Graph-based Deep Learning with Attention for Spatial Reuse Optimization in Dense WLAN Deployments

Stephen Azeez, *Student Member, IEEE*, Shagufta Henna, *Staff Member, IEEE*

***Abstract*—While IEEE802.11ax is designed to support self-configuration and adaptation functionality in dense deployment in order to enhance dynamic network conditions, effective transmission power modulation is a crucial concern in the design and efficiency of wireless networks. Predominantly, collisions tend to be hard to evade if nodes transmitting at low transmission power rate are not noticed by nodes transmitting at high transmission power levels. Therefore, the presence of varied transmit powers can generate asymmetrical connectivity situations, thereby resulting to throughput bottlenecks. In the absence of proper power management, co-existing IEEE 802.11ax access points will suffer co-channel interference, degrading throughput and yielding less achievement to the overall network. It is therefore essential to take into account the impact of the variable transmit power features and the neighbor nodes in the strategy to optimize the network performance. This work proposes the affinityGNN-attention to model the neighborhood variable transmit power impact on access points to be optimized through the attention mechanism to generate a more expressive framework representation of the node state. In our experiments, it was shown that attention module integration improves the prediction accuracy and robustness of the baseline affinityGNN model.**

***Index Terms*—IEEE802.11ax, network management, transmission power, interference identification, legacy, heterogenous network, throughput, multi transmission.**

I. Introduction

With the exponential growth of wireless networks, wireless access point self-management and adaptive adjusting capabilities have become increasingly important. Wireless network performance optimization is a critical requirement for intelligent wireless local area network (WLAN) devices. Nonetheless, the WLAN's tremendous deployment efficiency is also jeopardizing its potential development. Users are becoming more demanding, and the density of networks and clients is rising, thus the existing state-of-the-art in wireless technology is likely to soon fall short of supporting the ultra-dense deployment of WLAN access points (APs) and stations (STAs). Although the next-generation standard IEEE802.11ax uses a new PHY layer technology called Orthogonal Frequency Division Multiple Access (OFDMA) to enhance the performance and scalability of ultra-dense networks for a variety of transmission needs. Yet, when the amount of APs and legacy nodes increases, the ultra-dense throughput of 802.11 ax suffers.

There have been lots of discussion on how to increase performance in dense networks. The latest 802.11ax network analysis reveals that interference patterns and frequency-selective attenuation, along with medium access inefficiencies and configurations, technically affect the throughput in legacy spectrum in a dense network operation. In order to reduce network interference, high efficiency devices seeks to allow simultaneous transmissions in overlapping networks, while at the same time hoping to boost total throughput. The concept of dynamic transmission power regulation and dynamic sensitivity thresholds has generated a great deal of interest. Asymmetric links may also result in unjust conditions due to sensitivity and power control [1]. Dynamic sensitivity and transmission power modification have been found to improve network efficiency and help mitigate the effects of the hidden and exposed terminal difficulties [2].

This paper explores the effect of variable transmission power in the presence of signal sensitivity for throughput maximization in network co-existence. For instance, it is well recognized that a high transmission power influences the APs power usage, creating interference between nodes operating simultaneous frequency. In contrast, a low transmission power will adversely affect the receiver's signal-to-noise ratio. In order to ensure transmission, the power level must be high enough to guarantee transmission, while remaining low enough to save energy on the nodes. While the increased transmission power does increase the strength of the signal received, it also has a few drawbacks. An increase in transmission power may trigger the creation of new weak connections with increased signal strength that are not yet strong enough to build new reliable connections. Consequently, in order to extend network lifetimes and improve performance, each node must have transmission power regulation.

