated traffic modeling in the space domain [10–13]. None of the existing papers, however, considers the statistical properties of traffic and parameter fitting. The modeling methodology used in the existing papers blindly changes the TGIPs to generate traffic patterns with unknown statistics and measures network performance metrics at the output. We are not aware of any research in the literature that considers the statistical modeling of spatial traffic.

As a first step to fill this void, the main contributions of this paper are as follows: 1) A spatial traffic modeling approach with adjustable statistical properties, i.e., heterogeneity and correlation with BSs, is introduced which regulates spatial traffic with only two parameters. 2) The main statistical properties of the spatial traffic distribution are characterized and formulated. It is shown that besides first-moment statistic (i.e., mean density), the normalized second-moment statistic (CoV) and joint moment statistics with BSs (specifically correlation coefficient), can be used to describe UE distribution with sufficient accuracy. 3) Stochastic geometry is used to generate

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.

The authors would like to acknowledge Ziyang Wang and Dr. Marc St-Hilaire, Carleton University, for their valuable comments.

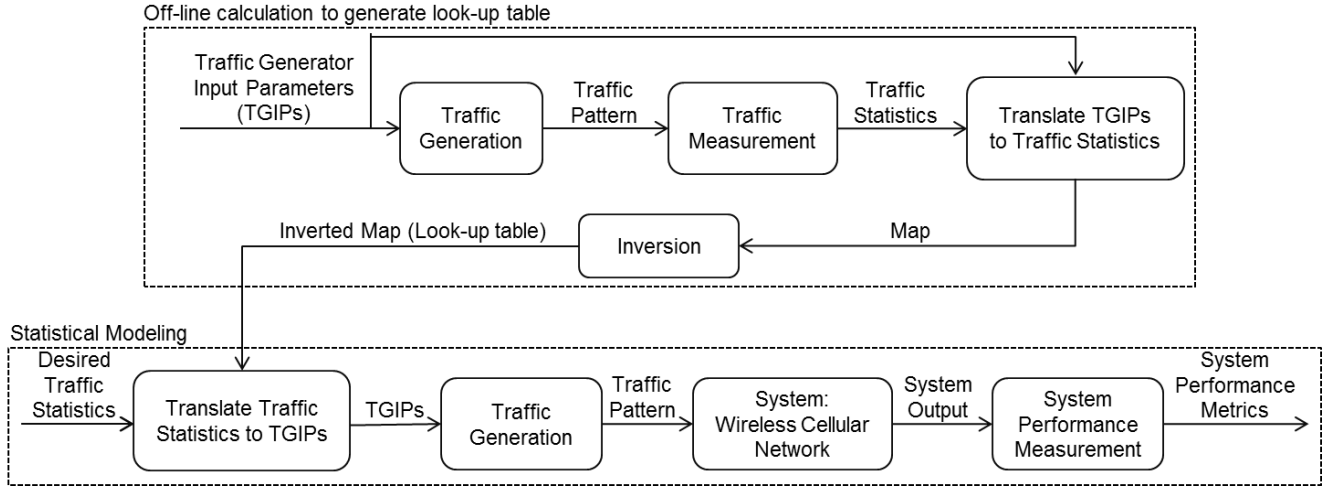


Fig. 1. The desired statistical properties of traffic as the modeling inputs are translated to the appropriate TGIPs and traffic with known characteristics is generated to be used for network performance analysis (lower box). The look-up table for translation is generated in advance via off-line calculations (upper box).

realistic and adjustable traffic models with desired properties, and the effects of realistic traffic modeling on the performance of wireless cellular networks is illustrated.

The remainder of this paper is organized as follows: Section II introduces the proposed traffic modeling methodology. Spatial traffic statistics are investigated in Section III. Section IV describes the traffic generation method used in this paper. Experimental results are presented in Section V, and the paper is concluded in Section VI.

