

Pairwise Interaction Processes for Modeling Cellular Network Topology

David B. Taylor, Harpreet S. Dhillon, Thomas D. Novlan, Jeffrey G. Andrews

Abstract—In industry, cellular tower locations have primarily been modeled by a deterministic hexagonal grid. Since real deployments are rarely regular, the even spacing between nodes in the grid and constant *Voronoi cell* areas make the hexagonal grid unrealistic. In this paper we use tools from spatial statistics to show that a purely random node placement and a hexagonal grid distribution with the points perturbed also have unrealistic spatial relationships between nodes, and that pairwise interactions between nodes are necessary, and in most cases sufficient, for modeling spatial qualities of cellular networks. We detail the benefits of using pairwise point interactions in modeling both a coverage-centric tower deployment and a capacity-centric tower deployment. We propose using pairwise and saturated pairwise interaction point processes from the *Gibbs process* family of point processes: the *Strauss Hardcore process* for inhibitive point patterns and the *Geyer Saturation process* for clustered point patterns. Due to its relationship with the coverage areas, we also propose that the *Voronoi cell* area distribution can be used as a test statistic in general spatial modeling of cellular networks.

I. INTRODUCTION

Driven by exponential increase in network traffic and expectations of ubiquitous access, modern cellular networks continue to evolve from highly regular voice/coverage-oriented deployments to spatially dense data-driven deployments with multiple tiers of access points [1].

Evaluating the performance of wireless networks is essential to understanding system-wide trends and furthering research and development of new algorithms, which requires capturing the complexity of real deployments as accurately as possible. One of the most significant technical challenges in modeling cellular networks arises from the interdependence of base station locations among themselves as well as with the user locations. The resulting distribution of the interference is sensitive to this dependence and has a critical impact on performance in modern networks which are primarily limited by inter-cell interference.

A. Related Work

By far the most popular approach in academia and industry has been to model base station locations according to a deterministic hexagonal grid model [2]. This approach has been successful in creating well-defined simulation scenarios that allow calibration across organizations and is used by standards bodies like the 3rd Generation Partnership Project (3GPP).

D. B. Taylor, H. S. Dhillon, T. D. Novlan and J. G. Andrews are with WNCG, the University of Texas at Austin (email: david.taylor@utexas.edu, dhillon@utexas.edu, tdnovlan@utexas.edu, jandrews@ece.utexas.edu). This research has been supported by Crown Castle International and by AT&T Laboratories.

However, the network topology of real deployments is rarely regular due to regulatory factors restricting site acquisition and placement, topological factors which cause the propagation environment to vary significantly over a deployment area, and economic and demographic factors such as unequal demand for capacity in urban versus rural regions.

A regular grid approach can be scaled for larger or smaller cell sizes by varying the inter-site distance to adjust to different simulation scenarios, but this does not change the homogeneity of the deployment. One approach to introduce non-uniformity is to maintain the overall grid structure but induce a random perturbation of the base station locations [3]. Additionally, a novel non-regular but deterministic model called the Springwald layout was proposed in [4]. The layout is based on Archimedes's spiral, making it simple to implement in simulations while providing scalability similar to the hexagonal grid, but with varying inter-site distances. The choice of these parameters is however heuristic and it is not clear whether they capture the interdependence of deployments in their clustering or dispersion. Also system evaluations have been performed using actual base station locations [5]. However, the use of real locations is not a scalable approach as networks are constantly evolving and this approach only gives insights into a specific deployment.

A fundamentally alternative approach is the use of random spatial models, i.e., the *Poisson point process* (PPP) [6]. This new approach is advantageous both in capturing the non-uniformity of deployments, especially applicable in the context of heterogeneous networks, and in its ability to derive tractable expressions for key performance metrics of coverage and average rate as a function of system parameters instead of relying on detailed and complex system simulations [7]–[10]. Although it is shown to be about as accurate as the grid model when compared to an actual cellular network deployment in [8], one drawback of either approach is that they do not take into account inter-point interactions arising from the site-planning process conducted by the operator. In order to model performance as accurately as possible, it is important to first understand the exact nature of the spatial processes that describe these deployments. Therefore, the main goal of this paper is to extract useful features from actual location data that can be used to develop more realistic models for general macrocellular deployments.

B. Contributions

The main contributions of this paper are as follows:
Realistic Spatial Models for Cellular Networks. Since macrocellular deployments are in general neither too regular

nor too clustered, it provides an ideal setting to use *Gibbs models* [11]. As a main contribution, we show that pairwise interaction is in general necessary and in most cases sufficient to model certain real world deployments. Furthermore, we show that simple pairwise and saturated pairwise interaction models, such as *Strauss process* for inhibitive deployments and its generalization *Geyer saturation process* for clustered deployment, are enough to model wide variety of deployment scenarios. Recently the *Geyer Saturation process* model was also used to accurately model the spatial characteristics of an outdoor Wi-Fi deployment maintained by Google in Mountain View, California [12].