Based on the original affinityGNN approach [3], this work will leverage an attention based on weighted graph framework to learn interactions in co-existing networks. The contribution of this work is proposing a attention-based affinityGNN model to leverage the presence of interference due to variable transmission power at each access points compared with the initial affinityGNN model. With attention, the weights for all the input weighted summations are redistributed, which improves the accuracy of prediction and ensures the best utilization of resources. Particularly, an affinityGNN-attention model is proposed, which could outperform the base affinityGNN model

The authors are with the Faculty of Engineering and Technology at the Letterkenny Institute of Technology, Ireland (emails: azeezsteven@gmail.com, shagufta.henna@lyit.ie)

Manuscript received April, 2022; revised April 2022.

operating over such a use case.

A. Related Works

Numerous efforts have been made to evaluate network throughput in the presence of variable transmission power for the purpose of identifying and minimizing interference to improve network throughput [4] [5] [6]. The advantages of power management in lowering co-channel interference (CCI) levels have been widely studied [2] [7]. Krunz et. al [2] illustrates power control techniques that consider the MAC concept in its architecture. These techniques encompass a group of algorithms that typically leverage transmission power to regulate the network's topological features which employs interference information to limit the power rates of successive transmissions. As detailed in [3] [8] [9], merely enabling wireless APs to transmit at varying power rates in a distributed coordinate function will certainly raise hidden terminals, resulting in additional collisions and re-transmissions attributable to the contention-based access scheme. Conversely, in ultra-dense, using varied transmit powers for every access points might result in unbalanced links, which can create throughput bottlenecks [10]. To mitigate interference, Oteri et. al [11] devised a transmission power distribution approach that functions alongside inter-cell coordination. Rather than focusing solely on transmission power control to prevent interference, the utilization of coordinated transmit power modulation to improve performance and fairness in ultra-dense wireless networks. ElBatt et al. [12] envisaged a power distribution framework to optimize transmission rate, which is parallel to the data transmission power control. Gomez et al. [13] studied the effects of variable transmission power on the physical and network interconnected nodes in multihop wireless network. As regards power usage and connectivity, their research found that variable-range transmission power outperforms prevalent transmission power. As deep learning and AI techniques continue to advance, there has been a clear trend in using deep neural networks to solve wireless prediction challenges [14]- [22]. In recent years, attention-based approaches have seen significant advancements and have demonstrated effectiveness in wireless network application, driving us to apply a graph-attention-based method to network co-existence collision problem. The Graph Attention QNetwork (GAQ) was introduced in [23] to evaluate the effect of the neighboring node on the node be maximized using the attention ramework. it also provides a lowdimensional embedding vector with greater expressive capacity to characterize the to be optimized node state in the RL model. In this paper [24], a model named LA-ResNet is developed that handles spatio-temporal analysis and predicts wireless network traffic using an attention mechanism. The temporal and spatial aspects of wireless network traffic data are modeled, and associated attributes are enhanced, using this method. Wireless network traffic data can thus be successfully recorded in terms of spatio-temporal correlation. A graph attention framework, ST-GAT [25], to capture the spatial relationships using LSTM network for extracting temporal domain elements. The suggested approach, in comparison to earlier relevant studies, is capable of capturing dynamic spatial relationships of traffic networks. The reviews of existing literatures on interference and network collision solutions that focus on transmit power control have revealed that there is an extensive lack of a ubiquitous solution to the dynamic transmit power problem; thus, improvements are possible and necessary in this area. As a result, this article outlines an encompassing solution where collision is detected such that an interacting AP is informed of the clients signal strength and assesses the transmission power levels influence.

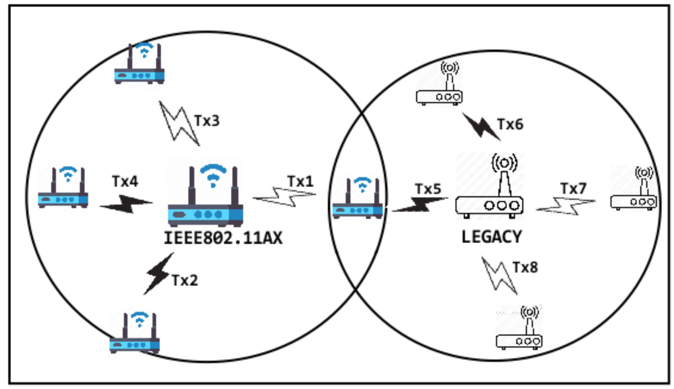


Fig. 1. Figure showing coexistence of modern high-efficiency (HE) device, IEEE802.11ax, and legacy with the presence of variable transmission power