## II. STATISTICAL TRAFFIC MODELING METHODOLOGY

Figure 1 demonstrates the proposed traffic modeling methodology in the space domain. A similar methodology has been used in the literature for traffic modeling in the time domain [6]. The desired statistical properties of traffic as the modeling inputs are translated to the appropriate TGIPs (parameter fitting) and traffic with known characteristics is generated to be used for network performance analysis. The translation look-up table is developed in advance by inversion of the map from TGIPs to measured statistical properties.

For the measurement of traffic patterns, first, an appropriate measure must be selected. In the time domain, Inter-Arrival Time (IAT) is the most popular and well accepted measure. In [14], equivalent measures for statistical measurement of spatial traffic are introduced by the same authors. It is shown that when modeling UE distribution as a two-dimensional (2D) point pattern in the space domain, Voronoi cell areas ( $V$ ) and Delaunay cell edge lengths ( $E$ ) associated with the points can be measured as equivalents of IAT in the time domain.

Perfect modeling requires all statistics including the probability distribution functions (PDFs) and auto-correlation functions of traffic measure to match between the generated traffic patterns and the real traffic traces. However, a perfect match is not practical and requires extremely complicated models with a very large number of parameters. Therefore, simplified models are commonly used which only consider the first few

moments of traffic. In the time domain, usually the mean, CoV, and the third moment as well as auto-correlation are considered to match with the real traffic trace. Spatial traffic statistics are described in Section III.

## III. SPATIAL TRAFFIC STATISTICS

Spatial traffic in wireless cellular network can be modeled as a 2D point pattern  $U \in \mathbb{R}^2$  which is generated by a generator point process  $\Phi_U$  with density function  $\Lambda(x, y)$  at any point  $(x, y) \in \mathbb{R}^2$ . In this paper, the mean,  $\mu_E$ , and the CoV,

$$C_E = \frac{\sigma_E}{\mu_E}, \quad (1)$$

are the desired statistics of traffic where  $E$  is the Delaunay cell edge lengths associated with each point and  $\sigma_E$  is the standard deviation of  $E$  (the third-moment and auto-correlation are stated as future extensions in Section VI). It is shown in [14] that for a Poissonian distribution, the normalized CoV is  $C = 1$ , for a sub-Poisson distribution (more homogeneous than Poisson)  $0 \leq C \leq 1$ , and for a super-Poisson distribution (more heterogeneous than Poisson)  $C \geq 1$ .

Along with the mean and the CoV, a very important statistical property of traffic in space which affects the network performance and does not exist in the time domain is the bias or dependence of the UE distribution to the BS distribution. BS locations in wireless cellular network can be modeled by a 2D point pattern  $B \in \mathbb{R}^2$ , generated by a point process  $\Phi_B$ , whose Voronoi tessellation divides the entire field into Voronoi cells associated with each BS consisting of all area closer to that BS than to any other. A sample realization is shown in Fig. 2(a).

To measure the joint distribution of UEs and BSs, every point  $(x, y)$  in the field is associated with a Potential value  $P(x, y) \in [-1, +1]$ . The  $P$  function must have the following properties:

- $P(x, y) = +1$  for cell-center points,

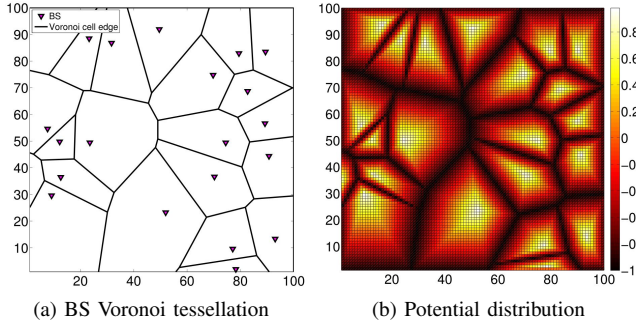


Fig. 2. Voronoi tessellation of BSs divides the whole field into cells associated with each BS consisting of all points closer to that BS than to any other. Every point in the field is then associated with a Potential value between  $-1$  and  $+1$ .

