Unified and Non-parameterized Statistical Modeling of Temporal and Spatial Traffic Heterogeneity in Wireless Cellular Networks

Meisam Mirahsan, Ziyang Wang, Rainer Schoenen, Halim Yanikomeroglu, Marc St-Hilaire

Abstract—**Understanding and solving performance-related issues of current and future (5G+) networks requires the availability of realistic, yet simple and manageable, traffic models which capture and regenerate various properties of real traffic with sufficient accuracy and minimum number of parameters. Traffic in wireless cellular networks must be modeled in the space domain as well as the time domain. Modeling traffic in the time domain has been investigated well. However, for modeling the User Equipment (UE) distribution in the space domain, either the unrealistic uniform Poisson model, or some non-adjustable model, or specifc data from operators, is commonly used. In this paper, stochastic geometry is used to explain the similarities of traffic modeling in the time domain and the space domain. It is shown that traffic modeling in the time domain is a special one-dimensional case of traffic modeling in the space domain. Unified and non-parameterized metrics for characterizing the heterogeneity of traffic in the time domain and the space domain are proposed and their equivalence to the inter-arrival time, a well accepted metric in the time domain, is demonstrated. Coefficient of Variation (CoV), the normalized second-order statistic, is suggested as an appropriate statistical property of traffic to be measured. Simulation results show that the proposed metrics capture the properties of traffic more accurately than the existing metrics. Finally, the performance of LTE networks under modeled traffic using the new metrics is illustrated.**

Index Terms—**Traffic Modeling, Stochastic Geometry, Point Process, Voronoi Tessellation, Wireless Cellular Network.**

I. Introduction

The statistics of Signal-to-Interference-plus-Noise Ratio (SINR) are the key to the performance of wireless cellular networks. The signal strengths and interference depend strongly on the network geometry, i.e., the relative positions of the transmitters and the receivers. So, in wireless cellular networks, spatial properties of traffic as well as temporal properties of traffic have direct effects on network performance. Modeling traffic in the time domain has been investigated well in the literature [1–3]. In the space domain, on the other hand, the unrealistic uniform Poisson modeling based on IMT-Advanced evaluation guidelines [4], non-adjustable models, or traced data from specific service providers are commonly used.

M. Mirahsan, Z. Wang, R. Schoenen, H. Yanikomeroglu and M. St-Hilaire are with the Department of Systems and Computer Engineering, Carleton University, Ottawa, ON, Canada (e-mail: {mirahsan, wangzi, rs, halim}@sce.carleton.ca, marc_st_hilaire@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.

Real UE distributions appear due to various reasons, but they are never pure Poisson point processes. Studying UE distributions of more extreme characteristics and their impact on performance is thus an important issue. Heterogeneous scenarios in space have been studied recently [5], but mainly focusing on the Base Station (BS) or small cell deployment, not on the UE statistics. The requirement is a continuously scalable model from zero heterogeneity (lattice) to extreme cases (e.g., clustering).

Unlike existing models which require many parameters to specify and thus lead to scientific work being incomparable, in this paper we come up with just two parameters to sufficiently describe a heterogeneous scenario. First, the user density (homogeneous component). Second, the CoV of inter-point distance metrics for specifying the deviations from homogeneity. This turns out to be the equivalent of the CoV in temporal traffic, known from traffic and queueing theory.

With these first and second order statistics, we are able to specify scenarios which are then analyzed in a wireless context. In order to get there, we propose a mapping and normalization procedure, which allows existing methods for point process generation to be used, but hide their complexity so that only the CoV is required as parameter.

Figure 1 shows the traffic modeling procedure used in this paper. Modeling traffic comprises of two main steps: 1) Generating a traffic pattern, and 2) Capturing and inference of the statistical properties of a given pattern. Different traffic generators receive different input parameters. So, a translation of the Traffic Generator Input Parameters (TGIP) to the desired traffic properties is essential. For statistical inference of traffic, two questions should be answered: a) what is the right metric to analyze? In the time domain, inter-arrival time is a well accepted and commonly used metric. In the space domain, on the other hand, the equivalent of this metric doesn't exist in the literature. In this paper, equivalents of inter-arrival time in the space domain are introduced. b) Which statistics of the chosen metric are to be characterized? First-order statistics like the mean, and infinite-order statistics like probability density function can be considered. We suggest CoV (C), the normalized second-order statistic, defined as the ratio of standard deviation to mean, which is easy to calculate and at the same time appears to be sufficient to capture the main characteristics (heterogeneity) of the traffic.