B. Motivation

The coexistence problem due to interference is discussed in this paper, and it is based on earlier work in the literature [3]. The Affinity model is good at learning networks interaction, however it doesn't take into account the weighted properties of the to be optimized AP neighbors when it comes to improving the overall performance of a wireless network with the effect of transmission power variability. As a result, considering the influencing factors are employed as features to maximize individual IEEE802.11axAPs without taking into account the influence of their neighboring APs, the wireless network throughput will be suboptimal.

Consider a Wireless heterogeneous Network G on N access points with E as the signal's interference edge of various signals comprising of the varying weighted power that occur in the network. As the weights is attached to each edge of the signal distance at a certain distance d, there exist a channel assignment for a channel Q of an edge E denoted by (E, Q). The channel assignment is said to be interfered by a set of assigned channels P, if there is a channel assignment (E, Q) $\in$ P such that edges are within the interference range at distance d with variable power control for some channel Q with W(.) being the function of assigned edges in G represented by W(P).

II. System Model

For our network model, the power that an access point receives during uplink multi-user transmissions have variations across all stations, hence, importance estimating it impacts on wireless communication. Individual APs determines the weighted transmission power exclusively based on the received signal strength indication of the linked access points containing network information in the basic service set, given a transmission based on distance. In conformity with these distance, APs closer to the access points use less power, resulting in fewer exposed nodes, whereas APs further away consume the most power. When IEEE802.11ax and legacy devices coexist, signal interference is depicted as a geometric weighted graph, with the edges reflecting all the signals received. Each interference edge will be assigned a weight that corresponds to the amount of power received and to account for the weight impact of an interference signal from IEEE802.11ax and legacy APs.

A. Transmission Power

The amount of energy spent is known to be significantly determined by the transmission power. A very high transmission power impacts the transmitter's power usage, causing interference between nodes running on the same frequency. Conversely, in a too low transmission power, the receiver's signal-to-noise ratio is impacted [26]. Hence, transmission power regulation must be implemented within every node to enhance the network span. Suppose we take P_{ij} as the minimal power required for a transmission interaction of a certain AP_i to reach AP_j in a dense network. The maximum distance to which another AP can be positioned for a single transmission from AP to reach AP_j is the transmission range of AP_i. Transmissions sent by each node over its communication range will be received by all other nodes in the communication range. The transmissions sent by AP_j will have power $P_1 = P_{ij}$, while AP_j and AP_k will both have power $P_2=\max(P_{ij},P_{ik})$. The nodes can adjust the power of its transmission between zero and the maximum transmission power level. As AP_i varies its transmission power, P_λ, based on the distance between the nodes, its transmission power will be as follows:

$$P_\lambda(i,k) = P_\lambda(i,j) + P_\lambda(j,k) \tag{1}$$

Based on distance and edge features, the optimal transmission power, P_0, required for transmitting from i to j with n number of nodes is given as:

$$P_0^n = \sum_{i,k=1}^{n} P_\lambda(i,k) \tag{2}$$

B. Distance, Coverage and Connectivity

We evaluated an unvarying random distribution of AP placements in a poison process distribution such that each AP interacts with its closest AP (if it is in its coverage area specified by the receiving device's sensitivity), which is usually the AP with the highest receiving power. Also, simply because to capacity constraints, more access points are required in high-density deployments. To avoid co-channel interference, lower transmit power levels and careful channel planning are required [27] [28]. The signal intensity in a wireless network varies inversely with the distance between two nodes. In a uniformly distributed deployment, the reference communication distance is given by

$$\eta(d) = 2\pi\rho d e^{-\pi\rho d^2} \tag{3}$$

where Ψ symbolize the locations of AP, ρ as the density of APs, and d denoting communication distance.

