How Stochastic Geometry and Machine Learning Coexist in Wireless Networks: Collaboration or Competition?

Ruibo Wang[1*] **Zhengying Lou**[1*] **Lijie Hu**[1] **Di Wang**[1] **Mohamed-Slim Alouini**[1]

[1] Computer, Electrical and Mathematical Sciences and Engineering Division, King Abdullah University of Science and Technology, Thuwal, Saudia Arabia

{ruibo.wang, zhengying.lou, lijie.hu, di.wang, slim.alouini}@kaust.edu.sa

Abstract

Stochastic Geometry (SG) and Machine Learning (ML) are considered two highly effective methods for designing and evaluating the performance of next-generation large-scale wireless networks. SG is a model-driven approach that leverages previous research experience and mathematical derivations, while ML is a data-driven approach that learns from available datasets. Recently, it is indeed surprising that these two distinct methods have been frequently interacting in the field of wireless communication, coexisting through cooperation and competition. In existing studies, three types of interactions have been identified: (i) ML methods optimize the SG-based point process (PP) to bring it closer to the expected distribution; (ii) The SG framework can be utilized to analyze or accelerate the convergence of ML-based methods; (iii) ML can substitute SG in symmetric scenarios for performance evaluation and has been proven to be an evolution of SG for real-world scenarios. Furthermore, we design a novel and comprehensive case study called terrain-based coverage estimation, which encompasses all three types of interactions.

Introduction

With the exponential growth in demand for wireless communication capabilities, the performance evaluation of next-generation wireless networks has become a crucial task in the field of wireless communication. Stochastic Geometry (SG) and Machine Learning (ML) are considered the two most popular approaches for designing and analyzing wireless networks (Zappone, Di Renzo, and Debbah 2019).

SG is a powerful model-driven analytical framework. It is a mathematical tool that is particularly well-suited for analyzing large-scale random network topologies, which has recently gained significant attention. As one of the few methods that can provide analytical results for interference, SG-based approaches offer higher accuracy when analyzing the performance of interference-dominated densely deployed networks compared to other methods (Zappone, Di Renzo, and Debbah 2019).

ML is a typical data-driven approach. With sufficient real-world data, ML methods have strong generalization ability and they can learn better representations of data (Hmamouche et al. 2021). In recent years, wireless communication network-related statistics have unprecedented availability. Therefore, ML-based models have attracted more attention in various wireless communication scenarios, including power allocation, channel prediction, and coverage evaluation (Mondal et al. 2022).

Recently, the above two different methods have had frequent interactions in wireless communication. Table 1 shows the existing research involving both SG and ML methods. The contribution of the article and comparisons with existing literature are as follows:

- This is **the first summary to discuss the interaction between SG and ML.** As intuitively reflected in Table 1, the interaction between SG and ML is limited to technical research, so the scope of the content involved is far less than that of this paper.
- **The research perspective of this paper is unique** compared to the current non-technical summaries. To be specific, the existing summaries focus on the comparison between AI-based methods and model-based methods (Zappone, Di Renzo, and Debbah 2019), thus only slightly overlapping with the viewpoints of this paper.
- Currently, terrain-based network performance analysis primarily relies on optimization, model-based methods, and basic machine learning techniques, such as variants of linear regression (Wang et al. 2024). Apart from our case study, there is still no literature that **applies deep learning to terrain-based network performance analysis.**
- Furthermore, existing literature related to terrain-based network performance limited their scope to the evolution of ML as SG (Mondal et al. 2022; Wang et al. 2024), and neglected the collaboration between the two. In contrast, **our case study is more comprehensive.** Specifically, it compares the advantages and disadvantages of ML and SG through three interaction modes (C_1)-(C_3), enhancing the accuracy of performance evaluation.

ML-Aided PP Optimization

Optimization Objective

SG is a mathematical tool that enhances the tractability of traditional probabilistic wireless communication models at the cost of limiting the distribution of transmitter and receiver locations. In the SG framework, the homogeneous Poisson point process (HPPP) is one of the

*These authors contributed equally.

