ISBN: 978-981-94-1156-6
2024 the 14th International Workshop on Computer Science and Engineering (WCSE 2024)
doi: 10.18178/wcse.2024.06.054

Throughput Prediction of Dense WLAN Deployments Using Graph Attention Networks

Koushik P R, Neeraja S, Nidhishree U R, Likitha S and Rajasekar Mohan [+]

PES University, India

Abstract. With the increasing number of wireless devices in crowded Wi-Fi networks, interference has become a major issue, causing a drop in network speed. In response, this paper introduces an innovative approach that doesn't just predict but accurately forecasts network speed in dense Wi-Fi environments, using the power of Machine Learning techniques. Our focus is on the advanced IEEE 802.11ac/ax Wi-Fi operating at 5 GHz. Despite its improvements in speed and channel width, interference from overlapping Wi-Fi networks remains a persistent challenge. Our main goal is to provide the wireless community with a groundbreaking predictive model that can anticipate network speed before deployment begins. But we're not just stopping at prediction; our approach aims to optimize deployment settings to achieve the best possible network speed in crowded Wi-Fi environments. In essence, our research is like a guiding light, showing the way to improved performance and efficiency in Wi-Fi networks, even as the number of wireless devices continues to grow, and everyone seeks faster connections.

Keywords: Throughput, Dense WLAN, Machine Learning, IEEE 802.11ac/ax, Graph Attention Networks

1. Introduction

The rapid increase in the number of wireless devices has given rise to an insatiable demand for high-speed wireless connectivity, due to which the prevalence of dense WLAN (Wireless Local Area Network) deployments has surged. Dense deployments are characterized by a multitude of interconnected devices within a confined space and have ushered in a new era of connectivity but not without their own set of formidable challenges. Among these challenges, interference from neighbouring WLANs casts a looming shadow over the efficiency and performance of these networks, ultimately leading to a noticeable dip in overall throughput. This escalating challenge demands innovative solutions that can proactively navigate the complex web of interference factors and optimize the utilization of available resources. In response to this pressing need, we introduce a Machine Learning-based approach, a cutting-edge paradigm in wireless network optimization, to not only predict, but also systematically enhance throughput in the context of dense WLAN environments. Our approach harnesses the strength of cutting-edge machine learning algorithms and data-driven intelligence to orchestrate the complex flow of data packets in a dense wireless environment. By harnessing the capabilities of Machine Learning, we aim to mitigate interference and fine-tune deployment parameters for the optimum allocation of bandwidth, ensuring that each device in the network receives the best possible throughput. In this paper, we delve into the intricacies of this innovative approach, shedding light on the methodology, data generation, pre-processing, feature engineering, and the utilization of Graph Attention Networks (GAT). Through our research, we aim to contribute to the evolution of wireless networking in densely populated areas, where the thirst for uninterrupted high-speed connectivity continues to grow unabated.

2. Related Works

2.1. Channel Bonding

Research has been done to optimize the WLAN performances in scenarios where there is a need to allocate multiple channels for efficient transfer of data. The following papers are focused on the concept of channel bonding, a crucial aspect of IEEE 802.11ac/ax networks that aims to aggregate multiple channels for higher data rates. The study [7] evaluates the performance of channel bonding against frame aggregation and aggregation without bonding. It can be inferred that channel bonding outperforms both these techniques with

[+] Corresponding author. Tel.: 91 94806 13013.
E-mail address: rajasekarmohan@pes.edu.

the increase in distance between the nodes. This also highlights the challenge of availability of non-overlapping channels and provides an insight into optimizing WLAN performance. Paper [2] provides a comprehensive understanding of the factors influencing throughput in IEEE 802.11 ac/ax networks. A significant decrease has been observed in throughput in case of WLANs with greater traffic load when their secondary channels (SCH) are shared with multiple other WLANs, despite the total system load being constant. This emphasizes the intricate correlation between channel bonding and network performance of WLANs.