Voronoi Cell Area Distribution as a Possible Metric. Second, we propose to use Voronoi cell area distribution as a possible metric for spatial modeling of cellular networks due to its relationship with the base station coverage regions. For example, an irregular deployment is expected to have higher variance in the cell areas than the regular one. Using this proposed metric, we show that the state-of-the-art grid model along with its popular variants involving cell site perturbations do not accurately capture spatial dependence and hence is not the most accurate model, even for simulation studies. As discussed in the sequel, other popular summary statistics, such as signal-to-interference-ratio (SIR) distribution and *Ripley's K-function* are unable to reach this conclusion.

II. BACKGROUND – GIBBS MODELS

Gibbs models offer a remarkably general, flexible and powerful family of distributions for modeling point processes [11]. Since they are constructed from the probability densities, they are particularly useful in modeling the inter-point interactions in point patterns. To highlight this point, we consider a point pattern $\mathbf{z} = \{z_1, z_2, \dots, z_{n(\mathbf{z})}\}$ confined in a bounded window \mathcal{W} , where $n(\mathbf{z}) = |\mathbf{z}|$. Since we are assuming \mathbf{z} to be a random point pattern, $n(\mathbf{z})$ is not fixed and may as well be 0. Although *Gibbs spatial processes* can be used to model general inter-point interaction, we focus our attention on a simple “pairwise interaction model” in this paper, which is sufficient for macrocellular modeling. For this model, the probability density $f(\mathbf{z})$ can be expressed in the product form as follows:

$$f(\mathbf{z}) = \alpha \left[\prod_{i=1}^{n(\mathbf{z})} \phi(z_i) \right] \left[\prod_{i<j} \psi(z_i, z_j) \right], \quad (1)$$

where α is the normalizing factor to ensure the probability density integrates to unity, $\phi(z)$ is the function modeling the “first order trends” and $\psi(z_i, z_j)$ are the functions modeling the pair-wise interaction.

Now we introduce two special cases for our study: the *Strauss process* [13], which is useful in modeling inhibition, and its generalization the *Geyer saturation process* [14], which models both inhibition and clustering. The *Strauss process* is defined by taking $\phi(z_i) = \beta$ and defining $\psi(z_i, z_j)$ as follows:

$$\psi(z_i, z_j) = \begin{cases} 1, & \|z_i - z_j\| > r \\ \gamma, & \|z_i - z_j\| \leq r \end{cases} \quad (2)$$

The probability density function is

$$f(\mathbf{z}) = \alpha \beta^{n(\mathbf{z})} \gamma^{s(\mathbf{z})}, \quad (3)$$

where $s(\mathbf{z})$ is the number of distinct pairs of points less than an interaction radius r units apart. Clearly $\gamma \leq 1$ models inhibition. However, this probability density is not integrable for $\gamma > 1$ and hence cannot be used to model clustering [15]. Nevertheless, this can be overcome by including a saturation limit in the exponent of γ to obtain the following density function:

$$f(\mathbf{z}) = \alpha \beta^{n(\mathbf{z})} \gamma^{\min(s(\mathbf{z}), t)}. \quad (4)$$

This is termed as the *Geyer saturation process*. It reduces to a PPP for $t = 0$ and a *Strauss process* for $t \rightarrow \infty$.

Gibbs models are simulated using *Markov Chain Monte Carlo* (MCMC) algorithms [16]. The main idea behind these algorithms is to define a *Markov chain* whose states are point patterns for the *Gibbs model*. The equilibrium distribution of the chain is designed to be the distribution of the point process that is to be simulated. For this work, we use the *Metropolis-Hastings algorithm* provided in Spatstat, an R package [14].

III. PROPOSED MODELS FOR MACROCELLS

In this paper we investigate the point patterns \mathbf{x} and \mathbf{y} which are subsets of the tower deployments of two of the U.S.’ 10 largest metropolitan areas. Point pattern \mathbf{x} is from a sprawling, landlocked city, while \mathbf{y} is from a coastal area. Since \mathbf{x} is an inhibitive point pattern and \mathbf{y} is a clustered point pattern, taken together they cover a wide range of tower deployments. Both point patterns have 165 nodes and have been mapped to a unit square as shown in Figs. 1 and 2.