References	Classification	Application Scenario	Function of ML or/and SG
(Blaszczyszyn and Keeler 2019) (Saha and Dhillon 2019)	(C_1) PP thinning	Link scheduling	ML: Determinantal thinning of PPs
(Wang, Kishk, and Alouini 2022)	(C_1) PP evaluation	Satellite constellation modeling	ML: Measure the difference between two PPs
(Zappone, Di Renzo, and Debbah 2019) (Liu 2019)	(C_2) Data augmentation	Secrecy and energy efficiency analysis	SG: Generate extra data sets for the neural network
(Salehi and Hossain 2021) (Lin et al. 2021)	(C_2) Convergence analysis	Federated Learning (FL)	SG: Provide analytical expressions for convergence
(Yan et al. 2019) (Yang et al. 2019)	(C_2) Strategy selection	Content popularity prediction	SG: Select access mode and scheduling strategy
(Zappone, Di Renzo, and Debbah 2019) (El Hammouti, Ghogho, and Zaidi 2018) (Liu et al. 2022)	(C_3) Substitution	Performance evaluation	ML&SG: Coverage and QoS evaluation
(Mondal et al. 2022) (Liu et al. 2022) (Wang et al. 2024)	(C_3) Evolution	Performance evaluation	ML&SG: QoS, coverage and data rate evaluation

Table 1: Classification and summary of studies on interactions between stochastic geometry (SG) and machine learning (ML) in wireless communication. (C_1) ML-aided point process (PP) optimization, (C_2) SG-aided learning, and (C_3) competition.

most basic and common models. In HPPP, the spatial distribution of communication devices exhibits homogeneity and independence. Having homogeneity and independence, the distances between the signal transmitter and receiver follow a simple exponential distribution (Alzenad and Yanikomeroglu 2019), which allows concise expressions of performance metrics such as coverage probability and channel capacity.

However, evidence shows that the real-world distribution of communication devices often does not adhere to the assumptions of homogeneity and independence (Saha and Dhillon 2019). For example, base stations (BSs) may be distributed more densely in densely populated areas, and a certain distance between BSs needs to be maintained to avoid interference with each other. In the upper-right corner of Fig. 1, the differences between real-world BS positions and the generated HPPP are illustrated. To handle this issue, several more sophisticated PPs, such as the Poisson cluster process, Cox point process, and Matern hard-core process, have been designed to match better network topologies and features. Most of the sophisticated use HPPP as the underlying point process for generation. Among them, the determinantal point process (DPP) artfully utilizes ML methods to generate the desired PP (Blaszczyszyn and Keeler 2019).

PP Thinning

Thinning is a classic topic in the ML field, whose target is to balance the quality and diversity of selected subsets (Blaszczyszyn and Keeler 2019). Taking BS thinning as an example, BSs with better communication link quality will be prioritized for selection to constitute the DPP. On the other hand, close BSs will cause significant interference with neighboring BSs. Therefore, the diversity can be measured by Euclidean distance (Saha and Dhillon 2019).

DPP is generated through a doubly stochastic approach, where it generates an HPPP as the underlying point process and selects a subset from this HPPP to form a DPP. The figure in the upper left corner of Fig. 1 depicts the above procedure through BS selection with diversity priority. In the underlying HPPP (blue circles), the locations of some BSs are close to others, leading to significant interference for neighboring BS. The subfigure above the arrow in the middle shows the detailed thinning procedure. In the generated DPP, there is a considerable distance between each point (purple star), and the DPP better aligns with the actual BS distribution. To summarize, DPP sacrifices part of the mathematical traceability such as independence, in exchange for a more accurate modeling distribution. DPP has been demonstrated to maintain relatively good tractability and exhibit a relatively simple expression for the distance distribution between transmitters and receivers (Blaszczyszyn and Keeler 2019).

PP Evaluation

Due to several studies focusing on approximating generated PPs to real distributions, how to quantitatively measure the accuracy of the generated PP becomes an open question. The authors in (Wang, Kishk, and Alouini 2022) proposed a Wasserstein distance-inspired ML method to compare the difference between two PPs or between a PP and a deterministic points set. The squared discrete distributions' Wasserstein distance measures the minimum energy required to move one PP to another.

However, moving a PP with N points to another has N factorial ways in general, which means that to find the solution with the minimum energy, the computational complexity of the exhaustive search is N factorial. Therefore, the authors in (Wang, Kishk, and Alouini 2022) further propose an ML-based algorithm whose complexity is only N square. It is worth mentioning that results obtained from this algorithm prove that the PP has better approximation in more densely deployed networks.

