**Proceedings Paper:**

El Hammouti, H, Ghogho, M and Zaidi, SAR (2019) A Machine Learning Approach to Predicting Coverage in Random Wireless Networks. In: Proceedings of 2018 IEEE Globecom Workshops (GC Wkshps). 2018 IEEE Globecom Workshops (GC Wkshps), 09-13 Dec 2018, Abu Dhabi, United Arab Emirates. IEEE . ISBN 978-1-5386-4920-6

https://doi.org/10.1109/GLOCOMW.2018.8644199

# A Machine Learning Approach to Predicting Coverage in Random Wireless Networks

Hajar El Hammouti*, Mounir Ghogho*† and Syed Ali Raza Zaidi†
*International University of Rabat, FIL, TICLab, Morocco
†University of Leeds, School of EEE, UK
{hajar.elhammouti,mounir.ghogho}@uir.ac.ma,s.a.zaidi@leeds.ac.uk.

***Abstract*—There is a rich literature on the prediction of coverage in random wireless networks using stochastic geometry. Though valuable, the existing stochastic geometry-based analytical expressions for coverage are only valid for a restricted set of oversimplified network scenarios. Deriving such expressions for more general and more realistic network scenarios has so far been proven intractable. In this work, we adopt a data-driven approach to derive a model that can predict the coverage probability in any random wireless network. We first show that the coverage probability can be accurately approximated by a parametrized sigmoid-like function. Then, by building large simulation-based datasets, the relationship between the wireless network parameters and the parameters of the sigmoid-like function is modeled using a neural network.**

***Index Terms*—Coverage probability, sigmoid function, neural networks, machine learning, stochastic geometry.**

## I. Introduction

Motivated by its tractability, researchers have widely adopted stochastic geometry to model wireless networks performance and understand their behaviors [1]. Valuable and insightful stochastic geometry-based analytical expressions can be found in the literature [2–7]. However, to ensure tractability, these expressions are based on many simplifying assumptions on the wireless network, which are often unrealistic [2]. For general network setups, deriving analytical expressions to predict performance is very often unfeasible.

The following three scenarios illustrate the above-mentioned limitations.

- *Correlated shadowing*: most stochastic geometry-based studies have either neglected shadowing in the channel modeling, or assumed it to be a spatially independent process, following a log-normal distribution. Indeed, when spatial correlation is considered, tractability of the analytical derivations is in general no longer possible with stochastic geometry theoretical tools. The authors in [8], [9] have considered correlated shadowing but assumed a particular shadowing model and ignored the path loss component to make the derivations tractable. Another approach that has been proposed in the literature is to approximate the interference, which includes a large number of log-normally distributed shadowing terms, by a gamma distribution which simplifies the derivations of the coverage probability; see for example [10].
- *Non-homogeneous base station distribution*: most stochastic geometry-based studies model the positions of the base stations using a homogeneous Poisson point process (HPPP). While this simplifies the analysis, it does not capture the repulsive nature of the spatial topology observed in real-world cellular networks; several works have shown that base stations locations are better modeled using a Matérn hard-core point process (HCPP) [5], [11]. This more realistic modeling however undermines the tractability of the analytical analysis of network performance [12].
- *Deterministic base station deployment*: deriving a closed-form expression for the average coverage probability in this scenario is a challenging task. Indeed, random spatial distribution models for BS positions simplify the analytical derivations. Introducing specific locations of BS in the network generally undermines the tractability of the derivations [6].

The main objective of this paper is to propose an easy-to-apply and practical approach to predicting network performance for any given network setup, which is characterized here by several network features describing the channel modeling, the base station distribution, user association scheme, etc. Our approach is data-driven and borrows machine learning tools to determine an accurate mapping between the network features and its performance. In this paper, the performance metric is the coverage probability of a typical user, but the approach could be applied to other performance metrics.

### A. Related work

Deriving a closed-form expression of the coverage probability as a function of the network parameters is often a daunting task. In general, simplifying and often unrealistic assumptions such as Rayleigh channels, uncorrelated shadowing, or closest base station user-association are often considered to simplify the analysis.