The main contributions of this paper are as follows: 1) Stochastic geometry as a common tool is used to explain

978-1-4799-4640-2/14/$31.00 ©2014 Crown

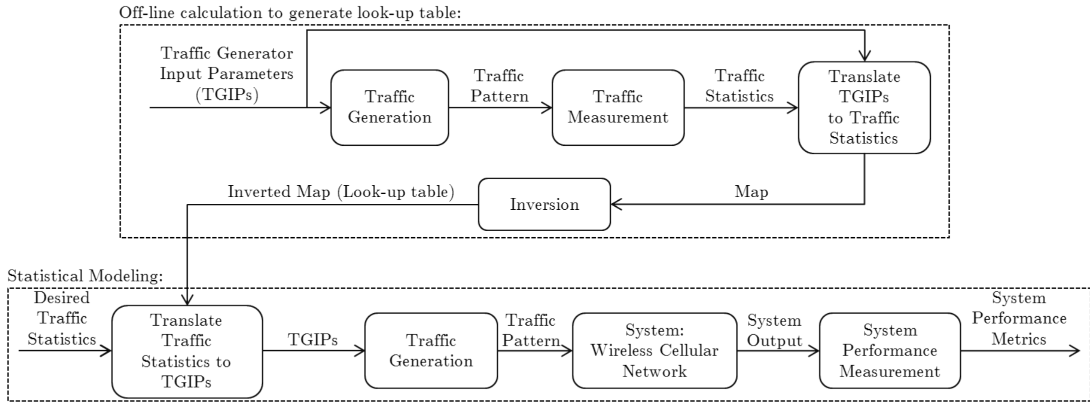

Fig. 1. Traffic modeling comprises of two main steps: 1) generating a traffic pattern, and 2) capturing the statistical properties of a given pattern. Different traffic generators receive different input parameters. So, a translation of the TGIP to the desired traffic properties is essential.

the similarities of traffic modeling in the time domain and the space domain. 2) Unified, non-parameterized and accurate metrics and models for capturing statistical properties of traffic in the space and time are proposed. 3) A continuously scalable traffic model from zero heterogeneity to extreme clustering is proposed.

In Section II, existing models for modeling traffic in the time domain and the space domain are investigated. In Section III, the proposed metrics and modeling for capturing statistical properties of traffic in time and space are introduced. Section IV presents the simulation results and Section V concludes the paper.

II. Related Work

Packet arrivals in time domain can be modeled by a one-dimensional (1D) point process. A fixed inter-arrival time (iat) between packets generates maximum homogeneity (lattice). Exponentially distributed iat generates complete randomness (Poisson). For generating sub-Poisson patterns (patterns with more homogeneity than Poisson) one way is to generate a perfect lattice and apply a random displacement (perturbation) on its points [6, 7]. Various models for generating super-Poisson patterns (patterns with more heterogeneity than Poisson) have been proposed in the literature which are mostly based on hierarchical randomness and Markov models [1–3].

A 1D point process in time domain can be measured mathematically in many different ways. One may use the interval counts $N(a, b] = N_b - N_a$ which is a density-based metric and divide the whole domain into smaller windows and count the number of process points in each window. A disadvantage of density-based metrics is that they are parameterized by the window size. Finding an appropriate window size is itself a challenging question and cannot be answered generally for all applications. Inter-arrival time $I_i = T_{i+1} - T_i$ is the most popular and best-accepted metric because it is distance-based rather than density-based and considers the distance between every two neighboring points in domain. Considering CoV, for 1D-lattice, the constant iat has $C_I = 0$. For a 1D-Poisson pattern, $C_I = 1$ since for an exponential distribution with parameter λ the standard deviation and the mean are both $\mu_I = \sigma_I = \lambda$. Sub-Poisson processes have $0 < C_I < 1$ and super-Poisson processes have $C_I > 1$.