Open Issues

The existing ML techniques are limited to optimizing two-dimensional PPs, but it is worthwhile to explore three-dimensional PPs for aerial and space network modeling. At present, some studies have made efforts to model non-planar

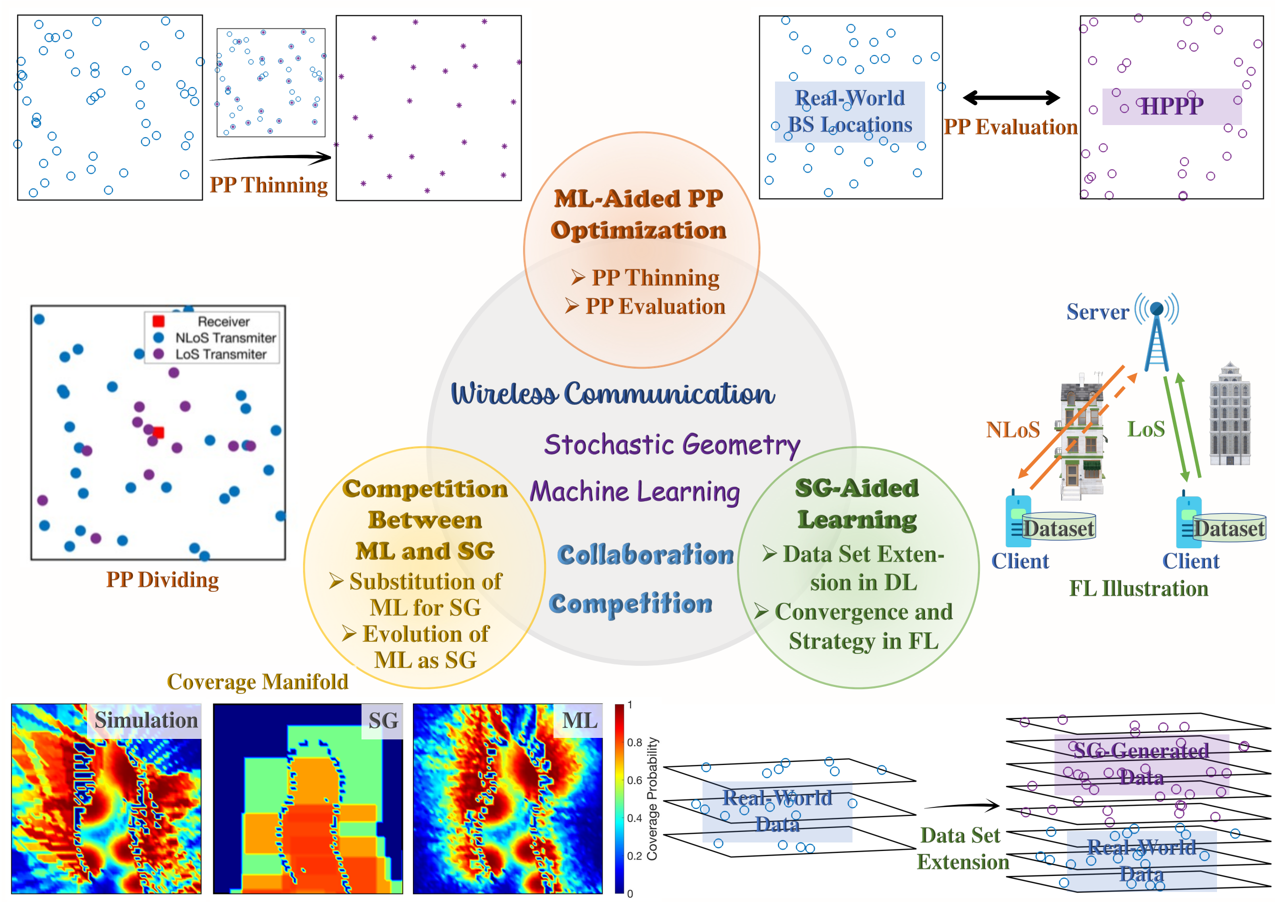

Figure 1: System diagram of the interaction between stochastic geometry (SG) and machine learning (ML). Point process (PP) thinning and evaluation are two common applications of ML-aided PP optimization; additionally, PP dividing is another promising application. SG helps federated learning (FL) by analyzing interference and SINR parameters to enhance convergence speed and link selection strategies. Next, SG-aided data augmentation reduces dependency on real data. Moreover, at the bottom left are terrain-based coverage manifolds obtained using simulation, SG, and ML approaches.

