Graph Approximations of Spatial Wireless Network Models

Janne Riihijärvi, Elena Meshkova and Petri Mähönen
Institute for Networked Systems, RWTH Aachen University
Kackertstrasse 9, D-52072 Aachen, Germany
email: {jar, eme, pma}@inets.rwth-aachen.de

ABSTRACT

Wireless networks are usually modeled as spatial objects, and more specifically with random processes generating the locations of individual nodes. This is in stark contrast to the fixed networks case, in which graph models are typically used. While expressive, spatial models are usually more difficult to analyze than graph based ones, which has resulted in adoption of oversimplified models in wireless networks research. In this paper we show how wireless networks can be modeled as graphs without losing the accuracy of spatial models. We argue based on measurements of wireless network performance that in a number of cases distance-dependent interactions between wireless nodes can be discretized with only small approximation error. This discretization immediately yields an approximation of a spatial wireless network model as a graph with multiple edge types. We study the structure of the arising graph models, and show that the choice of accurate underlying spatial model remains important. For accurate spatial models, the corresponding graph approximations have a relatively simple neighborhood structure, indicating that they can be used very effectively in performance evaluation and optimization applications.

Categories and Subject Descriptors

C.2 [**Computer-Communication Networks**]: Miscellaneous

General Terms

Theory, Performance

1. INTRODUCTION

One of the fundamental objectives of communications research is to provide *models* of network structure that can be used to understand and optimize the performance of networks. For fixed networks such models have traditionally been based on random graphs of different types, whereas recently there has been strong emphasis on models based on performance optimization [1–5]. Wireless networks on the other hand have usually been modeled as *spatial* objects, focusing on the distribution of node locations in some region $D \subset \mathbb{R}^2$ [6–8]. This has been done because performance of wireless networks depends directly on distances between the nodes, propagation conditions, and the amount of traffic and interference generated in a region. Particularly common assumption has been that the node locations are uniformly and randomly distributed in D. In [7] we developed techniques for more accurate modeling of node locations based on the use of *spatial statistics* [9, 10], and fitted stochastic location models on a variety of data sets. The resulting models are accurate, but somewhat challenging to work with. In particular, extensive Monte Carlo simulations are often needed for performance studies using these models.

In this paper we introduce a novel modeling method based on approximating spatial wireless network models by graphs with multiple types of edges. The motivation for developing such approximations is the fact that while wireless networks are indeed inherently spatial objects, their performance is *not* always tightly coupled to the distances between the nodes. For example, many earlier measurements on Wi-Fi and cellular networks have shown that for small networks there are rather sharp phase transitions in performance as a function of distances between the nodes. There often exists clear regime of high performance for a wide range of distances, followed by a "gray zone" with fluctuating performance, and finally a large region in which communication is no longer possible but nodes still can detect ongoing transmissions by carrier sensing. This indicates that instead of full spatial modeling of node locations, one can model the network as a graph with multiple edge types used to encode different types of potential interactions. In the above example, first type would correspond to the high performance regime, the second to the gray zone, and third to the sensing range.

Our work is closely related to two parallel lines of research that have emerged in recent years. First is the use of *random geometric graphs* [11] (of which the so-called *unit disk graphs* are a special case) to study especially the connectivity properties of networks. The main shortcomings of this approach have been that node locations are still usually assumed to be uniformly random, and the relationships between the nodes has been assumed to be purely binary (node pairs are connected if and only if they closer than some threshold distance r_{th} apart). However, to capture

PM^{2}HW^{2}N'11, October 31, 2011, Miami, Florida, USA.

the various interactions taking place in actual wireless networks multiple edge types are needed as argued above. The second line of work has been the study the study of performance of specific wireless networking protocols on specific instances of such graphs [12–15]. In particular, it has been shown that performance metrics of interest are highly sensitive to the structure of small neighborhood of the node and the corresponding edge types. Based on these observations, we focus in this paper specifically on studying which kinds of neighborhood structures, including edge types, typically arise as outcomes of different spatial network models. Our results thus directly yield *graph approximations* for spatial network models which can be directly used for similar applications as the full spatial models, but with significantly reduced complexity.

The rest of this paper is structured as follows. In Section II we provide a short overview of typical spatial wireless network models, and introduce the graph approximations of these in Section III. We then study the neighborhood structure of the arising models in Section IV. Finally, we draw our conclusions and outline topics for future work in Section V.