Even with such assumptions, advanced mathematical techniques are involved in order to compute the coverage probability. As reported in [4], five techniques are commonly used to calculate coverage probability in stochastic geometry based networks. In general these techniques either (i) rely on the Rayleigh fading assumption, (ii) consider only the dominant or a limited number of the nearest interferers, (iii) approximate the probability density function (pdf) of the sum of the interference, (iv) use Plancherel-Parseval theorem, (v) or finally, invert the moment of the generating function to obtain the pdf of the interference. These many complicated mathematical processes involved in coverage probability computation make it difficult to have an easy-to-apply and practical approach to coverage prediction. Under these circumstances, the development of a general framework that captures the real complexity of the network system and proposes accurate, yet simple and direct, coverage probability prediction is of a great importance.

In the last few years, there has been a large interest in machine learning (ML) techniques to provide accurate analytical models based on the statistical analysis of data [13]. The main success of machine learning techniques can be attributed to its ability to map various network parameters to the network's response. Unlike the theoretical tools provided by stochastic geometry, a ML based approach captures the real complexity of the network by running a large number of measurements and/or experiments and proposes a mapping between input features and the output feature (network performance).

In [14] and [15], comprehensive surveys on the potential use of machine learning in 5G networks and wireless sensor networks, respectively, are provided. In [16], the authors compare a measurements-based prediction model to the signal-to-interference-and-noise-ratio (SINR) theoretical model. The paper shows that the ML approach outperforms the traditional SINR based model by providing results that are closer to the real measurements. Their ML approach is also used to predict the achievable throughput, and can thus be used for resource allocation optimization. In [17], the authors describe the experimental environment and methodologies to model the throughput of a transmission control protocol connection. The experimental results therein show that the throughput can be predicted with a very high accuracy using a support vector machine model [18]. In [19], the authors show that operators and service providers can adapt their services and contents using prediction models based on user's experience feedback. In particular, a supervised ML technique is proposed to overcome video starvation in large-scale wireless networks. Other machine learning applications can be found in [20] and [21]. In [20], a machine learning approach is proposed for drones to build a radio map that supports their path planning and positioning. In [21], the authors propose a neural network based approach for a better handover decision in heterogeneous networks. This approach was shown to improve the quality of service perceived by the users. In [22], the authors propose a distributed deep neural network to learn the optimal power allocation for a device-to-device network. The main advantage of such an approach is to reduce the computational complexity caused by optimization-based algorithms.

### B. Contribution

In this work, we are interested in the prediction of coverage probability of a typical user in a random wireless network using machine learning. To the best of our knowledge, this has not been addressed before. The contribution of this paper is twofold.

1) First, by running a large number of simulations, we show that the coverage probability can be closely approximated using a parametrized *sigmoid*-like function.
2) Second, we propose to use the exceptional ability of neural networks (NN) to approximate complicated functions in order to estimate the parameters of the sigmoid-like function from the feature set that characterizes the random wireless network, namely: base stations spatial intensity, path loss exponent, Nakagami-channel parameter, log-normal shadowing variance, log-normal spatial correlation, the BS transmit power, and background noise variance. This modeling is carried out for two user-association schemes.

### C. Structure

The rest of the paper is organized as follows. The next section describes the studied system model. Section III presents the proposed coverage probability approximation. Section IV proposes a method to learn the model's parameters using NN, and presents accuracy results. Finally, concluding remarks and possible extensions of this work are provided.

## II. System Model

In this section, we explain the general framework of our simulations. We consider a cellular network where the base stations (BS) are randomly distributed on the 2D plane following either (i) a homogeneous Poisson point process (PPP) of intensity $\lambda$ (ii) or a HCPP extracted from a PPP with the same intensity $\lambda$, with a given radius of the guard zone $Rc$, (iii) or a deterministic base station deployment with the same intensity $\lambda$ (here, we consider the conventional hexagonal pattern) as described in Fig. 1. Without loss of generality, we assume a typical user at the origin of the 2D plane. We assume two BS-user association