UE locations in a wireless cellular network in space domain can be modeled by a two-dimensional (2D) or three-dimensional (3D) point process. A very inclusive review of Point processes in space domain is conducted in [8]. Fixed distance between points generates perfect homogeneity (lattice). Poissonian distribution generates complete randomness. For generating sub-Poisson patterns, one way is to generate a perfect lattice and apply a random perturbation on its points [6, 7]. For generating super-Poisson patterns, hierarchical randomness based on doubly stochastic clustering perturbation can be used. Clustering perturbation of a given (parent) process Φ consists of independent replication and displacement of points of Φ, with the number of replications of a given point $x \in \Phi$ having distribution $\Gamma(x)$ and the replicas' locations having distribution $\chi(x)$. All replicas of x form a cluster. A survey of super-Poisson processes in space domain can be found in [8].

As mentioned above, in time domain, distance-based metric inter-arrival time captures heterogeneity by one non-parameterized real value C_I. In multi dimensions, however, there is no natural ordering of the points, so finding the analogue of the inter-arrival time is not easy. There are many density-based heterogeneity metrics in the literature like Ripley's K-function and pair correlation function [8] but they are all parameterized. For introducing distance-based metrics, the problem is about defining the 'next point' or the 'neighboring points' in multi-dimensional domains. The first and simplest candidate for characterizing neighboring point in multi-dimensional domain is the nearest-neighbor. This leads to nearest-neighbor distance metric [9]. However, the nearest-neighbor distance metric in 1D time domain is not the analogue of the inter-arrival time because it is considering the $\min(I_i, I_{i+1})$ for every point T_i. It is shown in our simulation results that nearest neighbor fails to capture process statistics in multi-dimensional domains because it only considers the nearest neighbor and ignores the other neighbors. The next candidate is the distance to k_{th} neighbor. However, determining k globally is not possible because every point may have different number of neighbors.

III. Unified and Simplified Traffic Metrics

Given a point pattern $P = \{p_1, p_2, ..., p_n\}$ in d-dimensional space $\mathbb{R}^d$, the Voronoi tessellation $T = \{c_{p_1}, c_{p_2}, ..., c_{p_n}\}$ 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

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

The Voronoi tessellation in $\mathbb{R}^d$ has the property that each of its vertices is given by the intersection of exactly $d+1$ Voronoi cells. The corresponding $d+1$ points define a Delaunay cell. So the two tessellations are said to be dual. Figure 2 demonstrates a pattern of points with its Voronoi tessellation (dashed lines) and Delaunay tessellations (solid lines).

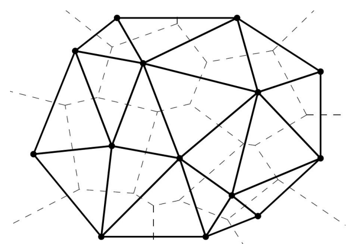

Fig. 2. Voronoi (dashed lines) and Delaunay (solid lines) tessellations.

Every two points sharing a common edge in Voronoi tessellation or equivalently every two connected points in Delaunay tessellation of a point process are called 'natural neighbors'. This gives an inspiration of neighboring relation in multi-dimensional domains and leads us to analogues of the well accepted inter-arrival time metric in multi dimensions. Various statistical inferences based on different properties of cells generated by these tessellations can be considered for measurement of a point pattern.

'Voronoi Cell Area' or 'Voronoi Cell Volume' V is the first natural choice. For a lattice process, all the cell areas in 2D or cell volumes in 3D are equal and $C_V = 0$. The statistics of the Voronoi cells for a Poisson point process (Poisson-Voronoi Tessellation) are well investigated in the literature [10–14]. Square rooted Voronoi cell area in 2D or cube rooted Voronoi cell volume in 3D also can be considered.

The next proposed metric is the Delaunay edge length E. The statistics of Delaunay tessellations is investigated in [15–17]. The mean value of the lengths of Delaunay edges of every point M can also be considered.

A Delaunay tessellation divides the space to triangles or tetrahedrons in 2D and 3D, respectively. The area distribution of the triangles or the volume distribution of tetrahedrons T can determine the properties of the underlying pattern.