PPs under the SG framework. Spherical PPP, spherical Binomial PP, Cox PP and orbit geometry model are proposed for satellite network modeling. However, the application of ML to assist in optimizing non-planar PPs has not yet emerged.

PP dividing is another promising application. In response to the inability of HPPP to capture the effects of terrain, the authors in (Wang et al. 2024) provided an air-to-ground line-of-sight (LoS) probability model, which is implemented in the left part of Fig. 1. An HPPP can be divided into a LoS sub-PPP and a non-line-of-sight (NLoS) sub-PPP. The transmitter with smaller elevation angle to (closer to) the receiver has a higher probability of not being blocked by buildings and, consequently, establishing an LoS link with the receiver. Nevertheless, this division is closely related to the type of terrain. The counting measure of LoS PPP in a high-rise building region should be significantly smaller than that in a suburban region. Therefore, ML methods are expected to study terrain features such as the density and height of buildings and depict these features on PP dividing.

SG-Aided Learning

In the previous section, ML provides assistance in the modeling of SG frameworks. Conversely, SG can assist in generating datasets and analyzing performance for ML.

Data Augmentation

In many cases, datasets collected from the real world are often limited in scale, since obtaining real-world datasets is often costly or challenging in wireless communication. Therefore, data generated by SG's models or analytical expressions can serve as a valuable supplement (Liu 2019). An illustration of data augmentation is shown in the bottom right of Fig. 1. Since SG-generated data is not as accurate as real-world one, SG-generated data is utilized for preliminary training for the neural networks (NN). A small amount of real-world data can be used for the second round of fine-tuning (Zappone, Di Renzo, and Debbah 2019).

There are three motivations for SG-aided data augmentation. Firstly, the addition of augmented datasets reduces the heavy reliance of NN training on large amounts of accurate real-world data. Secondly, NN can converge faster in the second round of training, because the initial weights of the network are obtained after preliminary training. Third, it aims to improve the NN's adaptability and generalization ability to different samples.

Convergence and Strategy in FL

FL is a typical topic at the intersection of ML and wireless communication. An illustration of FL is shown in the right part of Fig. 1. In FL, model training is performed on local devices, called clients. Each device utilizes its local data for training and only sends the updated model parameters to a central server for aggregation (Lin et al. 2021). Client-server links are usually wireless and not stable. For example, the link might be interrupted due to building blockage, as shown in Fig. 1. Therefore, analyzing link performance is crucial can be further used for studying convergence speed and designing link selection strategies.

To start with, we explain how coverage probability is used in convergence analysis. Coverage probability, mathematically defined as the probability that SINR is greater than a threshold, represents the probability of successful communication for a wireless link (Salehi and Hossain 2021). When the SINR of the communication link does not reach the coverage threshold, the server is unable to receive the transmitted training weights from the client, thus affecting the training of the global model (Lin et al. 2021). Considering the advantages of SG in the analysis of interference and SINR-related parameters, it can provide powerful assistance to FL.

Next, the influence of coverage probability and data rate on strategy selection could be explained. The server selects a subset of the users to upload the parameters for the global model in each round. Thus, strategy selection is also a subset selection problem, and the trade-off between quality and diversity is the core (Yan et al. 2019). The client-server links with higher coverage will be assigned with high priority. However, links with low SINR should also be transmitted at a certain frequency to ensure the diversity of the training data. Therefore, authors in (Yang et al. 2019) propose three link scheduling strategies that balance quality and diversity and give the analytical convergence rate results under different strategy selections based on the SG framework.

Open Issues

In the field of wireless communication, there are also numerous applications of other learning paradigms, such as transfer learning and reinforcement learning (Zappone, Di Renzo, and Debbah 2019). SG-embedded transfer learning and reinforcement learning are still waiting to be developed. Transfer learning and reinforcement learning share similar training and test processes with classical supervised learning. Therefore, SG can also be applied as a data augmentation method. Data augmentation is particularly crucial for reinforcement learning, as it requires multiple stages of continuous data driving. Once available, SG-augmented data can accelerate the training process of reinforcement learning and transfer learning frameworks with wireless communication.

SG-aided edge distributed learning is another potential research content. Similar to FL, the communication between core network devices and edge computing devices mainly relies on wireless channels. Compared to FL, edge distributed learning focuses more on latency analysis. Fortunately, the SG framework can provide analytical results for transmission latency.