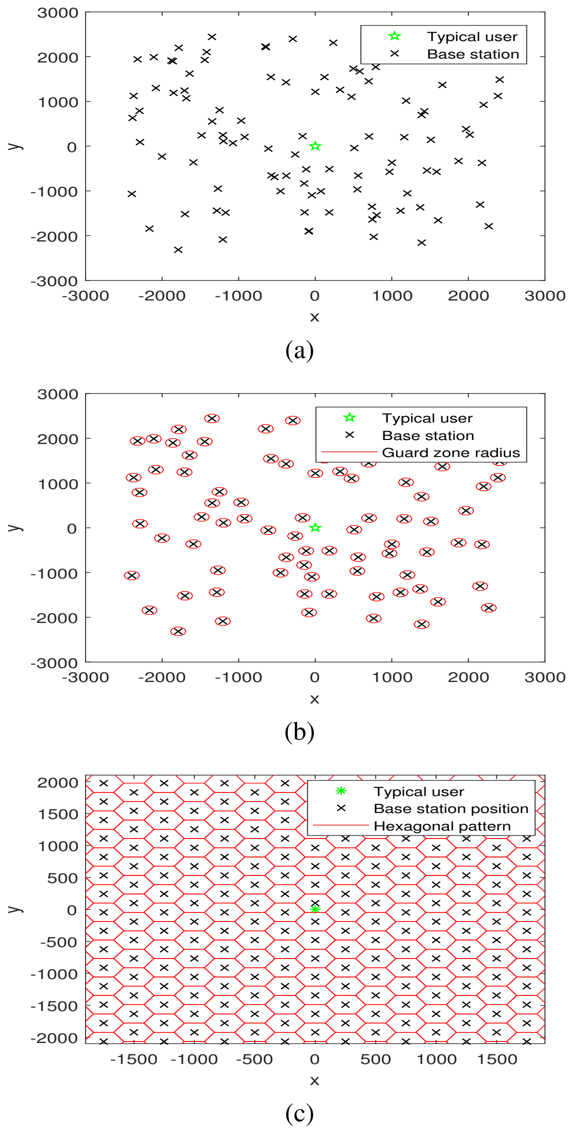


Fig. 1: (a) A PPP realization of base stations with $\lambda = 4.4{*}10^{-6}$ BS/$m^2$ , (b) A HCPP realization of base stations extracted from a PPP with $\lambda = 4.4{*}10^{-6}$ BS/$m^2$ when $R_c = 200$ m, (c) 256 BS positions following a hexagonal pattern.

schemes: (i) the user is associated with the nearest base station (Nearest Base station Association Scheme, NBAS), (ii) the user is served by the BS that provides the best signal-to-interference-and-noise ratio (SINR) (SINR Maximization Association Scheme, SMAS). Let $b_0$ denote the BS serving the typical user. In this paper, we focus on the downlink communication. The SINR experienced by the typical user is given by

$$\mathrm{SINR} = \frac{d_{b_0}^{-\alpha} G_{b_0} g_{b_0} P_{b_0}}{\sigma^2 + \sum\limits_{b \neq b_0} d_b^{-\alpha} G_b g_b P_b}, \tag{1}$$

where $P_b$ is the transmit power of BS $b$, $d_b$ is the distance between the typical user and BS $b$, $\alpha \in [2, 6]$ is the path loss exponent, $G_b$ is the channel power gain due to shadowing, $g_b$ is the small-scale fading power gain. In our simulations, we focus on the following setup: all BS transmit powers are equal to each other i.e. $P_b = P_{b_0} = P$, $\forall b$; the small-scale fading follows the general Nakagami distribution, i.e. $g_b \sim \Gamma(m, \frac{\omega}{m})$ follows a gamma law of parameters $(m, \frac{\omega}{m})$; the shadowing is log-normally distributed, i.e. $\log(G_b) \sim \mathcal{N}(0, \sigma_s^2)$. and is spatially correlated; the spatial correlation between two shadowing gains depends on the distance between the corresponding BS, as described by 3GPP in [23], i.e. the correlation between the shadowing gains associated with BS $i$ and BS $j$, $R(i, j)$, is an exponentially decreasing function of the distance separating the two BS, $\Delta d_{i,j}$, and so $R(i, j) = \exp(\frac{-\Delta d_{i,j}}{d_{\mathrm{cor}}})$ where $d_{\mathrm{cor}}$ is the correlation distance which controls the strength of the spatial correlation.

We are interested in the coverage probability of the typical user, which is defined as the probability that the SINR of that user is above a given threshold $\tau$, i.e.

$$p_c(\tau) = \mathbb{P}(\mathrm{SINR} > \tau). \tag{2}$$

In the case of hexagonal grid model for BS positions, we assume that the center of the grid is random in order to be able to use the same definition for the coverage probability as in the case of PPP and HCPP models.