2. MODELS OF SPATIAL NETWORK STRUCTURE

Typical models of spatial wireless network structure focus on the distribution of node locations in some region $D \subset \mathbb{R}^2$. The mathematical foundation for reasoning about such node location distributions is the *stochastic geometry of point processes* [16, 17]. Informally a point process N is a random collection of n locations $\{x_i\}_{i=1}^n \in D^n$, where in addition to the individual locations $x_i \in D$, the total number of them, $n \in \mathbb{N}$, can be random as well. The formal definition of a point process is usually given in terms of random counting measures assigning to each suitably regular subset $A \subseteq D$ a random integer $N(A)$, the number of points in A. However, for our purposes the more intuitive definition suffices, and we refer the reader interested in the rigorous mathematical framework of point processes to [16, 17].

The simplest example of a point process is the *homogeneous Poisson point process* (PPP) defined by two properties. First, the number of nodes $N(A)$ in a region $A \subseteq D$ follows a Poisson distribution with parameter $\nu|A|$, where $|A|$ denotes the area of A, and $\nu > 0$ is called the *density* of the process. Second, for disjoint regions A_1 and A_2 the point counts $N(A_1)$ and $N(A_2)$ are independent random variables. The homogeneous Poisson point process has very simple probabilistic structure as can be seen from these definitions, which has made it by far the most widely used model for the spatial structure of wireless networks. However, despite its popularity, it is commonly acknowledged that most actual network deployments do deviate significantly from the uniformly random nature of the Poisson model (see [18–20] for quantitative discussion). These deviations have usually been accounted for by developing models that are either more *clustered* (for nodes that tend to be located in groups, such as mobile terminals) or *regular* (for nodes that are rarely nearby each other, such as cellular network base station locations or planned Wi-Fi access point deployments). For space reasons we shall focus here on the regular case, and apply the corresponding models in a scenario where the points of the process correspond to locations of Wi-Fi access points. We refer the reader to [8, 21] for more detailed discussion on the clustered case, including discussion on performance implications.

Perhaps the simplest example of a regular point process is the *simple sequential inhibition process* (SSI). It is defined by two parameters, the total number of points $n = \nu|D|$, and a *hard core radius* r which defines a minimum allowed separation between the points. Realizations of an SSI are generated by throwing points one at a time uniformly on D, while enforcing the condition that no two points can be closer than distance r apart. This is continued until locations for all the n points have been generated, or placing further points becomes impossible due to too large network density or too large r.

The Poisson and SSI models are useful for gaining intuition on the influence of regularity on network performance. However, they are not very well suited for working with empirical data. For each model, the arising point patterns do not usually correspond well to actual network structures for any choices of parameter values, and it is not clear which criteria one should use for evaluating the goodness of fit. As one possible solution, we showed in [7] that the very general *Gibbs point processes* (GPPs) can be be used to correct both of these shortcomings. A GPP is defined by prescribing a *density function* $f(\boldsymbol{x})$, which assigns to each point configuration $\{x_i\}_{i=1}^n \in D^n$ a likelihood of that configuration to arise compared to the case of a homogeneous Poisson point process. In particular, if $f(\boldsymbol{x}) = 0$ the corresponding point configuration can never occur, if $0 < f(\boldsymbol{x}) < 1$ the given configuration can occur but is less likely to do so than in the PPP case, and, finally, if $f(\boldsymbol{x}) > 1$ the given configuration is more likely to occur than for PPP. The particular model used in [7] was the *Geyer saturation process*, in which each point x_i of the pattern $\{x_i\}_{i=1}^n$ contributes a factor of

$$\beta\gamma^{\min\{\zeta, s_r(x_i, \boldsymbol{x})\}} \tag{1}$$

to the density. Here $\beta > 0$ controls the density of the points, $\gamma > 0$ the type of interaction (whether clustered or regular), $s_r(x_i, \boldsymbol{x})$ denotes the number of points in $\boldsymbol{x}$ that are closer to x_i than distance r, and ζ is the *saturation threshold* bounding the contribution of the individual points and ensuring that the arising density is integrable. Such models can be fitted to location data sets using standard maximum pseudolikelihood techniques, and we showed by analyzing the residuals as well as various spatial statistics measures in [7] that for a wide variety of empirical data sets very good fits were obtained.

In the following we shall use these fitted Geyer models as a representative example of a realistic spatial wireless network model. Even more general Gibbs models can, of course, be developed. However, in our experience, the Geyer model is very well balanced. It does not have too many parameters so that overfitting would be a problem, but it is still able to model accurately a wide variety of network types and remains tractable analytically.