A. Proposed Summary Statistics

For hypothesis testing, we use the following three summary statistics:

SIR Distribution: Assuming equal tower transmit powers, the expression for SIR at a point z in the point pattern \mathbf{z} , connecting to the nearest node y is:

$$\text{SIR}(z, \mathbf{z}) = \frac{h_y d(z, y)^{-\alpha}}{\sum_{x \in \mathbf{z} \setminus y} h_x d(z, x)^{-\alpha}}, \quad (5)$$

where we assume $h_y, h_z \sim \exp(1)$ to model *Rayleigh fading* and α denotes the path loss exponent assumed to be 4 to model an outdoor scenario. The SIR distribution plays a crucial role in cellular analysis since its CDF defines the outage probability of a typical mobile in the network [8]–[10].

Voronoi Cell Area Distribution: The *Voronoi cell* of a node $z \in \mathbf{z}$ is defined as $\{y \in \mathbb{R}^2 : d(y, x) > d(y, z), \forall (x \in \mathbf{z} \setminus z)\}$. We propose to use this metric for spatial modeling of cellular networks due to its close relationship with the coverage regions of the base stations. For example, an irregular deployment is expected to have a higher variance in the coverage areas and hence of the Voronoi cell areas than a regular deployment. Additionally, cell sizes play a central role in the study of important performance metrics, such as hand-off rate.

Node Clustering Metric: To differentiate between an inhibitive and clustered point pattern, we define $g(r)$ as:

$$g(r) = \sqrt{E[n(\mathbf{z}, r)]} \quad (6)$$

where $n(\mathbf{z}, r)$ is the number of nodes within a distance r of a randomly chosen node. The function $g(r)$ describes clustering

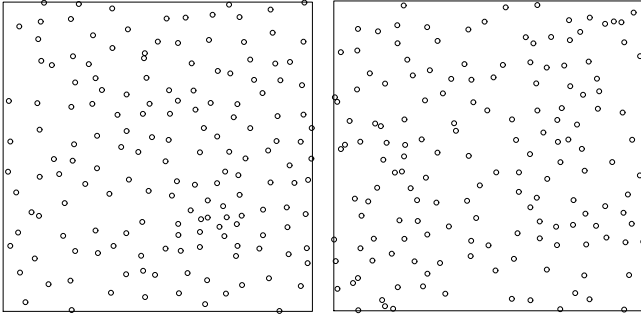


Fig. 1. (left): Point Pattern \mathbf{x} , a tower deployment which exhibits inhibition between nodes. (right): Realization of *Strauss Hardcore process* fitted to \mathbf{x} .

within a radius r from a typical point and has a linear trend. If $K(r)$ is Ripley's K function, $g(r)$ is equivalent to $\sqrt{\lambda K(r)}$. For a *Poisson process*, $g(r) = \sqrt{\pi r^2 \lambda}$. We used $g(r)$ to determine whether a point pattern was clustered and should be modeled with a *Geyer Saturation process*, or inhibitive and should be modeled with a *Strauss Hardcore process*. Clustered point patterns will have values of $g(r)$ greater than a *Poisson process*, while inhibitive point patterns will have values lower. This is true at lower values of r , as $g(r)$ for any homogeneous point process tends toward $\sqrt{\pi r^2 \lambda}$ as r becomes large. This metric is very similar to the *Besag-Ripley L-function* [17]. It should not be confused with the pair correlation function.

Other popular metrics used for studying point patterns that were not used in this paper include empty space distance distributions, nearest neighbor distributions, and the *J-function*, which relates the empty space distance and nearest neighbor distributions [18].

B. Fitted Models and Evaluation Method

Using $g(r)$ we first determine whether a given data pattern is inhibitive or clustered. For example, using this test we can easily see that the point pattern \mathbf{x} is inhibitive whereas \mathbf{y} is clustered. Therefore, we fit a *Strauss hardcore process* with parameters $r = .048$, $h = .0164$, and $\gamma = .1944$ to \mathbf{x} and a *Geyer saturation process* with parameters $r = .01$, $t = 2$, and $\gamma = 2.2255$ to \mathbf{y} . While all other parameters have been defined earlier, h is specific to the hardcore process and provides a minimum allowed distance between two points. We fit these pairwise interaction parameters using the method of maximum profile pseudolikelihood fitting and the method of maximum pseudolikelihood provided in *Spatstat* [19]. We test the hypotheses that \mathbf{x} and \mathbf{y} are realizations of *Strauss hardcore* and *Geyer saturation* processes, respectively, and show that these hypotheses cannot be rejected using the three metrics described earlier. The fact that the *Voronoi cell* area distribution, which is a higher-order property of a point pattern, of these second-order interaction processes closely approximates those of actual data highlights the sufficiency of using second-order interactions.