Voronoi and Delaunay tessellations can be applied on a 1D process which models traffic in time domain. In this case, the introduced distance-based metrics are converted to time domain metrics. Basic statistics of these metrics for a Poisson point process in one, two and three dimensions and their analogues in time domain are summarized in Table I.

In order to use the above mentioned metrics as an analogue of inter-arrival time, one needs to normalize their CoV to the CoV values of inter-arrival time in the time domain. For complete homogeneity case, the CoV values are already zero like inter-arrival time. To normalize the CoV values of complete random case to 1, it is required to divide the metrics by the values presented in Table I. Figure 3 demonstrates realizations of processes with sub-Poisson, Poisson and super-Poisson characteristics.

IV. Simulation Results

In a two dimensional 1000m $\times$ 1000m square field 1000 points are distributed. For every configuration, the simulations are repeated for 1000 ensemble drops. Simulation results for sub-Poisson and super-Poisson processes are presented in the following sub-sections and the final sub-section illustrates the performance of wireless cellular networks under modeled traffic using the new proposed metrics.

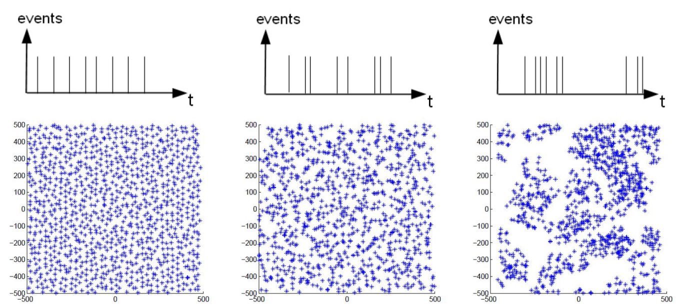

Fig. 3. Realizations of processes with sub-Poisson ($0 < C < 1$), Poisson ($C = 1$) and super-Poisson ($C > 1$) characteristics respectively from left to right in time domain (top) and space domain (bottom).

A. Sub-Poisson Processes

To generate a sub-Poisson process, a hexagonal lattice is generated and then a symmetric Gaussian perturbation is applied on its points with direction uniformly distributed in $[0, 2\pi]$. The perturbation distance is $l \sim Norm(0, \alpha L)$ where L is the original distance between every two neighboring points in the lattice. Figure 4 shows the CoV of the discussed distance-based metrics. To cancel the field edge effects, the edge Voronoi and Delaunay cells are cut at the border. With increase in the perturbation distance, the resulting process converges to the Poisson process [6, 7]. Various methods could be used to generate the primary lattice like square or hexagonal lattice. Also various methods could be used for perturbation like uniform or Gaussian distance from original location. It is shown in [7] that the difference is only in the convergence behavior. The CoV of all the metrics is normalized to be 1 at the Poisson end (right edge) by dividing them by their convergence value. Figure 5 shows the normalized results. The standard deviation of any random variable is a good indicator

TABLE I
Basic statistics of distance-based metrics for a Poisson point process in one, two and three dimensions and their analogues in time domain: i is the process point index, λ is the exponential distribution parameter for inter-arrival time and Λ is the mean intensity of point processes.

Distance-based metrics	Time domain analogue	Statistics	1D	2D	3D
Nearest-neighbor distance (G)	$\min\{I_i, I_{i+1}\}$	Mean (μ)	$0.5\lambda^{-1}$	$0.5\Lambda^{-0.5}$	$0.5539\Lambda^{-0.33}$
		Variance (σ^2)	$0.25\lambda^{-2}$	$0.0683\Lambda^{-1}$	$0.04\Lambda^{-0.66}$
		CoV (C)	1	0.653	0.364
Voronoi cell area/volume (V)	$\frac{I_i+I_{i+1}}{2}$	Mean (μ)	λ^{-1}	Λ^{-1}	Λ^{-1}
		Variance (σ^2)	λ^{-2}	$0.28\Lambda^{-2}$	$0.18\Lambda^{-2}$
		CoV (C)	1	0.529	0.424
Dealunay cell area/volume (T)	I_i	Mean (μ)	λ^{-1}	$0.5\Lambda^{-1}$	$0.147\Lambda^{-0.5}$
		Variance (σ^2)	λ^{-2}	$0.443\Lambda^{-2}$	$0.015\Lambda^{-1}$
		CoV (C)	1	0.879	0.833
Dealunay cell edge length (E)	I_i	Mean (μ)	λ^{-1}	$1.131\Lambda^{-0.5}$	$1.237\Lambda^{-0.33}$
		Variance (σ^2)	λ^{-2}	$0.31\Lambda^{-1}$	$0.185\Lambda^{-0.66}$
		CoV (C)	1	0.492	0.347