Competition Between ML and SG

Field of Competition

So far, the competition between ML and SG has been limited to the research domain of SG, that is, performance prediction of wireless networks. From the perspective of performance estimation, the analytical expressions derived from the SG framework can represent performance metrics as functions of network parameters. Similarly, ML-based methods, such as deep neural networks (DNNs), can also achieve mappings from network parameters to performance metrics. Therefore, multiple studies have compared the accuracy of both methods in performance evaluation. Furthermore, their competition includes two components: whether ML can replace SG in performance evaluation in symmetric scenarios and whether ML can provide more accurate estimates than SG in real-world scenarios. These two modes of competitions are known as complementary and evolutionary interactions between ML and SG.

Substitution of ML for SG

The motivation for substituting ML for SG is as follows. Sometimes, the SG framework is challenging to derive analytical expressions, especially in complex system models. However, ML methods can overcome this challenge by constructing parameterized models and learning the parameters from datasets. These models can be explicit expressions derived from analytical expressions obtained in previous research, or implicit representations like neural network models.

The authors in (El Hammouti, Ghogho, and Zaidi 2018) and (Liu et al. 2022) prove that neural networks can obtain performance evaluation results that coincide with SG's analytical results in symmetric scenarios. A symmetric scenario should satisfy the following three properties (Liu et al. 2022): (i) The distribution of transmitters follows some specific PP; (ii) The position of each transmitter is independent of each other; (iii) The transmitter's channel fading is independent of each other. The coverage probability curve is fitted by the Sigmoid function with a group of parameters in (El Hammouti, Ghogho, and Zaidi 2018). The parameters are estimated by training a neural network with the channel and system model parameters as input. The performance metric in (Liu et al. 2022) is estimated by some graph neural network without explicit form.

Evolution of ML as SG

The authors in (Mondal et al. 2022) present an evolution of ML as SG in real-time scenarios for performance evaluation. While there are numerous numerical simulation methods capable of achieving high-precision performance evaluation, such as Monte Carlo simulation, these methods are computationally expensive and their high complexity prevents real-time application. In contrast, once the analytical expressions of the SG framework or ML-based training are completed, both methods can obtain evaluation results with lower complexity (Liu et al. 2022). Therefore, numerical simulation results can be used to test the performance of SG and ML methods in the offline stage or generate datasets for ML training, while real-time performance analysis is better suited for ML and SG in the online stage.

The NN proposed in (Mondal et al. 2022) takes all BSs' positions as an input in the form of a manifold and maps it to a coverage manifold as an output. As a comparative method, the SG-based method converts the BSs' positions into the density of HPPP. The compression of position information and the introduction of HPPP with deviation from reality make SG coarse-grained. In contrast to the homogeneity characteristic in the SG framework, the coverage probability obtained by ML methods is location-related (Mondal et al. 2022; Wang et al. 2024). In this asymmetric scenario, ML approaches utilize position information more completely and have higher accuracy of evaluation while ensuring low complexity.

Open Issues

So far, the competition between ML and SG has mainly focused on comparing the accuracy of coverage analysis in simple scenarios. However, it is worth researching their performance in more realistic and complex scenarios, and considering other factors such as complexity.

The terrain is one of the most critical factors affecting signal transmission since signals will suffer severe attenuation when blocked by solid obstacles. The idea of coverage manifold estimation proposed in (Mondal et al. 2022) is a way to study terrain-based coverage probability. There are two potentially applicable methods to extend: incorporating the building topology information into the input matrix or adding terrain information processing layers on the basis of the original neural network.

In performance evaluation, ML and SG methods serve as alternatives to computationally expensive simulation methods. However, existing studies have not emphasized the complexity or time delay performance of these two methods. Furthermore, when considering more realistic scenarios, for example, a scenario that involves terrain information, performance evaluation becomes more complex, and complexity analysis is necessary.

Case Study: Terrain-Based Coverage Estimation

Background

The competition between ML and SG is currently conducted based on simplified system model, which somewhat affects the persuasiveness of the results. Therefore, we consider incorporating terrain information, which can lead to better predictions of system performance under different geographical conditions. The lower left part of Fig. 1 shows the terrain-based coverage manifolds at latitude 39.11° and longitude 22.33°. Datasets containing BS positions and buildings' topology can be obtained from the website www.opencellid.org and www.openstreetmap.org, respectively. In Fig. 1, simulation is realized by conducted by the ray tracing method, and the manifold of SG-based method is obtained by analytical expression provided in (Alzenad and Yanikomeroglu 2019). The next subsections will focus on how to apply ML to accomplish this task.