## III. Modeling the Coverage Probability

We have generated a large number of network setups with different values of the path loss exponent, BS intensity, Nakagami channel parameter, transmit power-to-noise ratio $\gamma = P/\sigma^2$, variance of shadowing and correlation distance, and for each network setup, we have generated a large number of network realizations. For each network setup, we have estimated the average coverage probability for typical values of the threshold $\tau$. Examples of these estimation results are provided in Fig. 2. For each network setup, curve-fitting using the non-linear least squares method and different fitting models is then performed to model the average coverage probability of the typical user versus the threshold $\tau$. The simulation results indicate curve-fitting provides more compact models when applied to the logarithm of the SINR. Hence, we define the following coverage probability function

$$\tilde{p}_c(\tau_{\mathrm{dB}}) := p_c(10^{\tau_{\mathrm{dB}}/10}), \tag{3}$$

where $\tau_{\mathrm{dB}}$ is the SINR threshold in dB. Our extensive simulations results led to the following proposition.

**Proposition** *The coverage probability of the typical user in a random wireless network can be accurately described by the following parameterised sigmoid-like function*

$$\tilde{p}_c(\tau_{\mathrm{dB}}) \approx \frac{1}{1 + \exp(-\beta_p \tau_{\mathrm{dB}}^p - \cdots - \beta_1 \tau_{\mathrm{dB}} - \beta_0)}, \tag{4}$$