- $P(x, y) = -1$  for Voronoi cell-edge points,
- $\iint_A P(x, y) dx dy = 0$ ,

where  $A$  is the Voronoi cell area. A simple polynomial function which has the above mentioned properties is used in this paper:

$$P(x, y) = \frac{-2d(x, y)^2}{D(x, y)^2} + 1. \quad (2)$$

In the above,  $d(x, y)$  is the distance of the point  $(x, y)$  to the closest cell-center, and  $D(x, y)$  is the length of the line connecting the closest cell-center to the cell-edge through point  $(x, y)$ . Figure 2(b) illustrates the Potential distribution associated to the Voronoi tessellation in Fig. 2(a).

Using the Potential function, the joint moments

$$\mathbb{E} [P^i \Lambda^j], \quad i, j \geq 0, \quad (3)$$

and joint central moments

$$\mathbb{E} [(P - \mu_P)^i (\Lambda - \mu_\Lambda)^j], \quad i, j \geq 0, \quad (4)$$

can be calculated, where  $\mu_P$  is the mean value of  $P$ ,  $\mu_\Lambda$  is the mean value of UE density  $\Lambda$ , and  $\mathbb{E} [x]$  is the expected value of  $x$  [15]. To have a normalized measure, we use correlation coefficient which is defined as

$$\rho = \frac{\sigma_{P\Lambda}}{\sigma_P \sigma_\Lambda}, \quad (5)$$

where

$$\sigma_{P\Lambda} = \mathbb{E} [(P - \mu_P)(\Lambda - \mu_\Lambda)] \quad (6)$$

is the covariance of  $P$  and  $\Lambda$ ,  $\sigma_P$  is the standard deviation of  $P$ , and  $\sigma_\Lambda$  is the standard deviation of  $\Lambda$ . So the  $\rho$  value can be calculated as

$$\rho = \frac{\iint (\Lambda(x, y) - \mu_\Lambda)(P(x, y) - \mu_P) dx dy}{\sqrt{(\iint (\Lambda(x, y) - \mu_\Lambda)^2 dx dy)(\iint (P(x, y) - \mu_P)^2 dx dy)}}. \quad (7)$$

For a UE pattern  $U$  (a realization of  $\Phi_U$ ), the correlation coefficient is defined as

$$\rho = \frac{\sum_{u \in U} P_u}{|U|}, \quad (8)$$

where  $P_u$  is the Potential value at point  $u$  and  $|U|$  is the number of points in  $U$ . A pattern with  $\rho = +1$  means that all UEs have gathered at the cell-centers, a pattern with  $\rho = 0$  means that UE distribution is independent from BSs, and a pattern with  $\rho = -1$  means that all UEs have gathered at the cell-edges.

#### IV. SPATIAL TRAFFIC GENERATION

Various point processes can be used as TG to model UE distribution in space. A thorough study of different point processes can be found in [16]. To model the clustering properties of UEs, cluster point processes are used in this paper. A cluster point process is generated by applying a clustering perturbation, defined below, on a point process.

##### Clustering Perturbation:

“Clustering perturbation of a given (parent) process  $\Phi$  consists of independent replication and displacement of points of  $\Phi$ , with the number of replications of a given point having distribution  $\Upsilon$  and the replicas’ locations having distribution  $\Psi$ . All replicas of an original point form a cluster” [16].

With combination of different parent processes, different replication distributions and different displacement distributions, various point processes with different properties can be generated. A list of popular cluster point processes can be found in [16]. To model the dependence between UE locations and BS locations, Shot-Noise Cox point processes, defined below, are used in this paper.

##### Shot-Noise Cox Point Processes:

“Shot-Noise Cox point processes are clustering perturbations of arbitrary parent processes with replication kernel  $\Upsilon$  having Poisson distribution” [16].