C. Overlapping Channels

During the transmission, IEEE802.ax-APs receives signals from numerous sources simultaneously, including legacy devices. Consequently, the data received becomes a mash-up of the signals from numerous sources. According to equation(1)&(3), if these access points are uniformly distributed then a transmitting IEEE802.11ax access point could detect a neighboring legacy AP on a channel of an overlapping basic service set (BSS). In an association-based BSS, all APs use the same channel and have an equal chance of accessing it. Interfering transmitters are defined as active nodes in un-associated BSSs for each AP in the association [3] [29]. Multiple overlapping channels must be analyzed more pragmatically by considering both pathloss attenuation and Rayleigh fading. The received power from AP_i at an AP_n is obtained by considering both pathloss attenuation and rayleigh fading.

$$P_{(d)} = TP_i \times L \times Q^n{}_{X_i} ||n - x_i||^{-\alpha} \tag{4}$$

The fading coefficient from the nth AP and ith AP is represented by $Q^n{}_X$, with TP_i depicting AP_i transmission power, L is the propagation loss (1m distance), α denoting the path loss exponent, including function $||.||$ as the geometrical distance function. With the propagation of the carrier sense threshold and transmission power being a constant, the transmission power of AP_i is given by

$Q^n{}_X$ represents the fading coefficient from the nth AP and ith AP, TP_i represents AP_i transmission power and L is the propagation loss (1m distance), which corresponds to the path loss exponent, as well as the function $||.||$ as the geometrical distance function. Given a constant propagation of the carrier sense threshold and transmission power, APi's transmission power is given by

$$TP_i = \frac{TP_0 \times \omega}{\gamma} \quad \forall i \tag{5}$$

such that ω symbolizes the initial carrier sense threshold of every APs, TP_0 denoting the initial transmission power of all APs and γ, as the carrier sense threshold function of a communicating device.

D. Signal-to-Noise/Interference Ratio

SINR threshold is obtained from propagating the probability and the AP density to determine whether a given AP is able to receive a signal, and factors in both the likelihood and density of the achievable transmission. To calculate the SIR for a wireless device, we consider both the power levels of interferences defined in an interference layer's edge weights, as well as the strength of the desired signal. There exist two criterion for a certain AP_j has two conditions, as in equation (6), to be threshold covered by the given AP_j for a station