3. GRAPH APPROXIMATIONS OF SPATIAL MODELS

We shall now discuss the problem of approximating a given spatial wireless network model using graphs with mulitple types of edges. Our approach is based on the observation that typical wireless networking technologies have distinct operational regimes, depending on the strength of the sig-

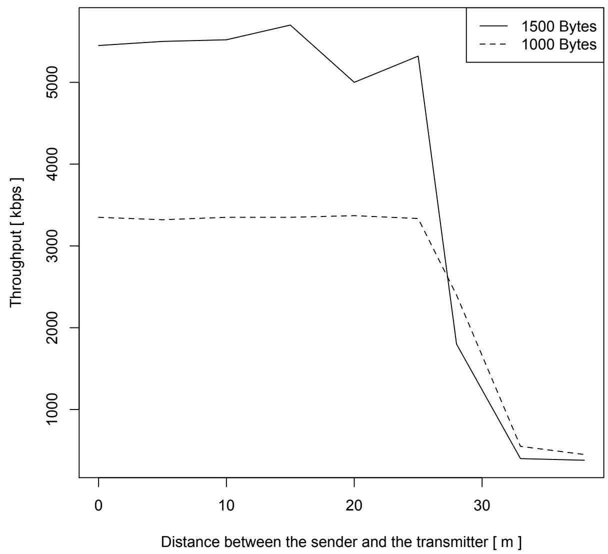

Figure 1: Throughput of a UDP connection over a IEEE 802.11b Wi-Fi link as a function of distance between the transmitter and the receiver for two different frame sizes.

nal the received from the transmitter. For example, Wi-Fi networks tend to exhibit similar performance in terms of average throughput and packet error rate over a wide range of distances between the transmitter and the receiver. This is illustrated in Figure 1, showing the throughput of a UDP stream over a IEEE 802.11b Wi-Fi link obtained from earlier measurements in an indoor testbed [22]. Similar results have been obtained for outdoor scenarios and for other Wi-Fi technologies as well, see, e.g., [23]. For both of the packet sizes considered in the measurements performance is practically independent of the distance until separation of the two nodes exceeds roughly 28 m, after which the quality of the connection degrades rapidly. The nodes do continue to interact over much wider range of distances however, due to the listen-before-talk or carrier sense mechanism used in Wi-Fi. Accordingly, for evaluating Wi-Fi performance, these results indicate that two or three edge types would suffice. The high performance regime and the carrier sense regime always need separate edge types. For outdoor scenarios the gray zone tends to be very narrow, so the corresponding type can often be omitted without reducing the accuracy of the arising model significantly. However, for more complex indoor propagation environment a separate type might be needed to account for it.

We formalize these observations by defining a *graph approximation G_N of a point process N* as follows. The vertex set of G_N is taken to coincide with the points in the realization of N. The edge between two vertices x_i and x_j is then assigned integer type $(n+1)-\sum_{k=1}^{n} \mathbf{1}(\|x_i - x_j\| \le d_k)$, where $\mathbf{1}(\cdot)$ is the indicator function, $\|x_i - x_j\|$ denotes the distance between x_i and x_j, the $\{d_k\}_{k=1}^{n}$ are the *threshold distances*, and type of $n+1$ indicates that x_i and x_j are not adjacent in the arising graph. This is a direct generalization of the usual definition of a random geometric graph [11], for which N is usually assumed to be the Poisson point process, and only one threshold distance is used[1]. The inherent assumption in this definition is that there are no significant anisotropies (dependencies of performance on the angles between the transmitter and the receiver) in either radio propagation conditions, or the antennas. For scenarios in which these assumptions do not hold with sufficient accuracy (use of directional antennas being the key example), more general approximation scheme is needed. This can be easily achieved by considering the statistics of the signal-to-noise ratio at the receiver for each potential transmitter-receiver pair, obtained by incorporating models of as many environmental characteristics as required.

4. NEIGHBORHOOD STRUCTURE OF ARISING MODELS

We shall now study the structure of graph approximations for the spatial wireless network models introduced in Section II. Our objectives are twofold. First, we illustrate the typical structures of the graph approximations, focusing on graph characteristics that are known to have major impact on the performance of the network being modeled. Second, we use these graph approximations as means to highlight the differences in network structure induced by the choice of the underlying spatial model.