The displacement kernel can be any distribution. Rayleigh distribution with uniform direction around original point is used in this paper. To generate parent processes with various correlation coefficient values with BSs, starting with a homogeneous PPP with mean density  $\mu_\Lambda$ , we define a heterogeneous UE distribution with density function defined as

$$\Lambda_i(x, y, b) = \mu_\Lambda w_i(x, y, b), \quad (9)$$

where  $i$  is the index for BSs,  $b$  is the bias parameter of UEs to the BSs ranging from  $b = -1$  (all UEs gathered at cell-edge) through  $b = 0$  (Poissonian distribution) to  $b = +1$  (all UEs gathered at cell-center), and  $w_i(x, y, b)$  is the weight of the point  $(x, y)$ . The function  $w$  must have the following characteristics:

- 1) If  $b = 0$ , then

$$w_i(x, y, 0) = 1, \quad \forall x \forall y; \quad (10)$$

this results in a Poissonian distribution for  $\Phi_U$ .

- 2) If  $b = 1$ , then

$$w_i(x, y, 1) = A_i \delta(x - x_{c_i}) \delta(y - y_{c_i}), \quad \forall x \forall y; \quad (11)$$

this results in all UEs gathering at cell-centers.

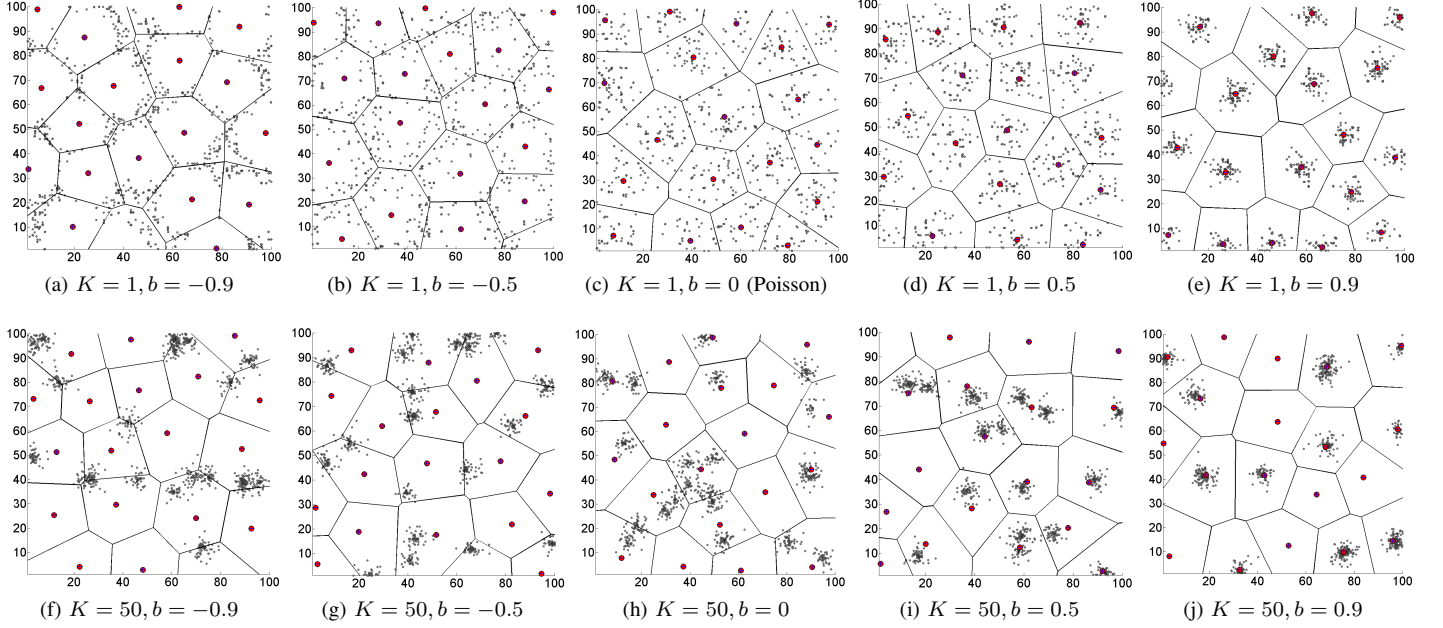


Fig. 3. Sample UE distributions with various statistical characteristics: Increasing  $b$  from  $-1$  to  $+1$ , UEs' bias is changed from cell-edge to cell-center and increasing  $K$  from  $1$  to  $50$ , UEs' clustering changes from Poissonian homogeneous to heterogeneous clustered scenarios.

3) If  $b = -1$ , then

$$w_i(x, y, -1) = \begin{cases} \frac{A_i}{|\mathbf{e}_i|}, & (x, y) \in \mathbf{e}_i, \\ 0, & \text{otherwise;} \end{cases} \quad (12)$$

this results in all UEs gathering at cell-edges.

In the above,  $A_i$  is the Voronoi cell area of BS  $B_i$ ,  $(x_{c_i}, y_{c_i})$  is the cell-center,  $\delta(\cdot)$  is the Dirac delta function,  $\mathbf{e}_i$  is the set of Voronoi cell-edge points, and  $|\mathbf{e}_i|$  is the perimeter of the Voronoi cell area.

Many different functions can be defined for  $w$  which have the above mentioned characteristics. In this paper, we choose  $w_i(x, y, b)$  as