We also compare \mathbf{x} to a hexagonal grid with points perturbed in a random direction by a distance d , where d is a uniform random variable between 0 and zr with r being the inner radius of the hexagonal cells. This model, which is

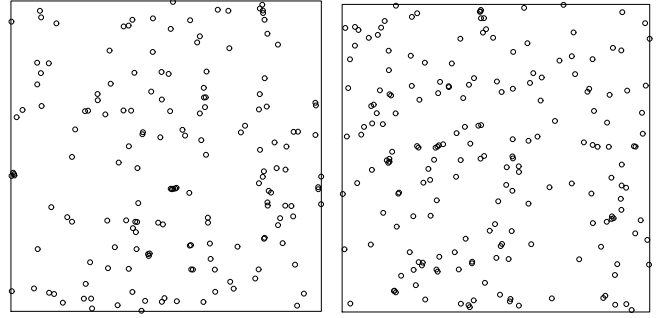


Fig. 2. (left): Point Pattern \mathbf{y} , a tower deployment which exhibits node clustering. (right): Realization of a *Geyer Saturation process* fitted to \mathbf{y} .

widely used in industry, can be no more effective than a PPP at modeling a clustered tower distribution. This is because as z goes infinity, this model becomes a PPP. In our modeling, we tuned z to .65, a value where it had a $g(r)$ similar to the point pattern \mathbf{x} . Even with this model's $g(r)$ tuned to be similar to \mathbf{x} 's, this process' *Voronoi cell* area distribution is much more regular than what we see in the real deployment, highlighting an important limitation of this model.

IV. HYPOTHESIS TESTING OF MODEL VALIDITY

For hypothesis testing, we generate 600 realizations each of a PPP (to test if the data sets exhibit complete spatial randomness), a hexagonal grid with perturbed points, the fitted *Strauss Hardcore process*, and the fitted *Geyer Saturation process*. For each of the three metrics used in this paper, we sample all 600 realizations of each process, and throw out the 30 highest and 30 lowest values to create 90% confidence intervals.

Using SIR: A PPP will rarely generate a node placement that has similar coverage probability to a coverage-centric deployment, which is likely to be more regular. Based on the SIR metric we can reject the hypothesis that \mathbf{x} is *Poisson* as seen in Fig. 3. The SIR of \mathbf{y} lies within the PPP confidence interval, so we cannot reject the hypothesis that \mathbf{y} is *Poisson* based on SIR. Neither can we reject the hypothesis that \mathbf{y} is a hexagonal grid with perturbed points based on this metric as seen in Fig. 4. Not only can we not reject the hypotheses that \mathbf{x} and \mathbf{y} are *Strauss Hardcore* and *Geyer Saturation processes*, respectively, but the SIRs of these processes match those of the actual data almost exactly as seen in Figs. 5 and 6.

TABLE I
VORONOI CELL AREA MEAN AND STANDARD DEVIATION

Process	μ	σ
Poisson	1	.525
Geyer Saturation	1	.611
Strauss Hardcore	1	.341
Perturbed Hex Grid	1	.212
Point Pattern	μ	σ
\mathbf{x}	1	.324
\mathbf{y}	1	.646

Using Voronoi Cell Area Distribution: Based on the cell area distributions, we can reject the hypothesis that \mathbf{x} is *Poisson*, but we cannot reject the hypothesis that \mathbf{y} is *Poisson*

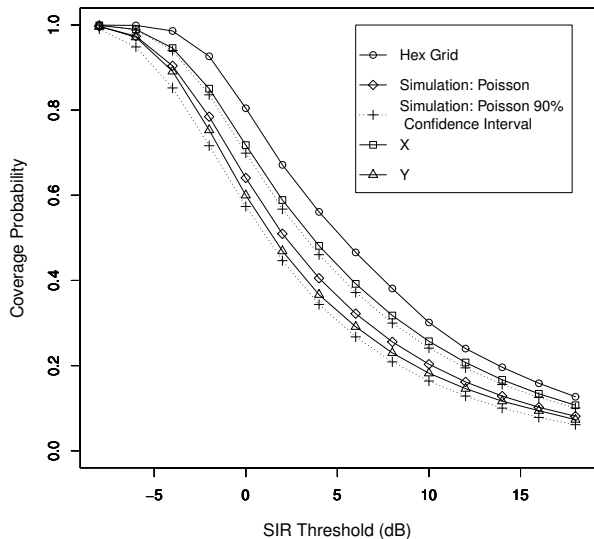


Fig. 3. Rejection of the hypothesis that x is *Poisson* by SIR.