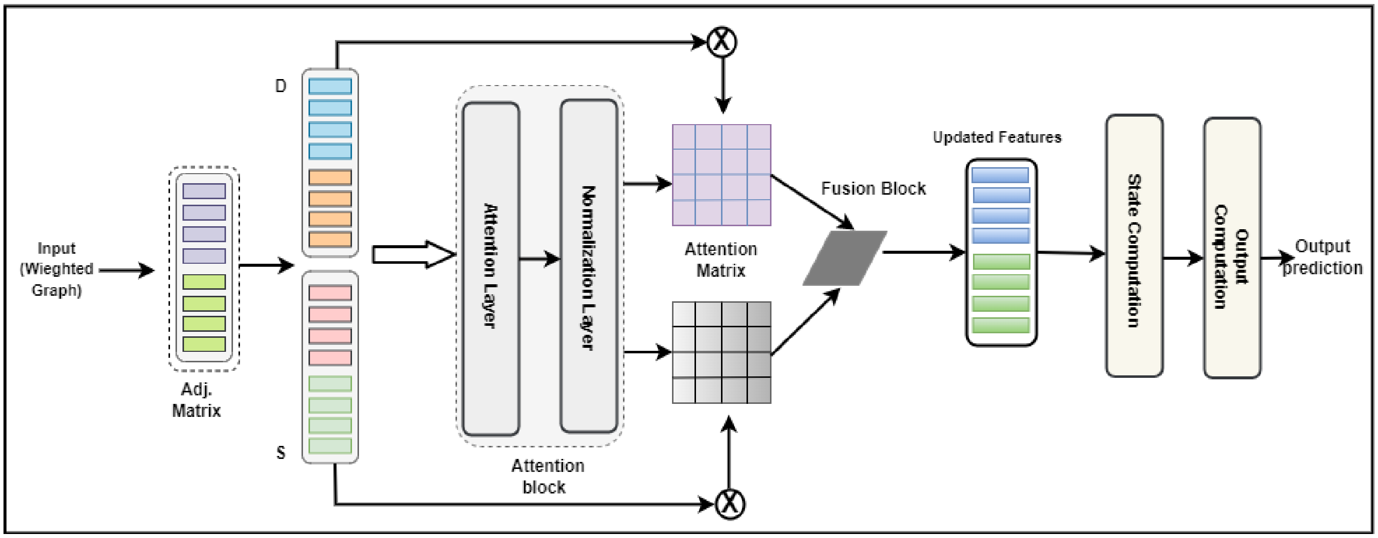


Fig. 2. In the architecture, we first feed the weight graph as input into the model thereby constructing skip affinity from the weighted adjacency matrix in which vector nodes is obtained foe direct-affinity (D), and skip-affinity (S). From each embedding vectors, the attention mechanism block adaptively captures the learned 'rich' representation between nodes. The extracted attention correlation coefficients matrix is concatenated with the initial node features for (D,S) then fused to obtain the updated node features to be fed into AfinityGNN to get the final output prediction.

positioned at a given point q. First, the signal power must exceed a threshold to be detected [3]. As for the second criterion, it is required that the SINR at q is greater than the threshold. Based on the value of θ, modulation technique and instantaneous rate can be determined

$$\text{SINR} = \begin{cases} PQ_sL(q) > P_0 \\ \frac{PQ_sL(q)}{\sum PQ_sL(q)} > \theta \end{cases} \tag{6}$$

Q_s is the shadowing/fading attainmet between the independent locations, P is the signal power, theta denoting the detection threshold and L representing the path loss at point q

III. GRAPH MODEL

The procedure for generating the network with edge weight is briefly described in this section. The graph theory principles and measures were used to design the weighted wireless network. Edges are undirected in this scenario as they reflect the bidirectional relationship between APs. To define and model the wireless network's interacting relationships, an undirected graph $G = (V, E, \omega)$ is employed. The vertex set for the n-th AP vertices is V={v_1, v_2,..., x_N}. The edge array $E\in$(V×V), where (v_i, v_j) means that node i and node j interacts and ω is array of edge weights. Not only are the influences of neighbors on the focusing node taken into account in this study, the value of the focusing node to its neighbors is also examined. As a result, we believe that the mutual contribution of the two APs is not equal and that it may be learned. Therefore, this study explored the weighted graph phenomenon.

Considering that the goal is to assign weights to the interference edges in the channel assignment so that each assigned weight indicates the power variable due distance of the received signal strength indicator, with the goal of identifying the interference dynamics of the channel assignment represented by each edge. To compute edge weights in measure terms of distance d, the RSSI measurements (in dBm) must be converted into a metric of distance linking the access points i and j

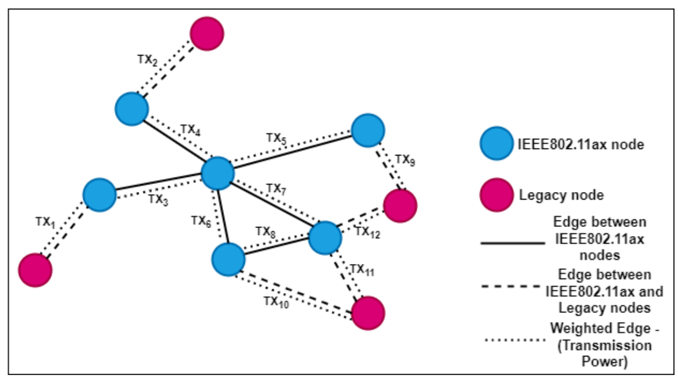


Fig. 3. An illustration of the graph model G

$$TX_{P(i,j)} = 10^{-RSSI_{i,j}/10} \tag{7}$$

Where $TX_{P(i,j)}$ corresponds to the resulting importance between node 'i' features to node 'j' in the network graph G.