$$\begin{cases} (1+b)N(d(x, y), 0, \frac{1}{|b|} - 1) \\ + (1-b)N(d(x, y), D(x, y), \frac{1}{|b|} - 1), & b \neq 0, \\ 1, & b = 0, \end{cases} \quad (13)$$

where  $N(d, \mu, \sigma)$  is the PDF value at  $d$  using the normal distribution with mean  $\mu$  and standard deviation  $\sigma$ .

Finally, the parent pattern is normalized to have  $N_c$  points which represent the UE clusters.

After generating the parent process, each point in the parent process is replicated  $k$  times, where  $k \sim \text{Poisson}(K)$ .  $K$  is the mean number of UEs per cluster which determines the clustering property of the process, and is defined as

$$K = N_v / N_c, \quad (14)$$

where  $N_v$  is the desired number of UEs. Figure 3 shows sample patterns with various  $b$  and  $K$  values.

To summarize,  $N_v$ ,  $K$ , and  $b$  are the TGIPs which internally regulate the desired spatial traffic statistical properties  $\mu$ ,  $C$ , and  $\rho$ , respectively.

## V. SIMULATION RESULTS

Section V-A presents the simulation results for traffic measurement and modeling, and Section V-B presents the simulation results for network performance analysis.

### A. Traffic Modeling Results

The first step towards statistical modeling of spatial traffic is to measure the statistical properties of traffic patterns generated by the TG for various values of the TGIPs. In a  $1000 \text{ m} \times 1000 \text{ m}$  square field, 19 BSs are assumed to be distributed Poissonian with a minimum distance limit of 200 m between any two BSs. The mean UE density  $\mu_\Lambda$  is fixed to be  $4.75 \times 10^{-4}$  UEs/m<sup>2</sup>. So, 475 UEs (25 UEs per cell) are distributed in the field on average. UEs are distributed using a shot-noise Cox point process whose input parameters are  $b$  (the bias of UEs to the BSs) and  $K$  (the mean number of UEs per cluster) as defined in Section IV. The traffic statistics which are measured are  $C_E$  (the CoV of the Delaunay cell edge lengths of the UEs) and  $\rho$  (the correlation coefficient between UEs and BSs). Figures 4 and 5 illustrate the measured traffic statistics ( $\rho$  and  $C$ ) for different values of the TGIPs ( $b$  and  $K$ ).