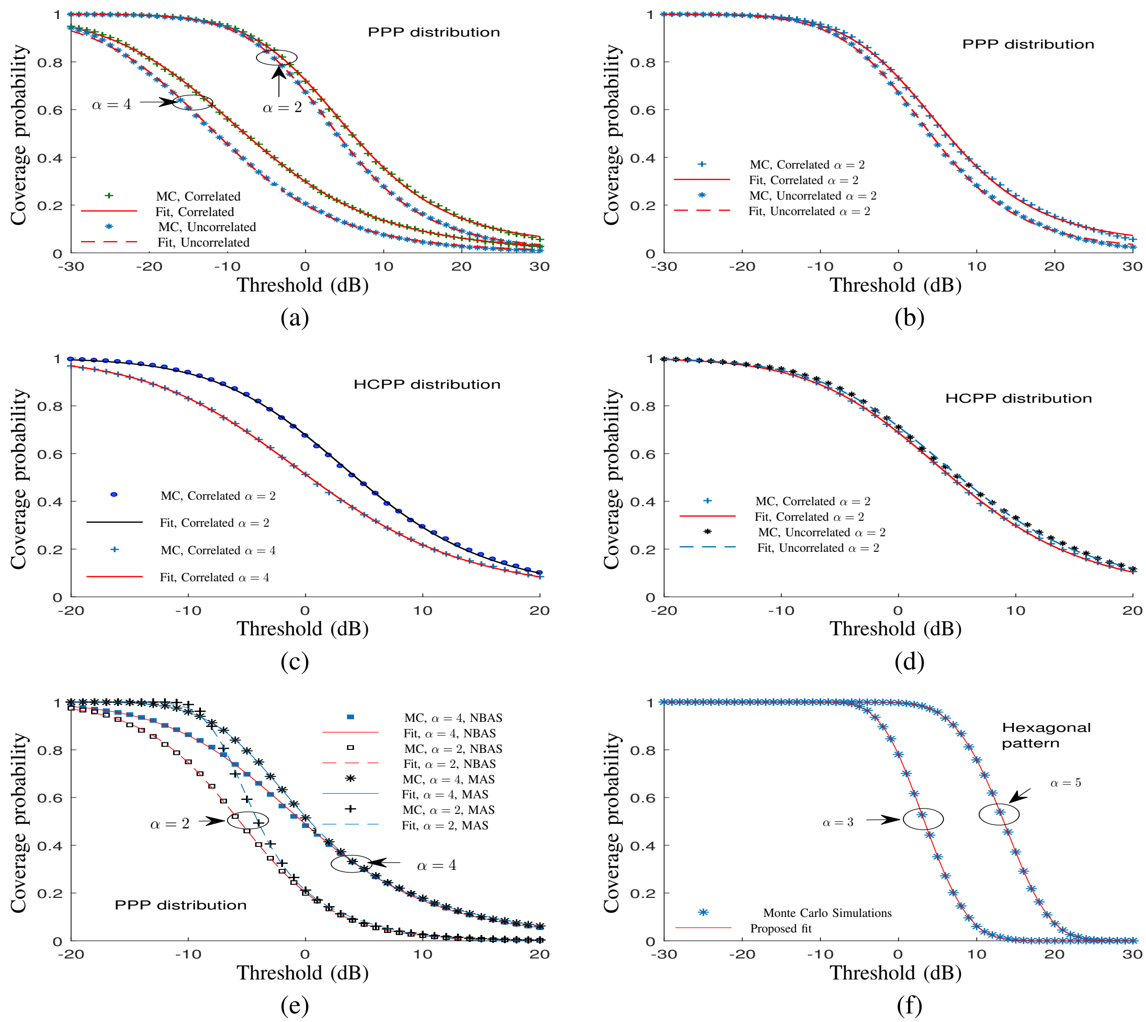


Fig. 2: (a) Coverage probability for correlated and uncorrelated shadowing and PPP base stations locations, with NBAS, $\sigma^2 = -100$ dB, $P = 1$ mw, $\lambda = 1.2 * 10^{-6}$ BS per $m^2$, $m = 2$, $d_{\text{corr}} = 150$ m, $\sigma_s = 50$ (b) Coverage probability for correlated and uncorrelated shadowing and PPP base stations locations, with NBAS, $\sigma^2 = 0$ mW, $P = 1$ mw, $\lambda = 1.2 * 10^{-6}$ BS per $m^2$, $m = 2$, $d_{\text{corr}} = 150$ m, $\sigma_s = 50$ (c) Coverage probability for correlated shadowing and HCPP base stations locations, with NBAS, $\sigma^2 = -100$ dB, $P = 1$ mw, $\lambda = 1.2 * 10^{-6}$ BS per $m^2$, $R_c = 200$ m, $m = 2$, $d_{\text{corr}} = 150$ m, $\sigma_s = 50$(d) Coverage probability for correlated and uncorrelated shadowing and HCPP base stations locations, with NBAS, $\sigma^2 = 0$ mW, $P = 1$ mW, $\lambda = 1.2 * 10^{-6}$ BS per $m^2$, $R_c = 200$ m, $m = 2$, $d_{\text{corr}} = 150$ m, $\sigma_s = 50$ (e) Coverage probability for correlated shadowing and PPP base stations locations, with different BS-user associations scheme, $\sigma^2 = -100$ dB, $P = 1$ mw, $\lambda = 1.2 * 10^{-6}$ BS per $m^2$, $m = 2$, $d_{\text{corr}} = 150$ m, $\sigma_s = 50$ ,(f) coverage probability for 256 BS and correlated shadowing, with NBAS, $\sigma^2 = -100$ dB, $P = 1$ mW, $m = 2$, $d_{\text{corr}} = 150$ m, $\sigma_s = 50$.

*where the value of vector* $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_p)$, *which is obtained through curve-fitting, depends on the network parameters.* □

In our simulations, the maximum value of the degree of the polynomial in the above sigmoid-like function was $p = 3$. More precisely, $p = 2$ was sufficient when the user association scheme was based on the nearest BS, and $p = 3$ when it was based on SINR maximization.

As seen from the examples described in Fig. 2, a very good fit can be provided using the sigmoid-like approximation. Using 500000 network setups, i.e. randomly selecting network parameters from typical value intervals, and millions of network realizations, the average value of R-square of the sigmoid-like modeling is around 0.98.

The result in the above proposition allows to predict the coverage probability for a typical user for a given network setup but to determine the values of the sigmoid-like function parameter vector $\boldsymbol{\beta}$, one would still have to run Monte Carlo simulations. In the next section, we will build large datasets consisting of network parameter values and the corresponding estimated values of $\boldsymbol{\beta}$, and use machine learning techniques to learn the relationship between the network parameters and $\boldsymbol{\beta}$. This would allow to predict the coverage probability for any network setup without running Monte Carlo simulations, as is the case when stochastic geometry-based closed-form expressions of coverage are available.