We consider two distinct scenarios to highlight the impact of the choice of the models and their parameters. First of these is an *Outdoor Wi-Fi scenario*, corresponding to the structure of the Google Mountain View Wi-Fi network studied in [7]. Second scenario corresponds to a denser network deployment, which we shall call the *Indoor Wi-Fi* scenario in the following. In both cases we focus on the placement and interactions between the Wi-Fi access points only. The parameters for the Poisson, SSI and Geyer models for these two scenarios are given in Table 1. The parameters for the Geyer model for the Outdoor Wi-Fi scenario are an outcome of the careful fitting procedure discussed in [7], whereas the parameters for other models have been chosen to yield the same overall network density and otherwise similar spatial structure. The parameters for the Indoor Wi-Fi scenario have also been constructed on the basis of the Outdoor Wi-Fi scenario by increasing the density of the network and reducing the interaction radius for the SSI and Geyer models. Due to space reasons we shall focus on the case of graph approximations with two edge types, ignoring the gray zone, with threshold distances of $(d_1, d_2) = (200\,\text{m}, 400\,\text{m})$ for the Outdoor Wi-Fi scenario, and $(d_1, d_2) = (50\,\text{m}, 100\,\text{m})$ for the Indoor Wi-Fi scenario. All of the results discussed below were obtained through 10^4 Monte Carlo simulation runs, conditioned to have a node in the middle of the simulation area to avoid edge effects. Results are reported with respect to this "typical node", formally corresponding to the estimates of the corresponding functionals of the Palm distribution of the underlying point process [16, 17].

[1]Random geometric graphs for which the threshold distance is one are also occasionally called *unit disk graphs* in the wireless networking literature. Generalizations to the case in which edges exist with certain probability are then called *quasi unit disk graphs*. The difference to the latter in our treatment arises from differentiating between multiple types of edges, enabling much more realistic modeling and approximation of the spatial structure of the underlying network.

Scenario	Model	Parameters
Outdoor Wi-Fi	Poisson	$\nu = 1.35 \times 10^{-5}$
	SSI	$\nu = 1.35 \times 10^{-5}$, $r = 100$
	Geyer	$\beta = 3.3 \times 10^{-5}$, $\gamma = 0.4$, $\zeta = 2$, $r = 150$
Indoor Wi-Fi	Poisson	$\nu = 2.7 \times 10^{-4}$
	SSI	$\nu = 2.7 \times 10^{-4}$, $r = 30$
	Geyer	$\beta = 6.8 \times 10^{-4}$, $\gamma = 0.65$, $\zeta = 2$, $r = 80$

Table 1: The considered scenarios and corresponding models and parameter values.

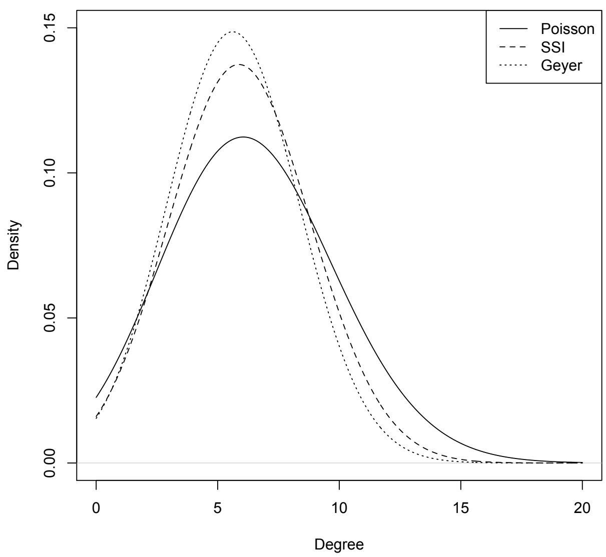

Figure 2: The overall degree distribution for the Outdoor Wi-Fi scenario.

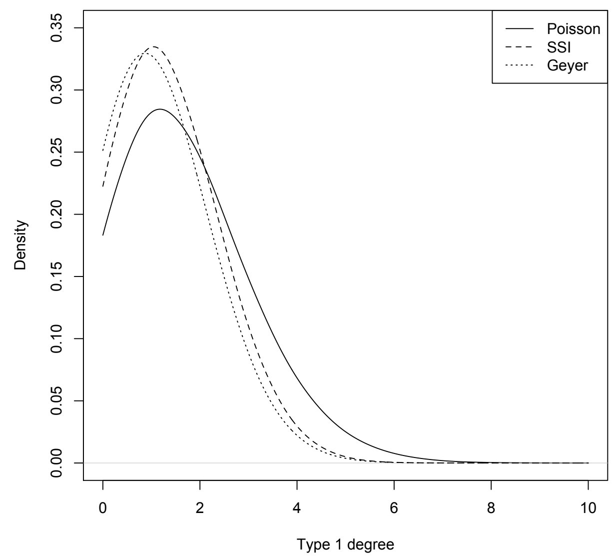

Figure 3: Distribution of type one degrees for the Outdoor Wi-Fi scenario.