The next step for modeling traffic is to use the inverted map as look-up tables and generate traffic with desired statistics. Figures 6 and 7 show the measured traffic statistics versus the desired traffic statistics.

As shown in the results, the measured traffic reflects the desired statistics accurately enough, despite the fact that only the first two moments and the correlation coefficient are fitted.

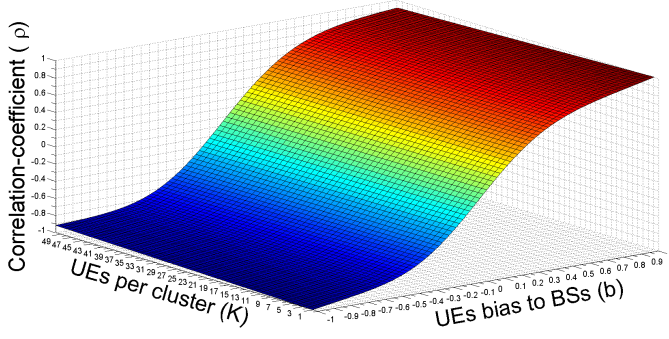


Fig. 4. Correlation coefficient is measured for various values of TGIPs. The results suggest that correlation coefficient is not affected by the parameter  $K$ . As a result, the map is invertible.

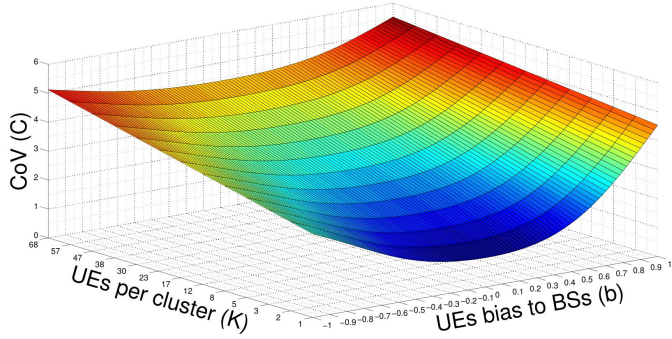


Fig. 5. CoV is measured for various values of TGIPs. Both  $K$  and  $b$  affect traffic CoV. For  $K = 1$  and  $b = 0$  the distribution is Poisson and  $C = 1$  as expected.

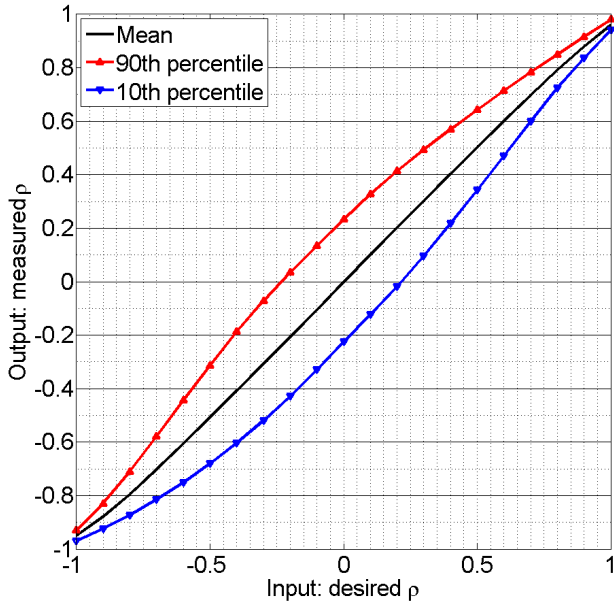


Fig. 6. The horizontal axis shows the desired correlation coefficient and the vertical axis shows the range of measured correlation coefficients from traffic patterns generated by TGIPs derived from the look-up tables ( $\rho$  is averaged for  $K = 1$  to  $K = 50$ ).