## IV. Learning the sigmoid-like model parameters

In order to determine the curve-fitting parameters $\boldsymbol{\beta} = (\beta_p, \ldots, \beta_1, \beta_0)$, we design and implement a machine learning system that applies a feed forward neural network to estimate the fitting parameters.A NN model is built for each of the spatial models of the wireless network (PPP, HCPP and hexagonal). For each spatial model, we first build large dataset consisting of network features, namely path loss exponent, base stations density, transmit power-to-noise ratio $\gamma = P/\sigma^2$, shadowing variance, and correlation distance, and radius of the guard zone in the case of HCPP, and the corresponding estimates of $\boldsymbol{\beta}$. We denote the network parameters by the vector $\boldsymbol{\theta}$, which is given by $\boldsymbol{\theta} = (\alpha, \lambda, m, \gamma, \sigma_s^2, d_{\text{corr}})$ in the case of PPP and hexagonal grid networks, and $\boldsymbol{\theta} = (\alpha, \lambda, R_c, m, \gamma, \sigma_s^2, d_{\text{corr}})$ in the case of HCPP. The input features of the neural network are the network parameters and the output features are the elements of parameter vector $\boldsymbol{\beta}$; see Fig. 3. The proposed NN is composed of $M$ layers with $N_m$ being the number of neurones of the $m$th layer. .

### A. Dataset Construction

In order to train our NN, we need to build a dataset. For this, we run a large number of simulations with various network setups, i.e different values of $\boldsymbol{\theta}$. For each scenario, we compute, using Monte Carlo (MC) simulations, the corresponding coverage for the typical values of the SINR threshold $\tau_{\text{dB}} \in [-30, 30]$ dB. By fitting the model described in equation (4) to the obtained coverage results, we collect the vector $\boldsymbol{\beta}$ that matches the studied scenario. By the end of this iterative process, we obtain the desired dataset. Table. I gives an idea about how our dataset is structured in the case of PPP networks; values of the network features are randomly drawn from the the typical intervals shwon in Table. II. Our dataset is constructed using Matlab and Simulink curve-fitting tools.

The generated dataset, denoted by $\mathcal{S}$, consists of

$$\mathcal{S} = \{(\boldsymbol{\theta}^{(1)}, \boldsymbol{\beta}^{(1)}), \ldots, (\boldsymbol{\theta}^{(n)}, \boldsymbol{\beta}^{(n)})\}, \tag{5}$$

where $\boldsymbol{\theta}^{(j)}$ and $\boldsymbol{\beta}^{(j)}$ denote respectively the $j$-th input and $j$-th output feature vectors, and $n$ is the size of the dataset. The dataset is randomly split between a training set, which represents 80% of the entire dataset, and test set.

### B. Cost function

The cost function allows to tune the NN parameters, denoted by vector $\boldsymbol{\psi}$, in order to obtain the best matching between the actual and predicted outputs. We choose this to be the conventional mean square error (MSE), given by the following expression

$$J(\boldsymbol{\psi}) = \sum_{k \in \mathcal{N}_T} \|\boldsymbol{\beta}^{(k)} - \boldsymbol{f}(\boldsymbol{\theta}^{(k)}; \boldsymbol{\psi})\|^2. \tag{6}$$

where $\boldsymbol{f}(\boldsymbol{\theta}^{(k)}; \boldsymbol{\psi})$ is the output of the NN to input $\boldsymbol{\theta}^{(k)}$ and thus the estimate of $\boldsymbol{\beta}^{(k)}$, and $\mathcal{N}_T$ refers to the training dataset. The optimum NN parameter vector $\boldsymbol{\psi}$ is obtained by minimizing the above cost function using a retropropagation algorithm.

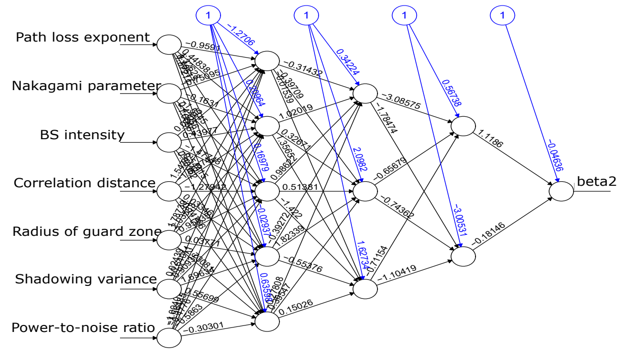


Fig. 3: Neural networks plots for $\beta_2$ estimation.