2.2. Network Simulators

In the paper [1], the authors highlight the importance of network simulators to replicate highly complex deployments and characterize future 5G/6G use cases. Simulators prove to be cost effective in replicating scenarios for large-scale real-world applications. Integration of simulators with machine learning (ML) assisted networks also proves to be beneficial in optimizing, and refining network performance in controlled, scalable environments. In paper [6], a study has been conducted to analyse the performance of throughput in a multi-BSS (Basic Service Set) scenario, using NS-3 simulator, with a particular focus on BSS Colouring, a novel spatial reuse technique. It also provides an insight into the variation of throughput in a realistic multi-BSS scenario with respect to the variation in various WLAN parameters, such as Modulation and Coding Scheme (MCS) index, Channel bandwidth, Number of Antennas, to name a few. The paper [5] is based on Komondor, a wireless network simulator built on the Comp C^{++} COST library. It is designed to simulate the characteristics of next-generation high-density WLANs, with a particular focus on the 802.11ax standard, with its main features being its ability to simulate techniques such as Distributed Coordination Function, Dynamic Channel Bonding, Modulation and Coding Scheme Selection and RTS-CTS interactions. The incorporation of ML-based agents makes it superior to other network simulators.

2.3. Machine Learning Techniques

An exhaustive survey covering more than 250 recent ML-based solutions aimed at enhancing various aspects of IEEE 802.11 performance, has been conducted in paper [3], which highlights the usage of Deep Learning (DL) methods, along with Reinforcement Learning (RL) and Supervised Learning (SL) to optimize network performance. The study is highly valuable for its broad scope, offering a panoramic view of how ML can be applied to improve Wi-Fi networks across multiple dimensions.

The paper [8] mainly focuses on the concept of Dynamic Channel Bonding (DCB) and its relevance in the context of densely deployed WLANs. To overcome the effect of Overlapping Basic Service Sets (OBSS) scenarios, the study underscores the role of machine learning in predicting and optimizing throughput in high-density WLAN deployments. Ultimately, it can be inferred that the best approach to deal with these deployments is by representing the data in the form of graphs.

The most preliminary machine learning model that introduces graphs is GNN, which is used for representing and analyzing data by leveraging graph structures to capture intricate relationships between data points. The paper [4] provides a foundation for understanding GNNs and highlights their potential applications in various domains where data relationships are graph-like in nature.

One of the notable techniques evolved from GNNs is the GCNs, which incorporates local clustering coefficients and attention mechanisms, which is referred to in paper [10]. These techniques are used to capture localized patterns and dependencies between nodes and assign different weights to them according to their significance in classification. This adaptability makes the model context-aware, leading to an accurate and efficient classification. This paper is vital in offering insights and techniques to boost GCN performance, and thereby, to the field of graph-based machine learning.

3. Methodology

3.1. Understanding the Parameters

The throughput obtained in a particular area depends on many factors, some of which have a greater impact on the performance. We have experimented with different values for various channel bonding models, central frequency, number of APs and the dimensions of the area in which the APs and STAs are deployed.

3.1.1. Channel Bonding

Channel bonding in WLANs boosts data rates by combining multiple channels. This offers increased bandwidth but comes at the cost of being more prone to interference from neighbouring networks. To maximize the benefits of channel bonding, it is vital to choose a proper channel bonding scheme. A typical average throughput performance is compared in Fig. 1. for different map dimensions.

1. **Only Primary:** Only the primary channel is used for transmission (Model 0)
2. **Always-Max:** Maximum number of channels available is used for transmission (Model 3)
3. **Always-Max log2:** Logarithmic consideration of channels used for bonding (Model 4)
4. **Always-Max (MCS):** Consideration of Modulation and Coding Scheme along with channels available for transmission (Model 5)
5. **Probabilistic-Uniform:** Uses a probability-based approach towards channels bonded (Model 6)