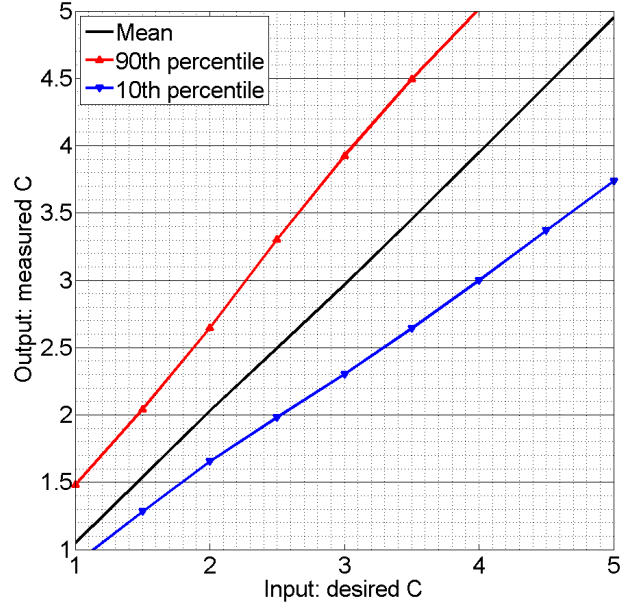


Fig. 7. The horizontal axis shows the desired CoV and the vertical axis shows the range of measured CoVs from traffic patterns generated by TGIPs derived from the look-up tables ( $C$  is averaged for  $b = -1$  to  $b = +1$ ).

## B. Performance Analysis Results

A static snapshot-based system-level simulation is used in this paper. The channel follows the model in IMT-Advanced guideline [9, 17] in which a Line-Of-Sight (LOS) and Non-Line-Of-Sight (NLOS) power-law path-loss model is applied. The downlink signal experiences log-normal shadowing but fast-fading is averaged out. Simulation parameters are listed in Table I.

TABLE I. Simulation parameters based on IMT-advanced [9].

Parameter	Value
Carrier frequency	2 GHz
Bandwidth	20 MHz
Shadowing	Log-normal, std.= 4 for LOS, std.= 6 for NLOS
UE speed	30 km/h
Total BS Tx power	49 dBm
BS antenna height	25 m
Antenna number	SISO
Antenna model	Omni-directional
BS antenna gain	17 dBi
UE antenna gain	0
Traffic model	Full buffer
Number of drops	1000
Scheduling algorithm	Proportional Fair

The mean UE rates for different values of traffic CoV is shown in Fig. 8 and the mean UE rates for different values of traffic correlation coefficient with BSs is shown in Fig. 9. As expected, with increase in traffic heterogeneity, mean user rate in the network is decreased because in an heterogeneous traffic situation some BSs are overloaded and have to share resources



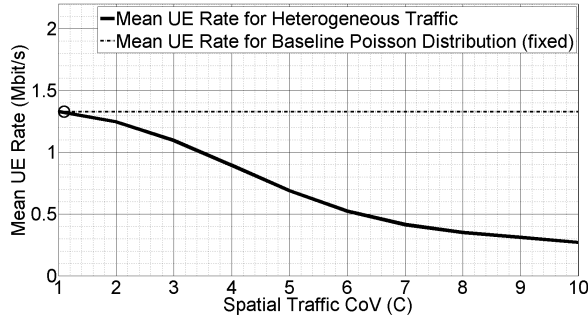


Fig. 8. With increase in traffic heterogeneity, the mean UE rate is decreased. The results are shown for  $\rho = 0$ .