We shall first study the degree distributions of the arising graph approximations. Figure 2 shows the overall degree distribution ignoring the types of the edges for the Outdoor Wi-Fi scenario[2]. The difference between the three underlying choices of spatial models is significant, with the Geyer model resulting in the most concentrated distribution, and Poisson model resulting in the highest variation in the degrees. These differences continue to persist in the individual degree distributions of type one shown in Figure 3 and type two given in Figure 4, although the SSI model becomes quite close in terms of behavior to the Geyer one.

The degree distribution of the arising graph approximation is an important statistic for determining several types of performance characteristics. For example, type two degree in Wi-Fi networks has a direct relation to the amount of contention experienced by the node, and thus indicates for a given traffic load how likely the node is to be able to access the channel. Thus, from these results one can infer that the SSI model and the corresponding graph approximations would yield similar results as for the Geyer model case, and at least be closer to the reality than the Poisson model and its approximations. Interestingly, this conclusion becomes false as the density of the network is increased. Figure 5 shows the distribution of type one degrees for the Indoor Wi-Fi scenario. Despite a careful choice of parameters, the results for the SSI case are significantly different from the both Poisson and Geyer alternatives. Figure 6 shows similar results for the type two degrees, for which the correspondence is once more much closer.

While the degree distribution often provides good first-order approximation of the performance characteristics of a wireless network, it does not contain information on the interactions *between* the neighboring nodes. Especially when the network becomes denser, these interactions start to dominate the performance characteristics, and more refined graph statistics are needed. In [12–15] it has been shown that espe-

[2]Strictly speaking the degree distribution is, of course, defined on the integers. We show here smoothed kernel density estimates to make it easier to see the differences between the degree distributions arising from the different models.

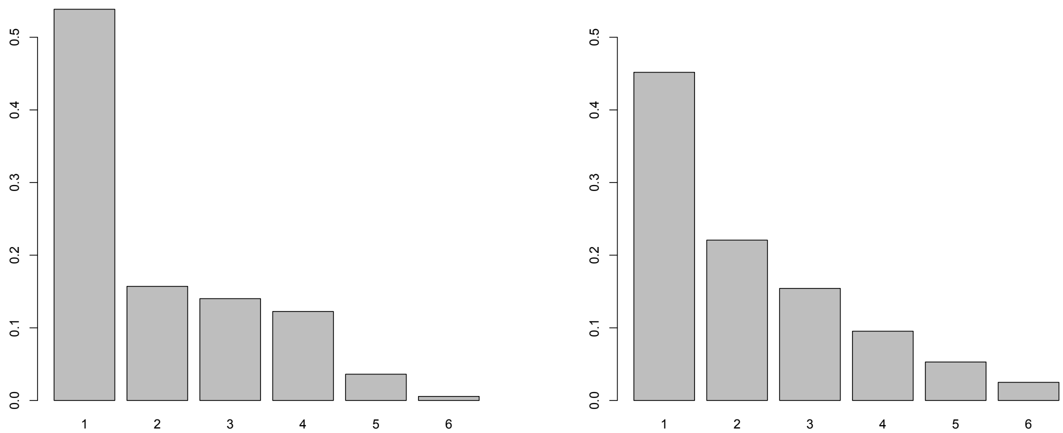

Figure 7: The rank-frequency plots for the Geyer model (left) and for PPP (right) for the occurrence of three-node network motifs for the Outdoor Wi-Fi scenario. See Figure 8 for illustration of the motifs.