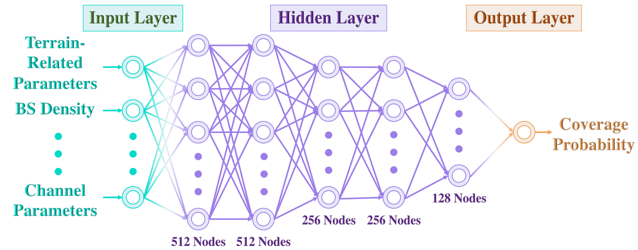

Figure 2: Deep neural network (DNN) architecture. The DNN takes the system model and channel parameters as input and the coverage probability as output.

The primary purpose of this case analysis is to help readers understand the mentioned interaction modes and extend their application scenarios. The case study is comprehensive, and it includes all three ML and SG coexisting modes. In the case study, we involve three tasks: (i) substitution and evolution interactions between ML and SG in terrain-based coverage estimation, (ii) data augmentation through SG modeling methods, and (iii) selecting more realistic data samples through PP evaluation. In addition, coverage probability is regarded as one of the most representative performance metrics. Studies have confirmed that the estimation coverage can be easily extended to the estimation of other performance metrics such as data rate and energy efficiency (Mondal et al. 2022).

DNN-Based Coverage Estimation

As a comparison method for SG, we use a DNN, whose structure is shown in Fig. 2, to estimate the coverage probability. A complete set of input parameters, their physical meanings, and values can be found in (Alzenad and Yanikomeroglu 2019). Compared with the ML-assisted model fitting method in (El Hammouti, Ghogho, and Zaidi 2018), the proposed DNN does not rely on specific models, which proves that ML methods can serve as excellent alternatives to model-driven approaches when analytical expressions are unable to provide.

Fig. 3 shows the substitution and evolution interaction between ML and SG. In this article, the number of samples used for the test process is fixed as 512. The loss on the vertical axis of Fig. 3 is defined as the absolute value of the difference between DNN's output and the expected coverage probability obtained by simulation. The "training" in the legend represents the real-time mean loss in the training process, while "validation" represents the mean loss of the trained DNN in the test process.

The brilliant blue line in Fig. 3 parallel to the horizontal axis is the loss between the label and the expected coverage probability given by SG's analytical expressions in (Alzenad and Yanikomeroglu 2019). In a symmetric scenario, we consider that the quantity of data samples is unlimited since it is relatively easy to generate data samples and labels. Under

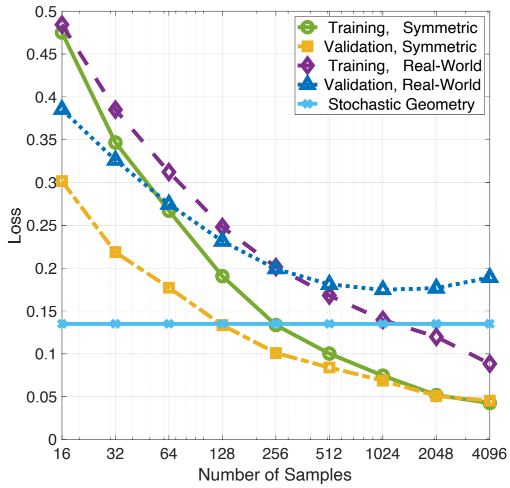

Figure 3: Substitution and evolution interaction between stochastic geometry (SG) and machine learning (ML). ML outperforms SG with sufficient data, showing the need for data augmentation.

this assumption, the estimate of coverage probability given by DNN is more accurate than SG-based method. The validation loss of the DNN converges to 4%, which is significantly lower than the loss obtained by the SG-based analytical expression (around 14 %). As long as more than 256 samples are used for training, the DNN outperforms the SG-based method in symmetric scenarios.

In real-world scenarios, the acquisition of terrain information is costly due to the limited availability of building topology in public data and significant amount of time for label acquisition. Therefore, we constructed a small training dataset with 512 samples for the real-world scenario by randomly sampling the receiving signal locations on the map. Then, we reduce loss by duplicating the training real-world data samples. As shown in Fig. 3, over-fitting occurs after more than two rounds of repeated training on the real-world dataset, and the validation loss decreases to about 17% and then rises again. In this case, the SG-based method has an advantage over the ML-based method, and data augmentation is necessary.