of its inaccuracy and can be divided by its mean value to be normalized. To compare the inaccuracy of the CoV of the metrics, the ensemble CoV (for 1000 drops) of the CoV of all metrics is depicted in Fig. 6.

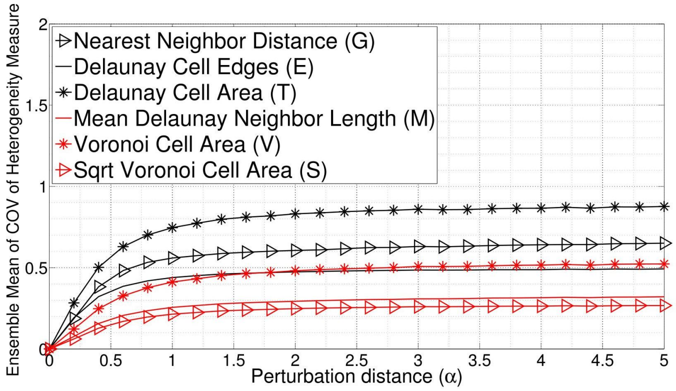

Fig. 4. Sub-Poisson: Ensemble mean of the CoV of the metrics.

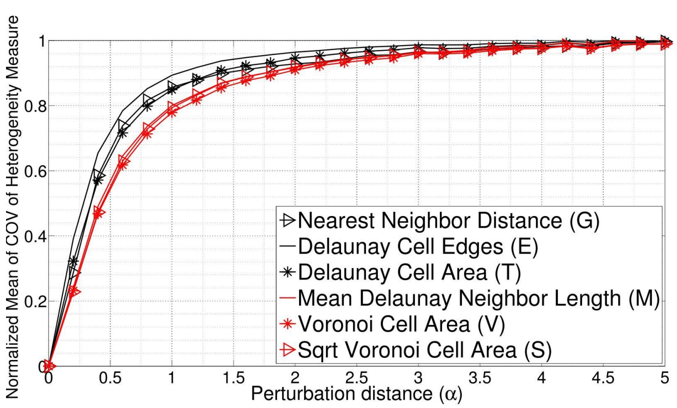

Fig. 5. Sub-Poisson: Normalized mean of the CoV of the metrics.

All the proposed unified metrics are generally more accurate than the existing 'nearest-neighbor' metric. Among the

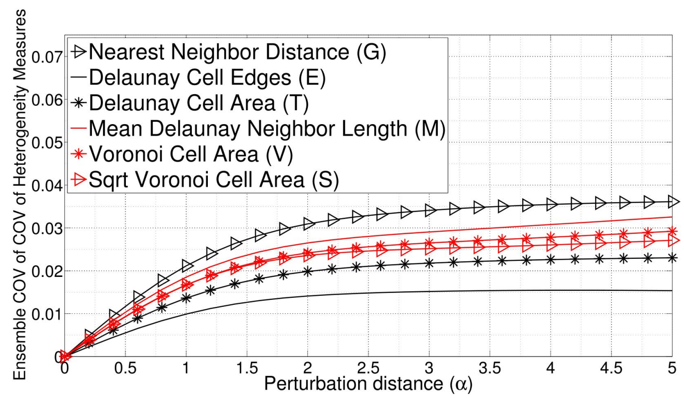

Fig. 6. Sub-Poisson: Ensemble CoV of the CoV of the metrics which measures the accuracy.

proposed metrics, 'Delaunay cell edge' shows the highest accuracy. 'Voronoi cell area' has the lowest slope in the normalized curves and converges last. This means that this metric can be used to generate a wide range of heterogeneity.