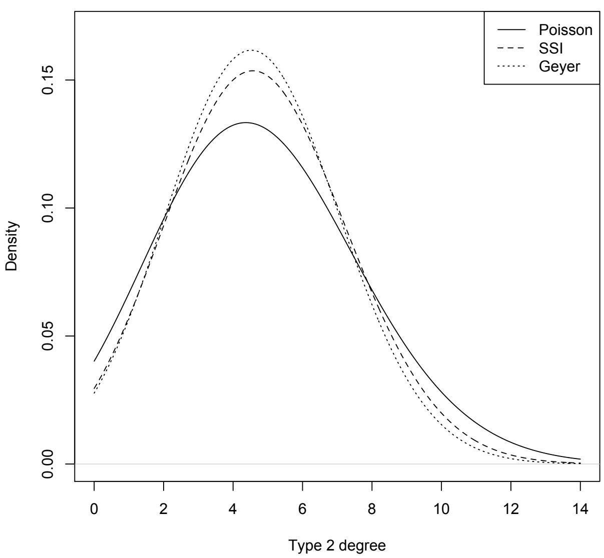

Figure 4: Distribution of type two degrees for the Outdoor Wi-Fi scenario.

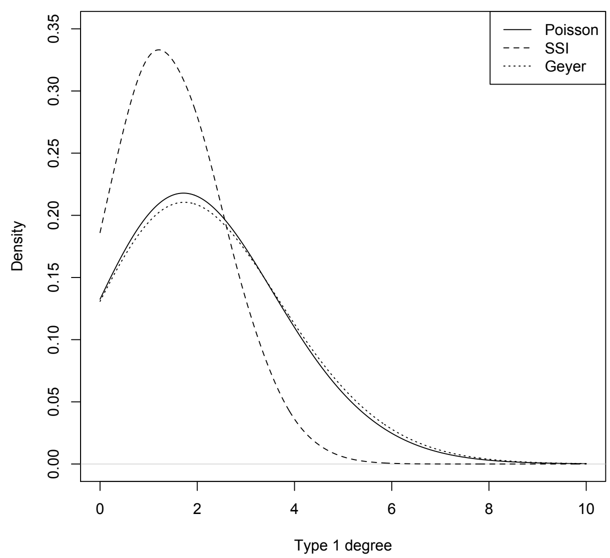

Figure 5: Distribution of type one degrees for the Indoor Wi-Fi scenario.

cially in the case of Wi-Fi networks, accurate results can often be achieved by studying the structure and occurrence of *small subgraphs* in the neighborhood of a typical node. This is similar to the observation made in many other research communities that such small subgraphs, usually called *network motifs* [24], often play a key role in the arising dynamics.

Based on these observations, we have studied the occurrence of different three- and four-node connected subgraphs or network motifs the typical node belongs to. We used the nauty toolkit [25] to classify these into their respective isomorphism classes, since only those have impact on the network dynamics. Figure 7 shows the rank-frequency plot for the occurrence of the different three-node subgraphs for the Geyer and Poisson models for the Outdoor Wi-Fi scenario. We see that out of the seven possible subgraphs only six actually occur in the results, and that the probabilities of occurrence are both highly non-uniform and significantly different between the two underlying spatial models. The occurring six motifs are illustrated in Figure 8.

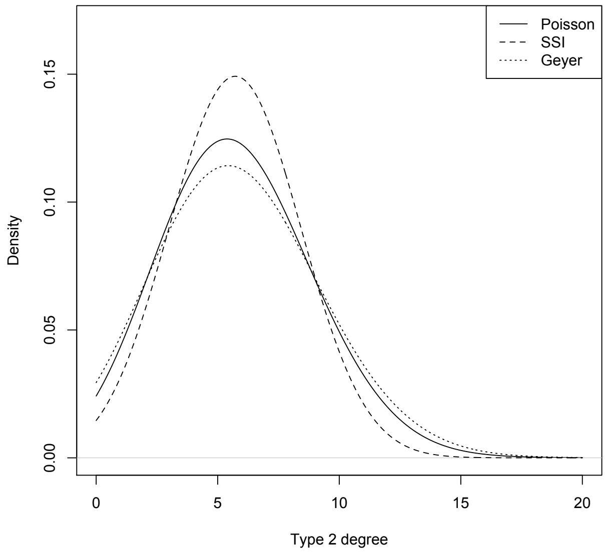

Figure 6: Distribution of type two degrees for the Indoor Wi-Fi scenario.

Figure 9 shows similar results but for the case of four-node subgraphs. Also in this case only 35 of the possible subgraphs actually occurred, a bit more than half of all the possible ones [15], and the corresponding probability distribution continued to be highly non-uniform. Also the differences between Poisson and Geyer cases continue to be quite clear. These results strongly indicate that for many types of wireless network deployments, the performance of the corresponding protocols and networking technologies depends heavily on how they behave in a small subset of the theoretically possible network configurations.