SG-Based Data Augmentation and ML-Based Sample Selection

The HPPP model is applied to generate locations of BSs and central locations of buildings. Based on the approaches proposed in (Wang et al. 2024) and (Alzenad and Yanikomeroglu 2019), we can compute the coverage probabilities corresponding to the generated point processes as a dataset quickly using analytical expressions under the SG framework. However, if the HPPP model is inaccurate, the training may produce some wrong initialization for DNN. A potentially feasible solution is to measure the difference between HPPP and real distributions.

We choose the Wasserstein distance to quantify the difference between HPPP and BS and the difference between HPPP and building topology (Wang, Kishk, and Alouini 2022). Each point in Fig. 4 records the Wasserstein distance between a randomly generated HPPP and a randomly chosen real-world sample at different densities. Because of the magnitude discrepancy between the building density and the BS density, the horizontal axis is unified as the normalized density. The normalized density of buildings and BSs is defined as the ratio of their density to $1241.4\mathrm{km}^{-2}$ and $167.1\mathrm{km}^{-2}$, respectively.

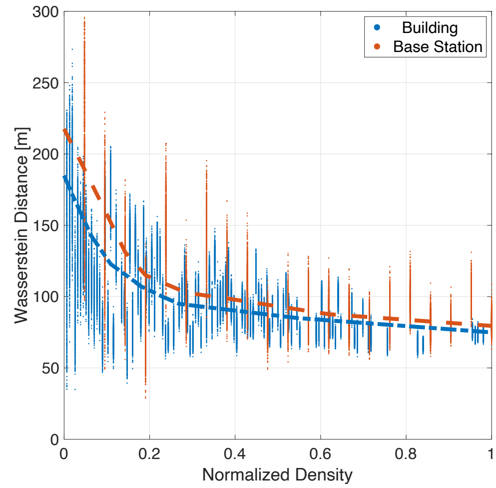

Figure 4: Measurement of differences between homogeneous Poisson point process and real-world samples. The differences decrease as density increases.

Then, we compare the coverage probability loss among the approaches of without data augmentation, augmentation, and selectivity augmentation. Selectivity augmentation represents that relatively accurate SG-generated samples below the dotted line in Fig. 4 have been selected as the augmented data, whereas augmentation means SG-generated samples are not selected. To balance accuracy and diversity, we keep the following criteria when mixing the augmented data with real-world data:

- Real-world samples are prioritized in training;
- At least half of the samples in the mixed dataset are from real-world samples;
- When the number of real-world samples is less than half, we duplicate real-world samples until they reach half.

We generate 10,000 samples by HPPP and SG analytical framework and select the most accurate 5,000 samples. Therefore, the number of samples available for data augmentation is sufficient, and there is no need for replication.

When training with more than 512 samples, the DNN trained with augmented data shows significant advantages. As shown in Fig. 5, the minimum validation loss decreases from 17.9% (without augmentation, 1024 samples) to 11.3% (augmentation, 2048 samples) and 7.2% (selectively augmentation, 2048 samples). However, the augmented dataset generated by the SG method more or less has some differences from the real dataset. Adding too many SG-generated samples in the training process will lead to a lower proportion of real-world samples in the dataset, resulting in an in-

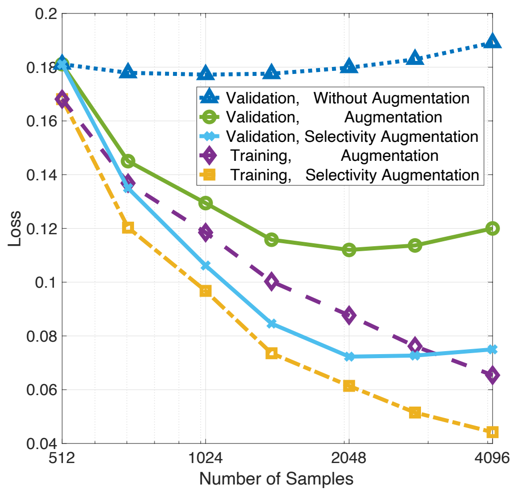

Figure 5: The influence of data augmentation on loss. Selective data augmentation greatly reduces loss, but too many stochastic geometry-generated samples can increase it due to real-world differences.

crease in validation loss. The results in Fig. 5 also demonstrate the effectiveness of the selectivity augmentation strategy, as it exhibits significantly smaller loss compared to the non-selectivity approach. Unfortunately, the model provided in (Alzenad and Yanikomeroglu 2019) is no longer aligned with the PP thinning situation and the analytical framework of this thinned PP and other sophisticated PPs (such as Poisson cluster process, Matern hard-core process) for terrain-based coverage estimation are still pending development, thus SG is not applied as a baseline for ML-based methods in Fig. 5.