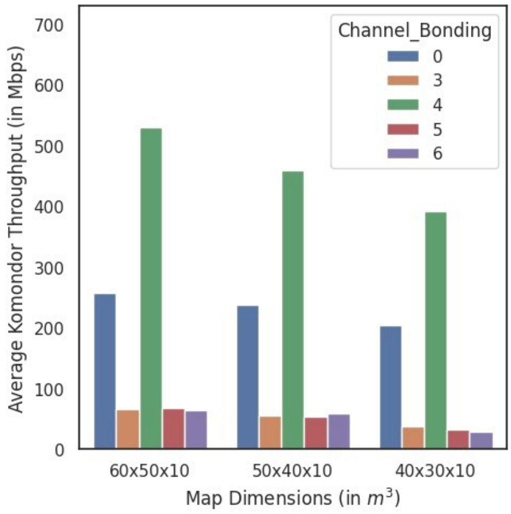

Fig. 1: Variation of Throughput with Respect to Channel Bonding

3.1.2. Number of AP and STA

Within a particular area, the number of APs and STAs deployed have a huge impact on the performance of densely deployed WLANs. Assuming the extreme case scenario where data transmission is constant and the number of associated STAs to every AP randomized, we have considered 8, 10, 12 APs deployed in three different volumes to assess network performance. The variation of average throughput in Mbps with respect to increasing values of number of access points (APs) in three different map dimensions is shown in Fig. 2.

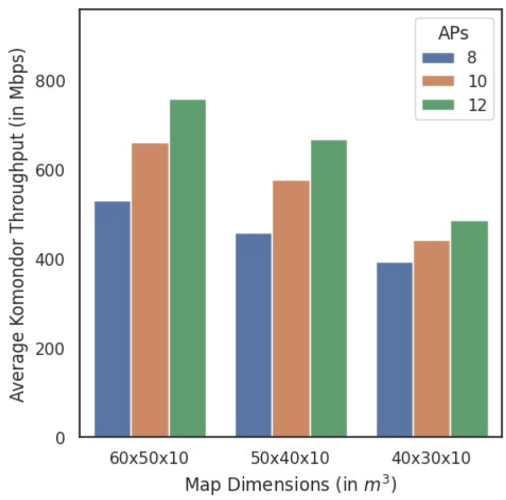

Fig. 2: Variation of Throughput with Respect to Number of Access Points

3.1.3. Central Frequency

The emergence of dual-band transmission in Wi-Fi has greatly improved the quality and performance of wireless networks. Dual band devices can operate on both 2.4 GHz and 5 GHz bands, enabling adaptability to different networking scenarios. The 5 GHz band is used for high performance applications due to larger number of non-overlapping channels, while the 2.4 GHz band can be used for longer coverage. The performance variation for different carrier frequencies is not significant as seen in Fig. 3.

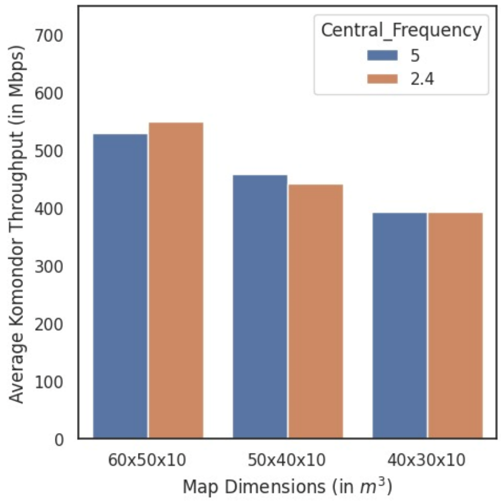

Fig. 3: Variation of Throughput with Respect to Central Frequency

However, it's important to note that the introduction of dual-band capabilities is a significant advancement in new Wi-Fi features. While it may not directly correlate with throughput in deployments, it plays a crucial role in influencing other parameters that can have an indirect impact on network performance.