5. CONCLUSIONS

In this paper we have studied the problem of approximating spatial wireless network models with graphs. Based on earlier measurement results we formulated a simple generalization of the well-known random geometric graphs that is often able to accurately model the possible interactions of wireless network nodes for several scenarios and technologies of interest. We then explored the structure of the arising graph models with especially studying of Wi-Fi networks in mind. The results show that the selection of the underlying spatial model remains to be important, as all the statistics observed showed major differences between the spatial models included in the study. We also observed that of all the possible interaction patterns, only a small portion actually occurred in the results, and that the probabilities of different patterns to arise were highly non-uniform.

These results have a number of applications. For example, the use of graph approximations is an effective way to decouple the performance evaluation of a protocol design for a wireless network from the underlying spatial model. One can study separately the influence of the occurrence of small subgraphs in the neighborhood of the typical node

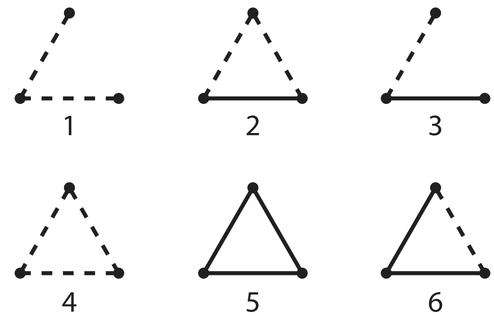

Figure 8: The three-node network motifs occurring in the Outdoor Wi-Fi scenario. Solid lines denote edges of type one, and dashed lines edges of type two.

on network performance from the probability of the given subgraph to actually occur. This allows especially the separation of analytical or simulation work from the underlying point process model, significantly reducing the complexity of carrying out performance evaluation for non-trivial spatial wireless network models.

Acknowledgment

The authors would like to thank RWTH Aachen University and the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) for providing financial support through the UMIC research centre. We would also like to thank the European Union for providing partial funding of this work through the FARAMIR project.

6. REFERENCES

[1] R. Albert and A. Barabási, "Statistical mechanics of complex networks," *Reviews of Modern Physics*, vol. 74, no. 1, pp. 47–97, 2002.

[2] M. Newman, "Random graphs as models of networks," *Arxiv preprint cond-mat/0202208*, 2002.

[3] W. Willinger, D. Alderson, and J. Doyle, "Mathematics and the internet: A source of enormous confusion and great potential," *Notices of the American Mathematical Society*, vol. 56, no. 5, pp. 586–599, 2009.

[4] J. Doyle, D. Alderson, L. Li, S. Low, M. Roughan, S. Shalunov, R. Tanaka, and W. Willinger, "The "robust yet fragile" nature of the Internet," *Proceedings of the National Academy of Sciences*, vol. 102, no. 41, pp. 14 497–14 502, 2005.

[5] R. V. Oliveira, D. Pei, W. Willinger, B. Zhang, and L. Zhang, "In search of the elusive ground truth: the internet's as-level connectivity structure," *SIGMETRICS Perform. Eval. Rev.*, vol. 36, no. 1, pp. 217–228, 2008.

[6] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic Geometry and Random Graphs for the Analysis and Design of Wireless Networks," *IEEE JSAC*, Sep. 2009.