B. Super-Poisson Processes

For generating a super-Poisson pattern, various processes can be used. Thomas process [8] is selected for this paper. In Thomas process, first, a number of cluster-heads are distributed Poissonian in the space. Then users are distributed around cluster-heads. With fixed number of points (fixed Λ), parameter β, the number of clusters, is changed to generate patterns with different heterogeneities. The points associated with every cluster are distributed using symmetric Gaussian distributed distance $N(0, \beta)$ from cluster-head and uniformly distributed direction in $[0, 2\pi]$ in the spheres centered at the cluster-heads. With increase in the number of clusters, the process finally converges to Poisson process because every point will be a cluster-head and cluster-heads are distributed by Poissonian distribution.

Figure 7 demonstrates the CoV of the discussed distance-

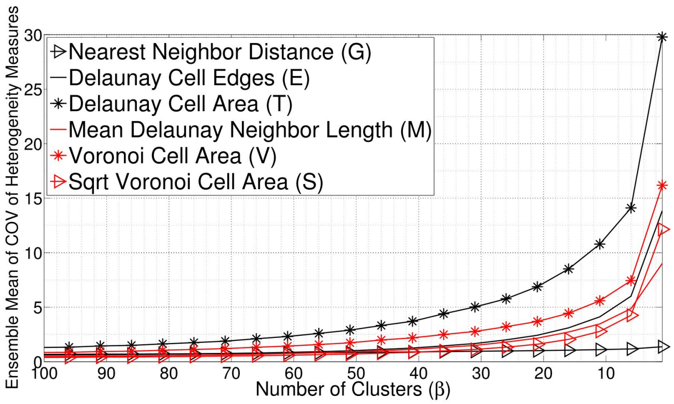

Fig. 7. Super-Poisson: Ensemble mean of the CoV of the metrics.

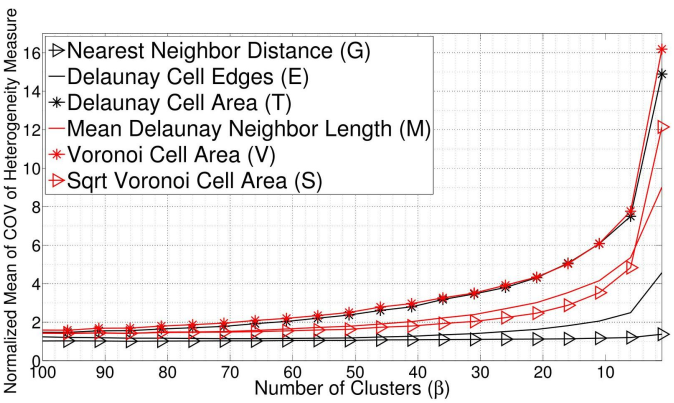

Fig. 8. Super-Poisson: Normalized mean of the CoV of the metrics.

based metrics. The CoV of all the metrics are normalized to be 1 at the Poisson end (left edge). Figure 8 shows the normalized results. The ensemble CoV (for 1000 drops) of the CoV of all metrics is demonstrated in Fig. 9.

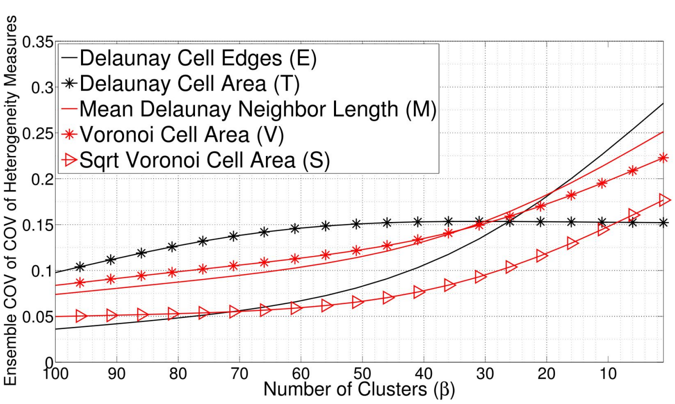

Fig. 9. Super-Poisson: Ensemble CoV of the CoV of the metrics.