3.2. Dataset

3.2.1. Dataset Generation

The dataset used in this study is generated using Komondor, an open-source wireless network simulator for simulating complex wireless network scenarios. These input files encapsulate a comprehensive array of configuration details, encompassing not only the characteristics of the Basic Service Set (BSS) but also the attributes of Access Points (APs) and Stations (STAs), in addition to essential deployment parameters. A total of 90 unique scenarios have been created by varying several parameters to create all possible combinations. This includes a combination of 5 Channel Bonding Models, 2 bands of Central Frequency, 3 different number of Access Points, and 3 Network layout dimensions. The simulator generates output files containing data on throughput per Station (STA), Received Signal Strength Indicator (RSSI), interference maps, airtime, and Signal-to-Interference-plus-Noise Ratio (SINR).

$$\text{Throughput}_{STA} = \frac{\text{No of data units transferred}}{\text{Transmission time}} \quad (1)$$

$$\text{SINR} = \frac{\text{Signal power}}{\text{Interference power} + \text{Noise power}} \quad (2)$$

$$\text{Throughput}_{AP} = \sum_{STAs\ in\ AP} \text{Throughput}_{STA} \quad (3)$$

$$\text{Airtime} = \left(\frac{\text{Packet Length}}{\text{Data Rate}}\right) \times \text{Transmission Time} \quad (4)$$

The features are prioritized based on their impact on the model's performance. Data cleaning is performed to address values that cannot be interpreted by the model.

3.2.2. Feature Engineering

Given the foundational basis of our models in the domain of Graph Networks, a crucial component of our research involves transforming the raw data into a format optimized for graph interpretation. Graphs, the central data structure in this context, are characterized by two fundamental elements: nodes and edges. Within this framework, the data is categorized into two primary types of features: Node Features and Edge Features, each of which plays a distinct yet complementary role in our graph-based modelling approach.

Node Features: Node type provides categorization and differentiation among individual nodes, allowing us to discern the roles and characteristics of entities within the network. The spatial coordinates, channel configurations, SINR (Signal to Interference plus Noise Ratio) and Airtime are the node features for the model.

Edge Features: Edge Type allows for the differentiation of edges, each potentially representing distinct interaction patterns or communication modes. The spatial distance, RSSI (Received Signal Strength Indicator) and interference can be characterized as the edge features for the respective nodes.

3.3. Machine Learning Algorithms

3.3.1. Graph Neural Network

Graph neural network is a type of neural network designed to process data represented in graph form, capturing the relationships and structure within the data. It can be thought as a graph where we have a set of nodes V and a set of edges E that connect these nodes. Each node may have associated features or attributes denoted as x_i, where i in V. Similarly, each edge may have attributes represented as e_{ij} for an edge connecting nodes i and j. A single GNN layer executes Message Passing, Aggregation and Update mechanism, where nodes exchange information with their neighbours to update their own representations.

3.3.2. Graph Convolution Network

A popular variant of GNNs is Graph Convolutional Networks (GCNs). In a GCN, information is propagated through the graph using convolutional operations, which allow nodes to aggregate and update their features by considering information from their neighbours. Multiple layers uncover even deeper patterns in the graph data. The learned feature representations can then be used for various graph-based tasks, such as node classification or link prediction.

3.3.3. Graph Attention Network

Graph Attention Networks (GATs) use an attention mechanism to process graphs for tasks like node classification. Unlike other methods, GATs consider the importance of neighboring nodes, leading to more accurate predictions. The attention mechanism assigns weights to connections based on node features, allowing GATs to capture complex relationships within the graph. The attention mechanism computes attention coefficients α_{ij} for each edge connecting node i to node j as follows:

$$\alpha_{ij} = \frac{\exp\left(\sigma\left(a^T[WX_i||WX_j]\right)\right)}{\sum_{k\in(i)} exp\left(\sigma\left(a^T[WX_i||WX_k]\right)\right)} \tag{5}$$

Here, W is a learnable weight matrix, || denotes concatenation, and $\mathcal{N}(i)$ represents the neighbours of node i. σ is an activation function applied elementwise like ReLU. The node representations are updated using the computed attention coefficients as follows:

$$h_i' = \sigma\left(\sum_{j\in\mathcal{N}(i)} \alpha_{ij} W h_j\right) \tag{6}$$

Here, h_i' represents the updated representation of node i after attention aggregation, and σ is an activation function. In summary, Graph Attention Networks (GAT) provide an efficient solution for processing data organized in a graph structure, and their ability to capture intricate relationships underscores their suitability for the analysis of networks like WLANs.