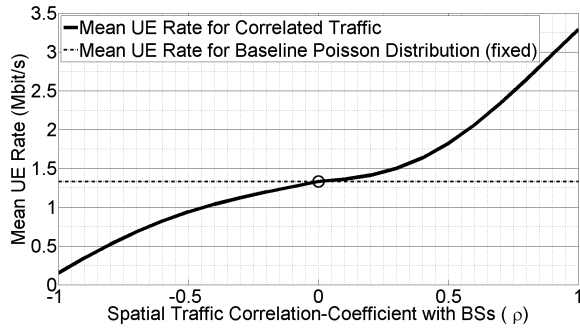


Fig. 9. With increase in traffic bias towards cell-edges, the mean UE rate is decreased and with increase in traffic bias towards BSs, the mean UE rate is increased. The results are shown for  $C = 1$ .

among a large number of user while other BSs don't use their full capacity and remain underloaded. Traffic with positive correlation coefficient with BSs results in higher user rates because users have higher Signal-to-Noise-plus-Interference-Ratio (SINR) and traffic with negative correlation coefficient with BSs results in lower user rates.

## VI. CONCLUSION

In this paper, traffic modeling in the space domain was investigated and a statistical approach was introduced. Only two parameters were used to regulate heterogeneity and correlation of traffic to the BSs. A clustering point process was used to generate realistic and adjustable spatial traffic. As a sample application, the effects of realistic traffic modeling on the performance of wireless cellular networks was illustrated. This work has numerous extensions. It is important to measure auto-correlation as a fundamental statistic which affects performance. It is also essential to consider other point processes and other functions for traffic generation. Various techniques such as the User-In-the-Loop (UIL) [18] can be used to improve network performance for heterogeneous and BS-correlated traffic. Finally, the joint time and space modeling is required to obtain an even more realistic traffic model.

## 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] H. Heffes and D. Lucantoni, "A Markov modulated characterization of packetized voice and data traffic and related statistical multiplexer performance," *IEEE Journal on Selected Areas in Communications*, vol. 4, no. 6, pp. 856–868, September 1986.

[4] A. Klemm, C. Lindemann, and M. Lohmann, "Modeling IP traffic using the batch Markovian arrival process," *Performance Evaluation*, vol. 54, no. 2, pp. 149–173, October 2003.

[5] A. Dainotti, A. Pescapé, P. S. Rossi, F. Palmieri, and G. Ventre, "Internet traffic modeling by means of hidden Markov models," *Computer Networks*, vol. 52, no. 14, pp. 2645–2662, October 2008.

[6] Y. Xie, J. Hu, Y. Xiang, S. Yu, S. Tang, and Y. Wang, "Modeling oscillation behavior of network traffic by nested hidden Markov model with variable state-duration," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 9, pp. 1807–1817, September 2013.

[7] A. T. Andersen and B. F. Nielsen, "A Markovian approach for modeling packet traffic with long-range dependence," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 5, pp. 719–732, 1998.

[8] G. Casale, E. Z. Zhang, and E. Smimi, *Interarrival Times Characterization and Fitting for Markovian Traffic Analysis*. Internat. Begegnungs- und Forschungszentrum für Informatik, 2008.

[9] ITU-R., "ITU-R M.2135: Guidelines for Evaluation of Radio Interface Technologies for IMT-Advanced," ITU, Tech. Rep., 2008.

[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] 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.

[13] M. Taranetz 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.

[14] 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.

[15] A. Leon-Garcia, *Probability and Random Processes for Electrical Engineering*. Reading, Massachusetts: Addison-Wesley Publishing Company, Inc, 1994.

[16] B. Błaszczyszyn, D. Yegorshwaran *et al.*, "On comparison of clustering properties of point processes," *Advances in Applied Probability*, vol. 46, no. 1, pp. 1–20, 2014.

[17] R. Schoenen, R. Halfmann, and B. Walke, "Mac performance of a 3GPP-LTE multihop cellular network," in *Proc. IEEE International Conference on Communications (ICC)*, Beijing, May 2008.

[18] R. Schoenen and H. Yanikomeroglu, "User-in-the-loop: Spatial and temporal demand shaping for sustainable wireless networks," *IEEE Communications Magazine*, vol. 52, pp. 196–203, February 2014.