### C. Neural network model

Different neural networks are implemented using *neuralnet* package in R studio. Comparing the different NN models, we observed that a three-layer NN is sufficient to achieve very high accuracy, and having more than three layers did not significantly improve the prediction accuracy.

In the case of PPP networks, after convergence, the part of the NN which predicts $\beta_2$ is shown in Fig. 3.

### D. Accuracy results

In order to illustrate the accuracy of the NN-base modeling, we select some PPP network scenarios and depict in Figure 4 the estimated coverage (based on Monte Carlo simulations) and the NN-based coverage probability prediction. As shown in the figure, the NN-based model provides a very accurate prediction of the coverage probability.

## V. Conclusion

In this paper, we have shown, through extensive simulations, that the coverage probability can be closely approximated by a parameterised sigmoid-like function. In order to determine the parameters of the sigmoid-ike function *directly* from the network parameters (e.g. path loss exponent, BS density etc), we have proposed to use a neural network model to characterise the mapping between these two sets of parameters. As a future work, we will compare the proposed approach to other existing approximations, and in particular, those that are

| Input features | | | | | | Output vector | | |
|---|---|---|---|---|---|---|---|---|
| $\alpha$ | $\lambda(BSper\text{Km}^2)$ | $m$ | $\sigma_s^2$ | $d_{\text{cor}}(m)$ | $\gamma$ | $\beta_0$ | $\beta_1$ | $\beta_2$ |
| 2 | 32 | 2 | 4 | 30 | 10 | $-3.57$ | $-0.0763$ | $8.05 * 10^{-5}$ |
| 4 | 22 | 2 | 4 | 37 | $\infty$ | $-2.72316444$ | $-0.03$ | 0.0005 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |

TABLE I: Dataset sample for $p = 2$.

| Parameter | Typical intervals |
|---|---|
| Density of PPP | $[0.44, 0.6] * 10^{-5}$ BS per $m^2$ |
| Path loss exponent | $[2, 6]$ |
| Power-to-noise ratio | $[10, +\infty[$ |
| Nakagami parameter | $[1, 4]$ |
| Variance of shadowing | $[4, 6]$ |
| Correlation distance | $[5, 37]$m |

TABLE II: Simulation settings

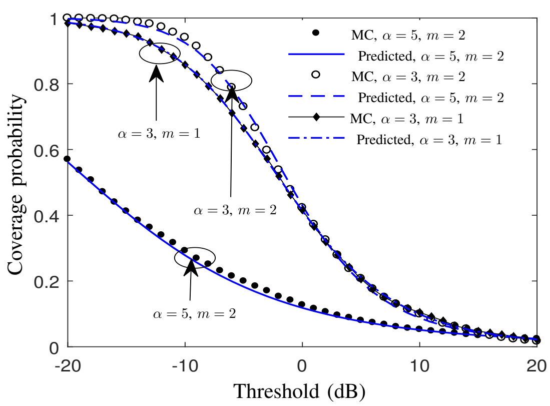


Fig. 4: Predicted coverage vs Monte Carlo simulations. For these simulations, $\lambda = 4.4 * 10^{-6}$ BS per $m^2$, $d_{\text{cor}} = 37m$, $\sigma^2 = -100$ dB, $P_b = 1$ mW.

based on the gamma approximation of the interference distribution.

## VI. Acknowledgment

This work has been partly funded by VLIR-UOS. We thank Prof. Pollin from KU Leuven for her valuable feedback on earlier versions of this paper.

## References

[1] H. ElSawy, A. Sultan-Salem, M.-S Alouini Alouini, and M. Z. Win, "Modeling and analysis of cellular networks using stochastic geometry: A tutorial," *IEEE Communications Surveys and Tutorials*, vol. 19, no. 1, pp. 167–203, 2017.

[2] J. Andrews, F. Baccelli, and R. K. Ganti, "A tractable approach to coverage and rate in cellular networks," *IEEE Transactions on Communications*, vol. 59, no. 11, pp. 3122–3134, 2011.

[3] Martin Haenggi and Radha Krishna Ganti, *Interference in large wireless networks*, Now Publishers Inc, 2009.