4. Experimental Results

4.1. Result Metrics

RMSE (Root Mean Square Error) gauges a model's prediction accuracy. It reflects the average magnitude of the difference between predicted and actual values. Lower RMSE signifies better model fit, as predictions closely align with observations. Conversely, higher RMSE indicates larger discrepancies, suggesting poorer model performance.

4.2. RMSE Curves for Tuning of Hyper Parameters

4.2.1. RMSE Curves for Varying Learning Rates

The learning rate in machine learning determines the step size at which a model updates its parameters during training. An appropriate learning rate can lead to faster convergence, while a poorly chosen one may result in overshooting or slow convergence, affecting model training and performance. In this paper, we conduct a performance comparison of the Graph Attention Network (GAT) model by systematically varying learning rates, aiming to identify the optimal learning rate. The RMSE values for different epochs for varying learning rates from 0.005 to 1 are shown in Fig. 4.

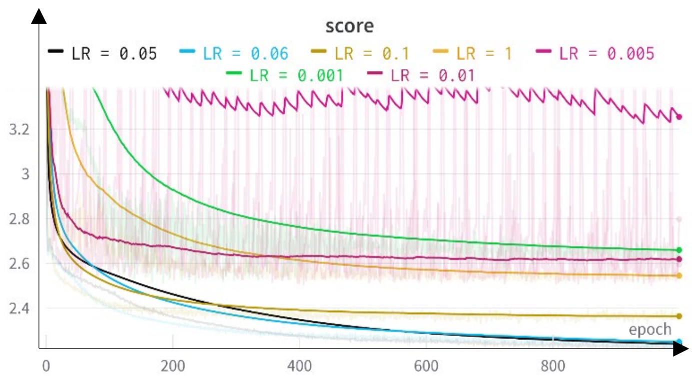

Fig. 4: RMSE Curve for Various Learning Rates

4.2.2. RMSE Curves for Varying Number of Hidden Layers

The number of hidden layers in a Graph Attention Network (GAT) significantly impacts its performance. Hence, selecting an appropriate depth is critical for achieving the desired trade-off between expressiveness and training stability. In this paper, we conduct a comprehensive performance analysis of GAT model by varying the number of hidden layers for evaluation. Fig 5. gives the RMSE variations for different numbers of hidden layers

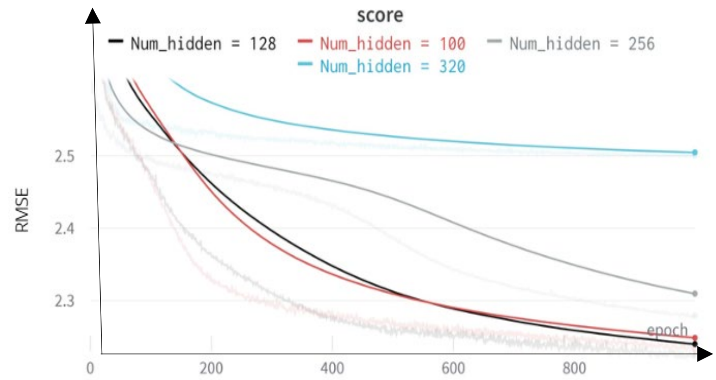

Fig. 5: RMSE Curves for Various Number of Hidden Layers

4.2.3. RMSE Curves for Varying Activation Functions

The Activation functions are critical components in neural networks, responsible for introducing non-linearity and enabling models to learn complex relationships. Here are some common activation functions we have included ReLU (Rectified Linear Unit), Leaky ReLU, ELU (Exponential Linear Unit), SELU (Scaled Exponential Linear Unit) and Mish. Fig 6. Gives the RMSE variations (across epochs) for different activation functions.