For super-Poisson processes, the nearest-neighbor distance metric fails to capture the properties of underlying point process as it remains almost the same with decreasing the number of clusters. Among all the proposed metrics, 'Voronoi cell area' has the lowest slope in the normalized curves and converges last. This means that this metric can be used to generate a wide range of heterogeneity. All the metrics already almost converge to Poisson process value at $\beta = 100$ (but for sure at $\beta = 1000$). Figure 10 shows the measured heterogeneity (model output) versus the desired heterogeneity (model input).

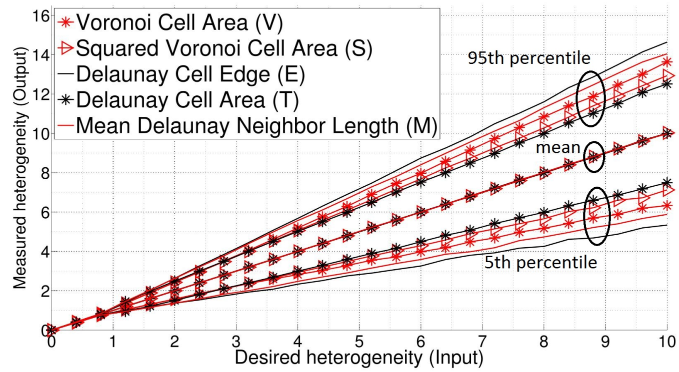

Fig. 10. Measured heterogeneity (output) vs. desired heterogeneity (input): Upper, middle and lower lines show 95th percentile, mean and 5th percentile respectively for each metric.

C. Wireless Cellular Network Performance Analysis

The traffic model has a direct effect on the performance of wireless cellular networks. The first-order statistic of the traffic is its mean value or intensity Λ. With increase in the number of users in the network or increase in the number of packets (or bits) generated by each user, the network resources are shared for more traffic demand and the rate achieved by each user (hence the average user rate) is decreased. The second-order statistic of traffic is the standard deviation. To demonstrate the effect of traffic standard deviation on the network performance, the traffic intensity must be fixed. CoV is the perfect statistic to do this because it normalizes the standard deviation by the mean. So, it cancels the effect of the mean value. To illustrate the performance of a wireless cellular network under modeled heterogeneous traffic, the CoV of the Voronoi cell area is chosen as the heterogeneity metric and is increased from 0 to 15. Simulation is done based on IMT-Advanced and LTE [18]. Table II shows the simulation parameters and Fig. 11 demonstrates the network performance metrics under various traffic heterogeneity levels. Figure 12 shows some statistics of user rates under modeled traffic.

With increase in the traffic heterogeneity, the mean user rate, the median user rate and coverage probability for the entire network are decreased monotonically. This is because in heterogeneous user distributions, some BSs serve more users than they are planned for and are under high pressure while some other BSs serve less users than they are planned for and their capacity is wasted.

V. Conclusion

This paper used stochastic geometry as a common modeling tool to explain the similarities of traffic modeling in

TABLE II
Simulation parameters based on IMT-advanced.

Parameter	Value
Number of Sectors	57 (wrap-around)
Cellular layout	3-sectorized hexagonal grid with 19 serving BSs
ISD	500 m
Scenario	ITU UMa
Carrier frequency	2 GHz
Bandwidth	20 MHz
Shadowing	Log-normal, std.= 4 for LOS, std.= 6 for NLOS
Average UE density	25 UEs/sector
UE speed	30 km/h
Total BS Tx power	49 dBm
BS antenna height	25 m
Antenna number	SISO
BS antenna gain	17 dBi
UE antenna gain	0
Traffic model	Full buffer
Number of drops	1000
Scheduling algorithm	Proportional Fair

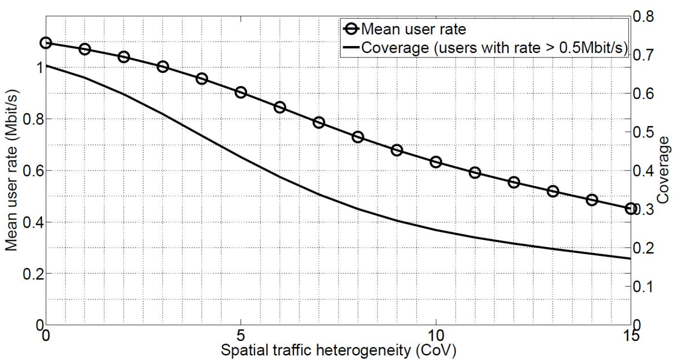

Fig. 11. Performance metrics (mean user rate [Mbit/s] and coverage) of the LTE network versus traffic heterogeneity level (CoV).