[4] H. ElSawy, A. 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*, vol. 15, no. 3, pp. 996–1019, 2013.

[5] H. ElSawy, H. Ekram, and M.-S Alouini Alouini, "Analytical modeling of mode selection and power control for underlay d2d communication in cellular networks," *IEEE Transactions on Communications*, vol. 62, no. 11, pp. 4147–4161, 2014.

[6] Martin Haenggi, "The meta distribution of the sir in poisson bipolar and cellular networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 4, pp. 2577–2589, 2016.

[7] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Mobile internet of things: Can uavs provide an energy-efficient mobile architecture?," in *IEEE Global Communications Conference (GLOBECOM)*, Washington DC, USA, Dec. 2016, pp. 1–6.

[8] F. Baccelli and X. Zhang, "A correlated shadowing model for urban wireless networks," in *2015 IEEE Conference on Computer Communications (INFOCOM)*, 2015, pp. 801–809.

[9] S. Szyszkowicz, F. Alaca, H. Yanikomeroglu, and J. Thompson, "Aggregate interference distribution from large wireless networks with correlated shadowing: An analytical–numerical–simulation approach," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 6, pp. 2752–2764, 2011.

[10] R. W Heath, M. Kountouris, and T. Bai, "Modeling heterogeneous network interference using poisson point processes," *IEEE Transactions on Signal Processing*, vol. 61, no. 16, pp. 4114–4126, 2013.

[11] A. Guo and M. Haenggi, "Spatial stochastic models and metrics for the structure of base stations in cellular networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 11, pp. 5800–5812, 2013.

[12] Ali Mohammad Hayajneh, Syed Ali Raza Zaidi, Des C McLernon, and Mounir Ghogho, "Performance analysis of uav enabled disaster recovery network: a stochastic geometric framework based on matern cluster processes," *IEEE Access*, 2017.

[13] Vasant Dhar, "Data science and prediction," *Communications of the ACM*, vol. 56, no. 12, pp. 64–73, 2013.

[14] S. Han, I. Chih-Lin, G. Li, S. Wang, and Q. Sun, "Big data enabled mobile network design for 5G and beyond," *IEEE Communications Magazine*, vol. 55, no. 9, pp. 150–157, 2017.

[15] M. Kulin, C. Fortuna, E. De Poorter, D. Deschrijver, and I. Moerman, "Data-driven design of intelligent wireless networks: An overview and tutorial," *Sensors*, vol. 16, no. 6, pp. 790, 2016.

[16] J. Herzen, H. Lundgren, and N. Hegde, "Learning Wi-Fi performance," in *Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, 2015, pp. 118–126.

[17] A. Samba, Y. Busnel, A. Blanc, P. Dooze, and G. Simon, "Throughput prediction in cellular networks: Experiments and preliminary results," in *1ères Rencontres Francophones sur la Conception de Protocoles, l'Évaluation de Performance et l'Expérimentation des Réseaux de Communication (CoRes 2016)*, 2016.

[18] D. T. Delaney and G. M. O'Hare, "A framework to implement IoT network performance modelling techniques for network solution selection," *Sensors*, vol. 16, no. 12, pp. 2038, 2016.

[19] Y. Lin, E. Oliveira, S. Jemaa, and S. Elayoubi, "Machine learning for predicting QoE of video streaming in mobile networks," in *IEEE International Conference on Communications (ICC)*, 2017, pp. 1–6.

[20] J. Chen, U. Yatnalli, and D. Gesbert, "Learning radio maps for UAV-aided wireless networks: A segmented regression approach," in *IEEE International Conference on Communications (ICC)*, 2017.

[21] N. Alotaibi and S. Alwakeel, "A neural network based handover management strategy for heterogeneous networks," in *International Conference on Machine Learning and Applications (ICMLA)*, 2015, pp. 1210–1214.

[22] J. Kim, J Park, J. Noh, and S. Cho, "Completely distributed power allocation using deep neural network for device to device communication underlaying LTE," *arXiv preprint arXiv:1802.02736*, 2018.

[23] 3GPP, "Study on 3D channel model for LTE (release 12)," Tech. Rep., Technical report, Technical report, Spreading and Modulation, http://www. 3gpp. org, 2017.