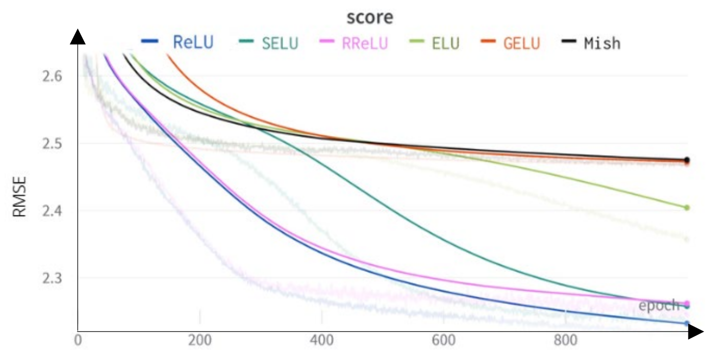

Fig. 6: RMSE Curves for Various Activation Functions

4.2.4. RMSE Curves for Varying Optimizers

The choice of optimizer in a Graph Attention Network (GAT) plays a vital role in model convergence and performance. Different optimizers can significantly impact training dynamics, affecting the network's ability to find optimal solutions and its overall performance in various graph-based tasks. Some optimizers used in our model include RMSProp, Adagrad, Adam, AdamW, SGD and NAdam. The RMSE variations over several epochs for different optimizers are depicted in Fig 7.

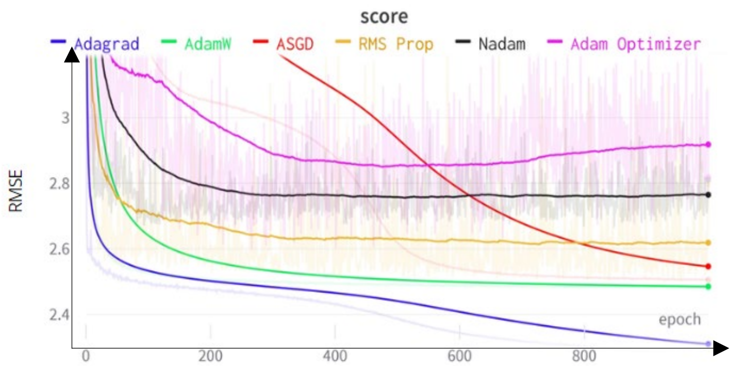

Fig. 7: RMSE Curves for Various Optimizers

4.2.5. RMSE Curves for Varying Weight Decays

The Weight decay, also known as L2 regularization, is a common technique employed in training deep learning models, including Graph Attention Networks (GAT). It serves as a regularization method to remove overfitting and improve the generalization ability of models. The primary objective of weight initialization is to mitigate the issues of layer activation outputs either exploding or vanishing during the forward propagation process. Here are the results of different weight decays as shown in Fig 8.

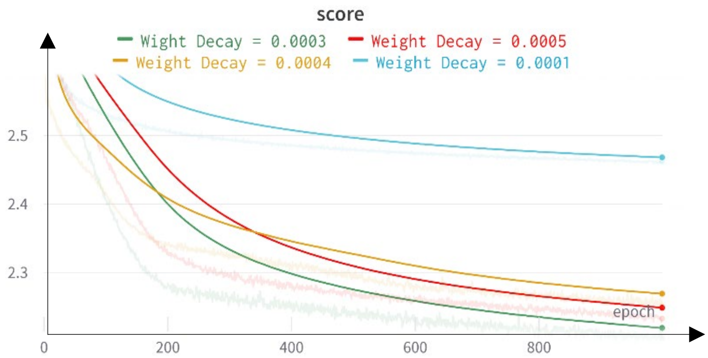

Fig. 8: RMSE Curves for Various Weight Decay Values

4.2.6. RMSE Curves for Different ML Models

Fig 9. illustrates the improvement, showcasing our model's lower RMSE curve, which directly translates to enhanced accuracy.