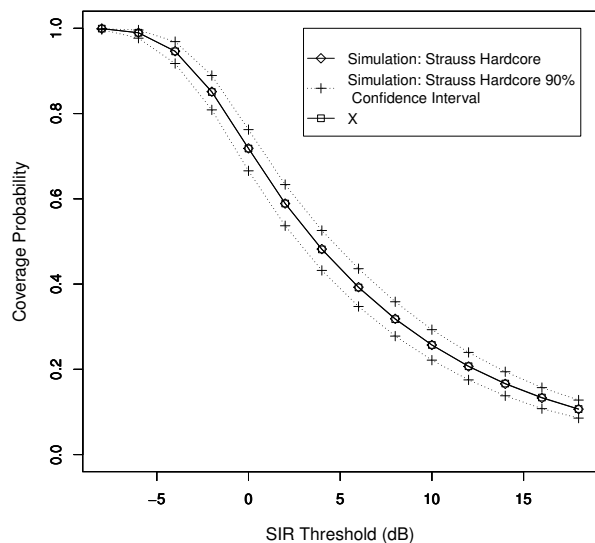


Fig. 5. x 's SIR lies within bounds of a *Strauss Hardcore* process.

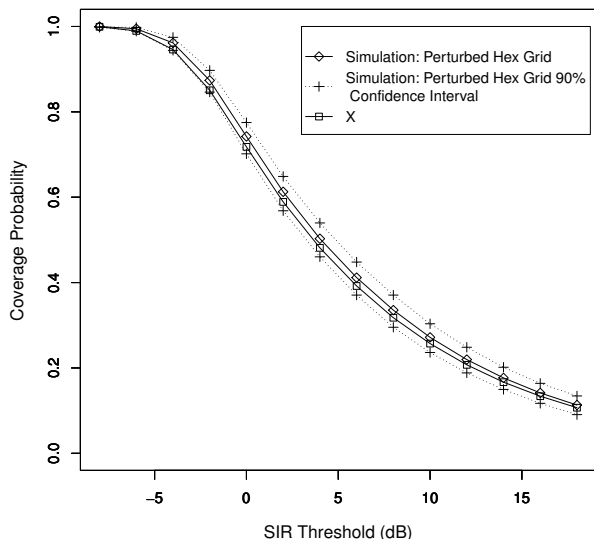


Fig. 4. x 's SIR lies within bounds of a hexagonal grid with perturbed points.

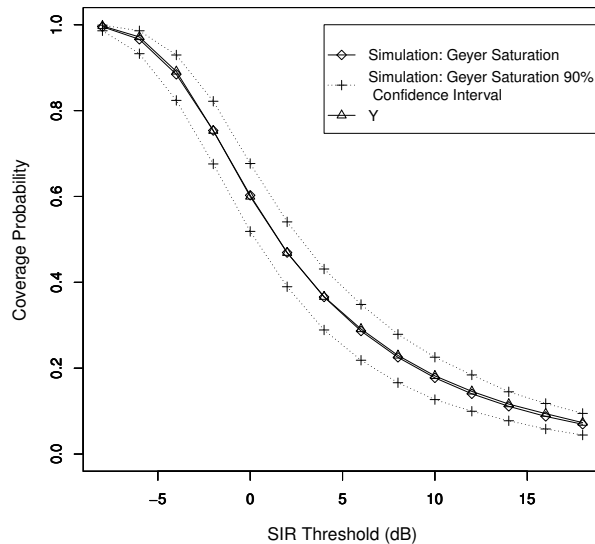


Fig. 6. y 's SIR lies within bounds of a *Geyer Saturation* process.

as seen in Fig. 7. We can reject the hypothesis that x is a hexagonal grid with perturbed points as seen in Fig. 8. This highlights a strength of using the *Voronoi cell* area distribution as a metric for evaluating the accuracy of model compared to actual tower deployments. In the case of the hexagonal grid with perturbed points, the *Voronoi cell* area distribution reveals it has spatial structure different than the real deployment x . We cannot reject the hypothesis that x is a realization of a *Strauss Hardcore process* or the hypothesis that y is a realization of a *Geyer Saturation process* as seen in Figs. 9 and 10.

Our measurement of the standard deviation σ of the reduced cell areas of a PPP is .525, close to .529 which is the standard deviation measured in the extensive simulation done in [20]. We include the standard deviation measurements for the other processes we used in Table I, where it can be seen that the standard deviation of *Voronoi cell* areas of x and y are closest to those of the *Strauss Hardcore process* and *Geyer Saturation*

process, respectively.

Using Node Clustering Metric $g(r)$: By using $g(r)$, we can reject the hypothesis that either x or y is *Poisson* as seen in Fig. 11. We cannot reject the hypothesis that x is a hexagonal distribution with perturbed points by this metric as seen in Fig. 12. This is because we chose the z parameter of point perturbation so that the model would closely match this point pattern by this metric. We cannot reject either the hypothesis that x is a realization of a *Strauss Hardcore process* or the hypothesis that y is a realization of a *Geyer Saturation process* as seen in Figs. 13 and 14.