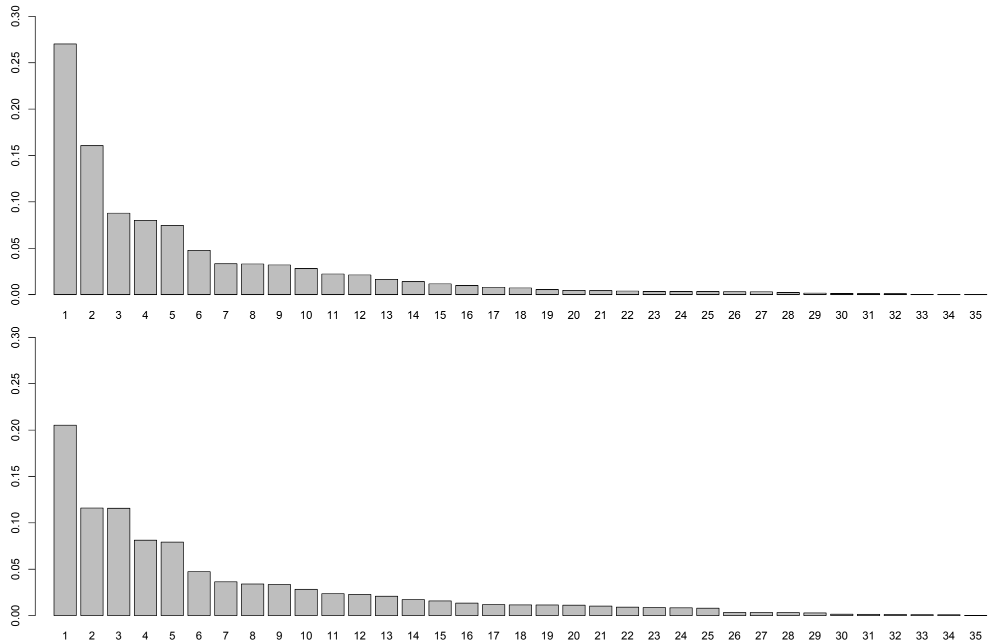

Figure 9: The rank-frequency plots for the Geyer model (top) and for PPP (bottom) for the occurrence of four-node network motifs for the Outdoor Wi-Fi scenario.

[7] J. Riihijärvi and P. Mähönen, "Modeling Spatial Structure of Wireless Communication Networks," in *NetSciCom 2010*, 2010, pp. 1–6.

[8] R. Ganti and M. Haenggi, "Interference and outage in clustered wireless ad hoc networks," *IEEE Transactions on Information Theory*, vol. 55, no. 9, pp. 4067–4086, 2009.

[9] B. Ripley, "Spatial Statistics." *John Wiley & Sons, Inc.*, 1981.

[10] N. Cressie, *Statistics for spatial data*. John Wiley & Sons, New York, 1993.

[11] M. Penrose, *Random geometric graphs*. Oxford University Press, 2003.

[12] M. Garetto, J. Shi, and E. Knightly, "Modeling media access in embedded two-flow topologies of multi-hop wireless networks," in *Proceedings of the 11th annual international conference on Mobile computing and networking*. ACM, 2005, pp. 200–214.

[13] M. Garetto, T. Salonidis, and E. Knightly, "Modeling per-flow throughput and capturing starvation in CSMA multi-hop wireless networks," *Networking, IEEE/ACM Transactions on*, vol. 16, no. 4, pp. 864–877, 2008.

[14] C. Hua and R. Zheng, "Starvation modeling and identification in dense 802.11 wireless community networks," in *INFOCOM 2008. The 27th Conference on Computer Communications. IEEE*. IEEE, 2008, pp. 1022–1030.

[15] S. Razak, V. Kolar, and N. Abu-Ghazaleh, "Modeling and analysis of two-flow interactions in wireless networks," *Ad Hoc Networks*, vol. 8, no. 6, pp. 564–581, 2010.

[16] D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic geometry and its applications*. Wiley, 1995.

[17] A. F. Karr, *Point Processes and Their Statistical Inference*, 2nd ed. Marcel Dekker, 1991.

[18] J. Riihijärvi, P. Mähönen, and M. Rübsamen, "Characterizing wireless networks by spatial correlations," *IEEE Communications Letters*, vol. 11, no. 1, pp. 37–39, January 2007.

[19] J. Riihijärvi, M. Petrova, and P. Mähönen, "Influence of node location distributions on the structure of ad hoc and mesh networks," in *Proc. of IEEE Globecom 2008*, New Orleans, USA, November 2008.

[20] M. Michalopoulou, J. Riihijärvi, and P. Mähönen, "Studying the relationships between spatial structures of wireless networks and population densities," in *IEEE Globecom 2010*, Miami, Florida, December 2010.

[21] M. Petrova, P. Mähönen, and J. Riihijärvi, "Connectivity analysis of clustered ad hoc and mesh networks," in *Proceedings of IEEE GLOBECOM 2007*, Washington D.C., USA, April 2007.

[22] M. Petrova, L. Wu, M. Wellens, and P. Mähönen, "Hop of no return: Practical limitations of wireless multi-hop networking," *Proceedings of RealMAN*, 2005.

[23] M. Wellens, B. Westphal, and P. Mahonen, "Performance evaluation of IEEE 802.11-based WLANs in vehicular scenarios," in *VTC2007-Spring*. IEEE, 2007, pp. 1167–1171.

[24] U. Alon, "Network motifs: theory and experimental approaches," *Nature Reviews Genetics*, vol. 8, no. 6, pp. 450–461, 2007.

[25] B. McKay, "Nauty user's guide (version 1.5)," 1990.