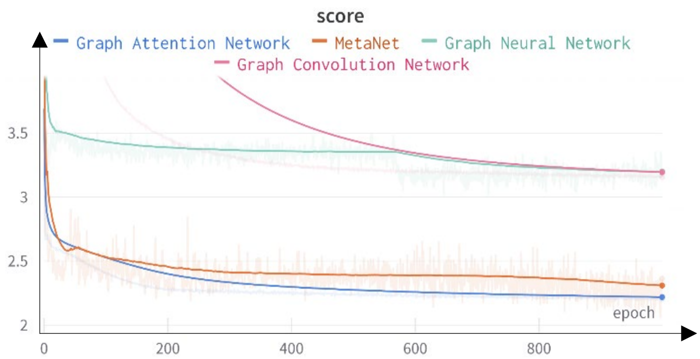

Fig. 9: RMSE Curves for Model Comparison

4.3. RMSE and Throughput Analysis of GAT Network

The accuracy of throughput prediction by the GAT model is determined by taking the simulator throughput as the reference. The average value of RMSE obtained by the best selection of hyper parameter values is estimated to be 2.203. This implies that, for a predicted throughput value, the actual throughput obtained after deployment of the WLAN lies within the range of (predicted throughput - 2.203, predicted throughput + 2.203) Mbps.

4.4. Scenario-wise Consolidated Analysis

The results of throughput predicted by the model for 90 scenarios has been consolidated below with the average throughput, range of throughput predicted, average and range of RMSE obtained for all the 18 main scenarios. The RMSE obtained for the prediction of throughput for all scenarios lies between 1.14 to 12.66.

Table 1: Consolidated result of Graph Attention Model's Throughput prediction and Accuracy

Scenario	Avg. Throughput (Range) (Mbps)	Avg. RMSE (Range) (Mbps)	Percentage Error (Range) (%)
Sce 3	306.41 (190.07 – 970.17)	2.41 (2.18 – 9.35)	0.786%
Sce 4	296.70 (190.94 – 970.53)	2.46 (2.38 – 9.99)	0.829%
Sce 5	245.88 (112.86 – 915.39)	2.64 (1.71 – 10.81)	1.073%
Sce 6	282.84	3.25	1.149%

	(112.87 – 960.89)	(1.88 – 12.66)	
Sce 7	130.08 (89.14 – 575.57)	1.89 (1.70 – 4.70)	1.444%
Sce 8	176.12 (88.01 – 719.71)	2.46 (2.23 – 6.05)	1.396%
Sce 9	198.62 (136.71 – 785.69)	2.42 (1.51 – 10.78)	1.218%
Sce 10	196.32 (120.66 – 700.51)	2.35 (2.03 – 10.39)	1.197%
Sce 11	248.61 (179.16 – 948.05)	2.41 (1.94 – 8.79)	0.969%
Sce 12	250.80 (151.34 – 887.41)	2.45 (2.14 – 6.05)	0.976%
Sce 13	175.67 (98.06 – 665.87)	2.10 (1.89 – 7.56)	1.194%
Sce 14	161.45 (84.40 – 670.86)	2.09 (2.50 – 6.52)	1.291%
Sce 15	253.82 (137.89 – 855.30)	2.12 (1.79 – 7.31)	0.835%
Sce 16	241.47 (153.13 – 876.00)	2.11 (1.92 – 7.43)	0.873%
Sce 17	154.19 (99.28 – 686.23)	1.92 (1.77 – 6.31)	1.247%
Sce 18	153.94 (62.01 – 623.06)	1.72 (1.58 – 5.85)	1.117%
Sce 19	166.52 (66.31 – 669.63)	1.63 (1.22 – 5.68)	0.978%
Sce 20	167.53 (78.94 – 666.17)	1.51 (1.14 – 5.48)	0.901%
	Mean Percentage Error		**1.0333%**

5. Ablative Experimentation

Ablative studies involve deliberately removing or deactivating specific components, variables, or factors within a system to evaluate their individual impact on overall performance or behavior. The accompanying table provides a comprehensive overview of the model's performance under various parameters. Interestingly, the Channel Bonding consistently yields similar results across different configurations (0, 3, 4, 5, and 6), suggesting that the choice of channel bonding has minimal effect on the model's behavior.