V. CONCLUSION

In this paper we tested the hypothesis that either of two different types of tower deployments were realizations of *Poisson processes*. We rejected this hypothesis using SIR, *Voronoi cell* area distributions, and a modified Ripley's K

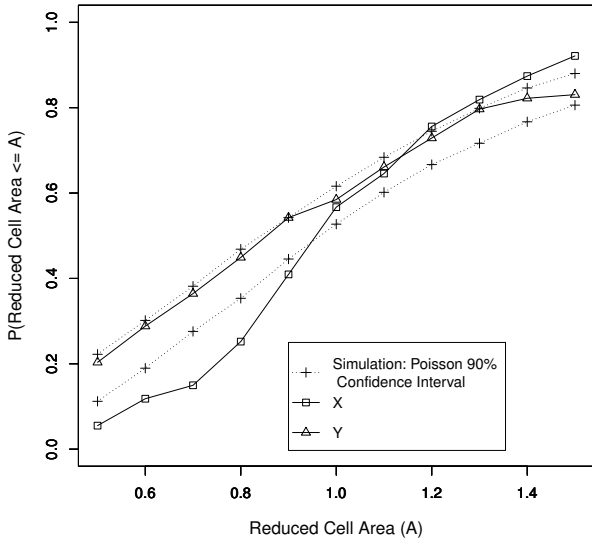


Fig. 7. Rejection of the hypothesis that x is *Poisson* distributed. y lies within bounds of a *Poisson* realization.

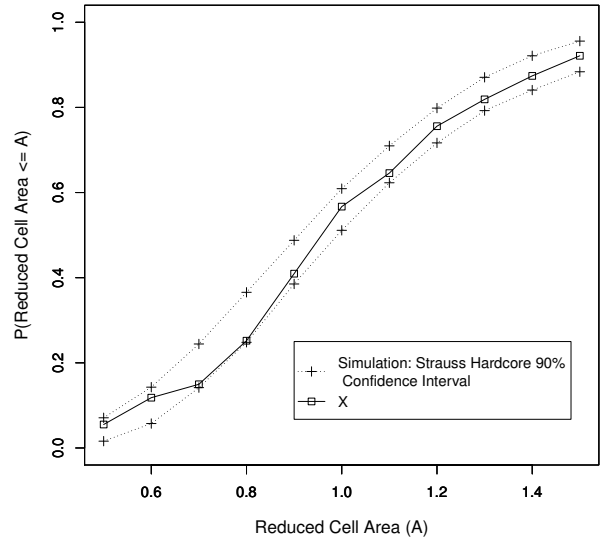


Fig. 9. x 's *Voronoi cell* area distribution lies within bounds of a *Strauss Hardcore* process.

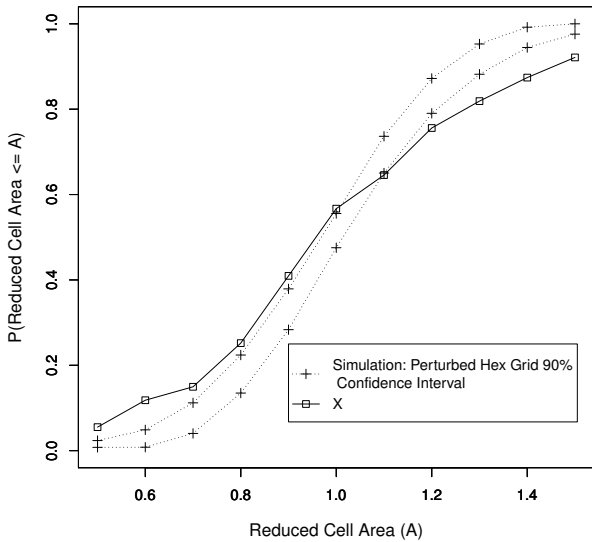


Fig. 8. Rejection of the hypothesis that x is a hexagonal grid with perturbed points by *Voronoi cell* area distributions.

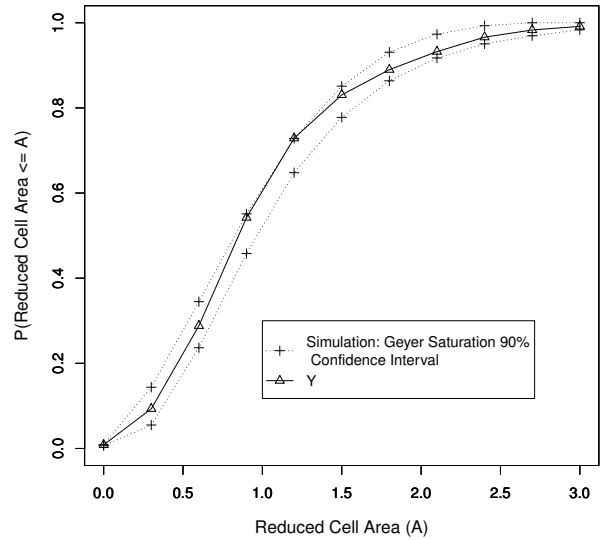


Fig. 10. y 's *Voronoi cell* area distribution lies within bounds of a *Geyer Saturation* process.

function. We showed that by using the same metrics we cannot reject the hypothesis that both point patterns were realizations of pairwise and saturated pairwise interaction processes fitted to them. Finally, we used *Voronoi cell* area distributions to show that variations of the hexagonal grid do not accurately model coverage cell size, an important spatial characteristic of cellular networks.