the time domain and the space domain. Unified and non-parameterized statistical traffic metrics based on Voronoi and Delaunay tessellations were proposed and their equivalence to the inter-arrival time was shown. CoV, the normalized second-order statistics, was suggested to be used to capture the main statistical properties of traffic. Results for the LTE network performance analysis show one important application case of the proposed traffic modeling.

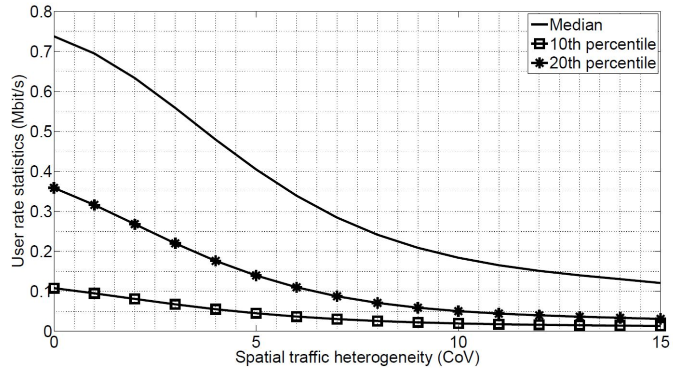

Fig. 12. Statistics of user rates of the LTE network versus traffic heterogeneity level (CoV).

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] 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.

[3] 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*, p. 1, September 2012.

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

[5] 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 Communications Surveys & Tutorials*, 2013.

[6] J. Rataj, I. Saxl, and K. Pelikán, "Convergence of randomly oscillating point patterns to the Poisson point process," *Applications of Mathematics*, vol. 38, no. 3, pp. 221–235, 1993.

[7] V. Lucarini, "From symmetry breaking to Poisson point process in 2d Voronoi tessellations: the generic nature of hexagons," *Journal of Statistical Physics*, vol. 130, no. 6, pp. 1047–1062, January 2008.

[8] B. Blaszczyszyn and D. Yogeshwaran, "Clustering comparison of point processes with applications to random geometric models," *arXiv:1212.5285*, 2012.

[9] P. J. Clark and F. C. Evans, "Distance to nearest neighbor as a measure of spatial relationships in populations," *Ecology*, vol. 35, no. 4, pp. 445–453, 1954.

[10] C. D. Barr, *Applications of Voronoi Tessellations in Point Pattern Analysis*. ProQuest, 2008.

[11] J. Møller and R. P. Waagepetersen, "Modern statistics for spatial point processes," *Scandinavian Journal of Statistics*, vol. 34, no. 4, pp. 643–684, September 2007.

[12] E. Gilbert, "Random subdivisions of space into crystals," *The Annals of Mathematical Statistics*, vol. 33, no. 3, pp. 958–972, 1962.

[13] K. Borovkov and D. Odell, "Simulation studies of some Voronoi point processes," *Acta Applicandae Mathematicae*, vol. 96, no. 1-3, pp. 87–97, 2007.

[14] M. Tanemura, "Statistical distributions of Poisson Voronoi cells in two and three dimensions," *FORMA-TOKYO-*, vol. 18, no. 4, pp. 221–247, November 2003.

[15] L. Muche, "Distributional properties of the three-dimensional Poisson Delaunay cell," *Journal of Statistical Physics*, vol. 84, no. 1-2, pp. 147–167, November 1996.

[16] P. Rathie, "On the volume distribution of the typical Poisson-Delaunay cell," *Journal of Applied Probability*, pp. 740–744, September 1992.

[17] R. Miles, "The random division of space," *Advances in Applied Probability*, pp. 243–266, 1972.

[18] R. Schoenen, R. Halfmann, and B. Walke, "MAC performance of a 3GPP-LTE Multihop Cellular Network," in *Proceedings of the IEEE ICC'08*, Beijing, May 2008.