IV. PROPOSED MODEL

This model employs an attention based aggregating mechanism on AffinityGNN to capture relational importance features with their attention signals on a static IEEE802.11ax communication network. AffinityGNN model equipped with an attention mechanism can provide access points -specific attention weights on its interaction features. The graph structure with its weights is incorporated the into attention mechanisms through AffinityGNN embedding which leverage the node2vec approach to learn the vertex representations. The

AffinityGNN embedding only provides static representations of IEE802.11ax network interactions, which could not represent the dynamic IEE802.11ax interactions among access points in the network.

There is an explicit attention mechanism included in AffinityGNN, which is based on GNN. Iteratively using features of every node for affinity calculation, attention-AffinityGNN learns the hidden features of each node using self-attention. A standard convolution in direct-affinityGNN and skip-affinityGNN contains the standardized gated fusion aggregation of the features of adjacent nodes in equation (8) & (9) respectively[3]

$$D^{l+1} = \sigma \sum (MD^{(l)} W^{(l)}{}_0, M_S S^{(l)} W'^{l}) \tag{8}$$

$$S^{l+1} = \sigma \sum (M_S S^{(l)} W^{(l)}{}_S, MD^{(l+1)} W'^{l}) \tag{9}$$

where M = $D^{\frac{1}{2}} A' D^{\frac{1}{2}}$ and M_S = $D^{\frac{1}{2}} A'_S D^{\frac{1}{2}}$. $W^l{}_0$, $W'^{(l)}{}_0$, $W^l{}_S$ are the propagated transformed weights. σ is rectified linear unit (RELu) activation function with the purpose of adding non-linearity to the propagation.

The aggregation function in equation (8)&(9) propagates information between nodes and updates the hidden state of nodes in order to output the final embedding as the AffinityGNN can automatically extract the node transition relationships in affinity graph. In the graph convolution, attenion-AffinityGNN substitutes the above convolution operation with an attention mechanism. In order to depict the mode at which features of each nodes at l^{th} layer are updated to those of l^{th}+1 layer, the model integrates the graph attentional layer's constituting component. A collection of node features in both direct $D^l \in \mathbb{R}^F$ and skip $S^l \in \mathbb{R}^F$ affiinty is fed into a attention layer, with F denoting number of features from each node. A pooled weight matrix $\psi \in \mathbb{R}^{F* \times F}$ is utilized to project the input towards other feature space of $F*$ dimension in order to transform the input features onto higher level features. Then, to measure/scores the relational importance between the nodes i and j if there is an edge between the node points, we define them as inputs to the attention layer to capture the attention coefficient score $A_{i,j}$ in each encoded state for direct affinity (D) and skip affinity (S) is given in equation 3 and 4 rspectively

$$A^D_{i,j} = Att(\psi D^{(l)}_i, \psi D^{(l)}_j) \tag{10}$$

$$A^S_{i,j} = Att(\psi S^{(l)}_i, \psi S^{(l)}_j) \tag{11}$$

such that $Att : \mathbb{R}^F \times \mathbb{R}^{F*} \Rightarrow \mathbb{R}$ is the attention layer, and $A_{i,j}$ denotes the resulting computed attention correlation coefficient. Since just the attention coefficients of the particular node and its first-hop neighbors, i.e the direct affinity nodes and neighbors, are usually computed to keep topological structure of the graph G in equation 8, $A^S_{i,j}$ in equation 9 tends to solve this limitation by computing the attention coefficients of the skip neighbors, i.e the second-hop neighbors.