Future work can be done estimating parameters of these pairwise interaction processes without knowing the actual node placement by, for example, using population or terrain data. These ideas can be extended to multi-tier networks. In a K -tier network, there can be up to $T(K)$, the triangular number sequence, different pairwise interactions, one for each pairing of the types of nodes.

REFERENCES

- [1] Qualcomm, "LTE advanced: heterogeneous networks," white paper, Jan. 2011.
- [2] 3GPP, "TR 25.814 v7.1.0: Physical layer aspects for evolved UTRA," Sept. 2006.
- [3] J. Niemelä and J. Lempiäinen, "Impact of base station locations and antenna orientations on UMTS radio network capacity and coverage evolution," in *IEEE Int. Symp. On Wireless Personal Multimedia Communications Conf.*, Yokosuka, 2003.
- [4] J. Turkka and A. Lobinger, "Non-regular layout for cellular network system simulations," in *IEEE 21st International Symposium on Personal Indoor and Mobile Radio Communications*, Sept. 2010, pp. 1929–1933.
- [5] S. Hua, P. Liu, and S. S. Panwar, "The urge to merge: When cellular service providers pool capacity," available online: arxiv.org/abs/1111.1964.
- [6] J. F. C. Kingman, *Poisson Processes*. Oxford University Press, 1993.
- [7] T. Brown, "Cellular performance bounds via shotgun cellular systems," *IEEE Journal on Sel. Areas in Communications*, vol. 18, no. 11, pp. 2443–2455, Nov. 2000.
- [8] J. G. Andrews, F. Baccelli, and R. K. Ganti, "A tractable approach

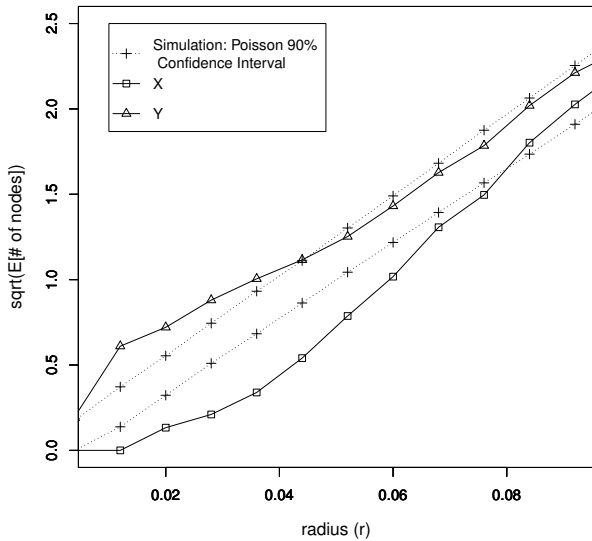


Fig. 11. Rejection of the hypothesis that x or y is *Poisson* by $g(r)$.

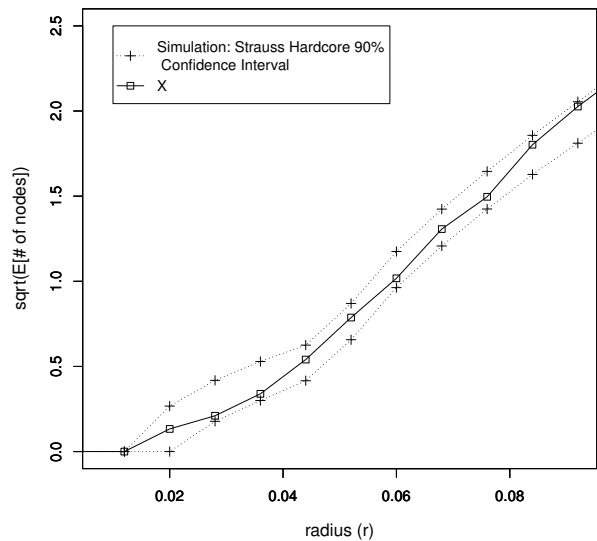


Fig. 13. x 's $g(r)$ lying within bounds of a *Strauss Hardcore* process.