Table 2: Comparison of RMSE Scores between ablative datasets and General model

Parameters	**Values**	**Train loss**	**Val loss**	**RMSE**
General	All	1.487	1.353	2.193
Channel Bonding	0	1.566	1.478	2.226
	3	1.566	1.478	2.214
	4	1.566	1.478	2.214
	5	1.570	1.478	2.210
	6	1.567	1.474	2.194
Central Frequency	2.4 GHz	1.572	1.477	2.204
	5 GHz	1.570	1.480	2.234

No of APs	8	1.477	1.352	2.186
	10	1.446	1.338	2.275
	12	1.479	1.349	2.043
Area of Deployment	40x30x10 m^3	1.222	1.074	1.721
	50x40x10 m^3	1.463	1.336	2.218
	60x50x10 m^3	1.721	1.626	2.486

The results demonstrate using 10 access points (APs) significantly improves model performance. Compared to 8 and 12 APs, the 10-AP configuration yields lower training and validation losses. This suggests that smaller deployment areas lead to better performance, highlighting the model's strength in confined spaces. These findings offer valuable insights for optimizing future deployments.

6. Acknowledgement

We would like to express our heartfelt gratitude to Punith S, Department of Electronics and Communication Engineering, PES University, Bangalore for his invaluable guidance and support during the course of the work.

7. References

[1] F. Wilhelmi, M. Carrascosa, C. Cano, A. Jonsson, V. Ram and B. Bellalta, “Usage of Network Simulators in Machine-Learning-Assisted 5G/6G Networks,” in IEEE Wireless Communications, vol. 28, no. 1, pp. 160-166, February 2021, doi: 10.1109/MWC.001.2000206.

[2] M. -S. Kim, T. Ropitault, S. Lee and N. Golmie, “A Throughput Study for Channel Bonding in IEEE 802.11ac Networks,” in IEEE Communications Letters, vol. 21, no. 12, pp. 2682-2685, Dec. 2017, doi: 10.1109/LCOMM.2017.2747511.

[3] S. Szott et al., “Wi-Fi Meets ML: A Survey on Improving IEEE 802.11 Performance With Machine Learning,” in IEEE Communications Surveys and Tutorials, vol. 24, no. 3, pp. 1843-1893, third quarter 2022, doi: 10.1109/COMST.2022.3179242.

[4] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Mon- fardini, “The Graph Neural Network Model,” in IEEE Transactions on Neural Networks, vol. 20, no. 1, pp. 61-80, Jan. 2009, doi: 10.1109/TNN.2008.2005605.

[5] S. Barrachina-Mun˜oz, F. Wilhelmi, I. Selinis and B. Bellalta, “Komondor: a Wireless Network Simulator for Next-Generation High- Density WLANs,” 2019 Wireless Days (WD), 2019, pp. 1-8, doi: 10.1109/WD.2019.8734225

[6] Performance study of High-Efficiency IEEE 802.11ax WLAN Standard using NS-3 simulator by Aparna Behara, T.G. Venkatesh, in 2022 IEEE 27th International Workshop on Computer aided modelling and Design of Communication Links and Networks (CAMAD)

[7] Vivek Parashar.; “Aggregation-Based Dynamic Channel Bonding to Maximize the Performance of Wireless Local Area Networks (WLAN)” in Wireless Communications and Mobile Computing Volume 2022, Article ID 4464447

[8] R. Mohan, K. V. Ramnan and J. Manikandan, “Machine Learning approaches for predicting Throughput of Very High and EX-treme High Throughput WLANs in dense deployments,” 2021 12th International Conference on Computing Communication and Net- working Technologies (ICCCNT), 2021, pp. 1-6, doi: 10.1109/ICC-CNT51525.2021.9580133

[9] Throughput Analysis of Dense-Deployed WLANs Using Machine Learning by Divya C V; Drishya Sriranjini B;Pavan Venkatesh Naik; Rajasekar Mohan, 2023 International Conference on Advances in Electronics, Communication, Computing and Intelligent Information Systems (ICAECIS)

[10] GCN with Clustering Coefficients and Attention Module Rakesh Kumar Yadav; Abhishek Abhishek; Sourav S; Shekhar Verma 2020 19th IEEE International Conference on Machine Learning and Applications (ICMLA)