Conclusion

SG and ML coexist in the network layer of wireless networks. The assistance of ML to SG is primarily focused on optimizing the distribution model of PPs, while the subsequent derivation of analytical results have not yet been addressed. SG assists NN training through data augmentation and provides performance analysis FL framework. The competition between ML and SG is limited to their common goal: performance evaluation. In the case study, we compared the accuracy of DNN and SG-based methods for terrain-based coverage probability estimation. At the same time, we verified that applying data augmentation through SG to improve DNN's performance is feasible.

References

Alzenad, M.; and Yanikomeroglu, H. 2019. Coverage and rate analysis for vertical heterogeneous networks (VHetNets). *IEEE Transactions on Wireless Communications*, 18(12): 5643–5657.

Blaszczyszyn, B.; and Keeler, H. P. 2019. Determinantal thinning of point processes with network learning applications. In *IEEE Wireless Communications and Networking Conference (WCNC)*, 1–8.

El Hammouti, H.; Ghogho, M.; and Zaidi, S. A. R. 2018. A machine learning approach to predicting coverage in random wireless networks. In *IEEE Globecom Workshops (GC Wkshps)*, 1–6.

Hmamouche, Y.; Benjillali, M.; Saoudi, S.; Yanikomeroglu, H.; and Di Renzo, M. 2021. New trends in stochastic geometry for wireless networks: A tutorial and survey. *Proceedings of the IEEE*, 109(7): 1200–1252.

Lin, Z.; Li, X.; Lau, V. K.; Gong, Y.; and Huang, K. 2021. Spatial Convergence of Federated Learning in Large-Scale Cellular Networks. In *IEEE 22nd International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 231–235.

Liu, X. 2019. Secrecy analysis with stochastic geometry: closed-form solution and machine learning. In *IEEE International Conference on Communications (ICC)*, 1–6.

Liu, Y.; She, C.; Zhong, Y.; Hardjawana, W.; Zheng, F.-C.; and Vucetic, B. 2022. Interference-Limited Ultra-Reliable and Low-Latency Communications: Graph Neural Networks or Stochastic Geometry? *available online:https://arxiv.org/pdf/2207.06918.pdf*.

Mondal, W. U.; Mankar, P. D.; Das, G.; Aggarwal, V.; and Ukkusuri, S. V. 2022. Deep Learning-Based Coverage and Rate Manifold Estimation in Cellular Networks. *IEEE Transactions on Cognitive Communications and Networking*, 8(4): 1706–1715.

Saha, C.; and Dhillon, H. S. 2019. Machine learning meets stochastic geometry: Determinantal subset selection for wireless networks. In *IEEE Global Communications Conference (GLOBECOM)*, 1–6.

Salehi, M.; and Hossain, E. 2021. Federated learning in unreliable and resource-constrained cellular wireless networks. *IEEE Transactions on Communications*, 69(8): 5136–5151.

Wang, R.; Kishk, M. A.; and Alouini, M.-S. 2022. Evaluating the Accuracy of Stochastic Geometry Based Models for LEO Satellite Networks Analysis. *IEEE Communications Letters*, 26(10): 2440–2444.

Wang, R.; Mondal, W. U.; Kishk, M. A.; Aggarwal, V.; and Alouini, M.-S. 2024. Terrain-Based Coverage Manifold Estimation: Machine Learning, Stochastic Geometry, or Simulation? *IEEE Open Journal of the Communications Society*, 5: 633–648.

Yan, S.; Qi, L.; Zhou, Y.; Peng, M.; and Rahman, G. S. 2019. Joint user access mode selection and content popularity prediction in non-orthogonal multiple access-based F-RANs. *IEEE Transactions on Communications*, 68(1): 654–666.

Yang, H. H.; Liu, Z.; Quek, T. Q.; and Poor, H. V. 2019. Scheduling policies for federated learning in wireless networks. *IEEE transactions on communications*, 68(1): 317–333.

Zappone, A.; Di Renzo, M.; and Debbah, M. 2019. Wireless networks design in the era of deep learning: Model-based, AI-based, or both? *IEEE Transactions on Communications*, 67(10): 7331–7376.