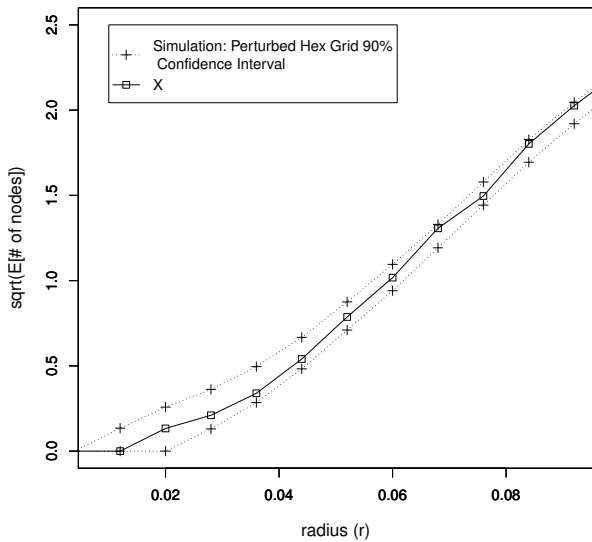


Fig. 12. x 's $g(r)$ area distribution lying within bounds of a hexagonal grid with perturbed points.

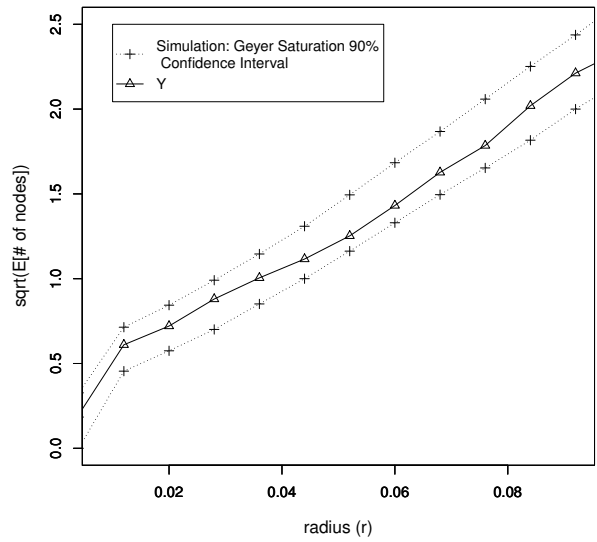


Fig. 14. y 's $g(r)$ lying within bounds of a *Geyer Saturation* process.

- to coverage and rate in cellular networks," *IEEE Transactions on Communications*, vol. 59, no. 11, pp. 3122–3134, Nov. 2011.
- [9] H. S. Dhillon, R. K. Ganti, and J. G. Andrews, "A tractable framework for coverage and outage in heterogeneous cellular networks," in *Proc., Inform. Theory and its Applications (ITA)*, San Diego, CA, Feb. 2011.
- [10] 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 Sel. Areas in Communications*, vol. 30, no. 3, pp. 550 – 560, Apr. 2012.
- [11] J. Møller and R. P. Waagepetersen, "Modern statistics for spatial point processes," *Scandinavian Journal of Statistics*, vol. 34, pp. 643–684, 2007.
- [12] J. Riihijärvi and P. Mähönen, "Modeling spatial structure of wireless communication networks," in *IEEE NetSciCom*, San Diego, CA, Mar. 2010.
- [13] D. J. Strauss, "A model for clustering," *Biometrika*, vol. 63, pp. 467 – 475, 1975.
- [14] A. Baddeley and R. Turner, "spatstat: An R package for analyzing spatial point patterns," *Journal of Statistical Software*, vol. 12, no. 6, pp. 1–42, 2005.
- [15] F. P. Kelly and B. D. Ripley, "A note on Strauss's model for clustering," *Biometrika*, vol. 63, no. 2, pp. 357 – 360, Aug. 1976.
- [16] J. Møller and R. P. Waagepetersen, *Statistical Inference and Simulation for Spatial Point Processes*. Boca Raton: Chapman and Hall/CRC, 2003.
- [17] J. Besag, "Contribution to the discussion of Dr. Ripley's paper," *Journal of the Royal Statistical Society*, vol. 39, pp. 193–195, 1977.
- [18] M. Van Lieshout and A. Baddeley, "A nonparametric measure of spatial interaction in point patterns," *Statistica Neerlandica*, vol. 50, pp. 344–361, 1996.
- [19] A. Baddeley and R. Turner, "Practical maximum pseudolikelihood for spatial point patterns (with discussion)," *Australian and New Zealand Journal of Statistics*, vol. 42, no. 3, pp. 283–322, 2000.
- [20] M. Tanemura, "Statistical distributions of poisson Voronoi cells in two and three dimensions," *Forma*, vol. 18, no. 4, pp. 221–247, 2003.