Subsequently, the mechanism of computing attention scores is employed by utilizing the softmax function to normalize the attention coefficients of node i over all neighbor nodes j, which is defined by a weight matrix $\psi \in \mathbb{R}^{F* \times F}$ with each node pair, in particular, having a self-loop that allows it to update itself.

$$\theta_{i,j} = softmax(A_{i,j}) \tag{12}$$

$$\theta^{D,S}_{i,j} = \frac{exp(A^{D,S}_{i,j})}{\sum_{K \in \mathbb{N}(i)} exp(A^{D,S}_{ik})} \tag{13}$$

where $\theta_{i,j}$ is the attention score indicating the importance of i to j in direct affinity and skip affinity. By fusing the respective self-attention measures, the feature vectors are transormed on each node

Algorithm 1 affinityGNN-attention

INPUT: AFFINITYGRAPH G = $(\mathbf{V},\mathbf{E})$
$\mathbf{A} \leftarrow$ Adjacency Matrix
$\hat{\mathbf{A}} \leftarrow Normalization(A)$
$\mathbf{D} \leftarrow Node2vec(\hat{A})$
$\mathbf{D^{(0)}} \leftarrow DW^{(0)}$
  For $each\ attention\ layer\ l, l = 1, ...\ in\ (i,j) \in G$**do**
    $\mathbf{D}^{(l)} \leftarrow \mathbf{D^{(1)}} W^{(l)}$
    COMPUTE ATTENTION VECTOR
    $\mathbf{A}^D_{i,j} \leftarrow Att(\psi D^{(l)}_i, \psi D^{(l)}_j)$
    $\theta_{i,j} \leftarrow softmax(A_{i,j})$
    LEARN INFORMATIVE FEATURE REPRESENTATIONS
    $\mathbf{D^{(1+1)}} \leftarrow D^{(l)} \theta_{i,j}$
  End
  COMPUTE OUTPUT
    $\mathbf{O} \leftarrow softmax(D^{(\ell+1)}, W^{(\ell+1)})$
  return O

$$\tilde{D} = \theta^D D, \quad \tilde{S} = \omega \theta^S S \tag{14}$$

such that ω is a weighting parameter alongside the feature-attention updates. The proposed node affinity attention component of the base model takes advantage of the relational importance between attributes vectorof both direct and skip affinity, therefore allowing multiple viewpoints for affinity node to exchange relative important information As a result, the need particular fusion of $\tilde{D}$ with $\tilde{S}$ in the presence of trainabe kernel coefficient k

$$\gamma = FusionGate(\tilde{D}, \tilde{S})^k \tag{15}$$

We apply a further fully connected feed-forward network to give the model with nonlinearity at the completion of the respective layers in the encoded state, matching the underlying structure of AffinityGNN forattention model. Point-wise feed-forward network,P_w is a variant of feedforward neural network consisting of two linear transformations with an activation function ReLU in between.

$$P_w = \sigma(\gamma W + \beta) \times W' + \beta' \tag{16}$$

Here, W, W' are the propagated weights, β, β' denoting binary inicator bias and γ being the resulting attention matrix. Finally, to determine the relative importance tendency, we employ a fully connected layer.

$$Z = softmax(W^* P_w + b^*) \tag{17}$$

where W^* and $b*$ are learnable parameters. We leverage the binary crossentropy loss over training data during the model training.

$$\mathcal{L} = \sum T_{i,j} log Z_{ij} + (1 - T_i j) log(1 - z_{pq}) \tag{18}$$

where T_{ij} indiates is the matching label indicator matrix by mini-batching during training. The Adam optimization approach is used to optimize the model.

Algorithm 2 skipAffinityGNN-attention Algorithm

INPUT: AFFINITYGRAPH G = $(\mathbf{V},\mathbf{E})$
$\mathbf{A} \leftarrow$ Adjacency matrix of network AffinityGraph $\mathbf{G}$
$\mathbf{A_D} \leftarrow Normalize(\mathbf{A})$
$\mathbf{A_S} \leftarrow Normalize(\mathbf{A}\mathbf{A}^\mathbf{T})$
$\mathbf{D} \leftarrow Node2Vec(\mathbf{A_D})$, $\mathbf{S} \leftarrow Node2Vec(\mathbf{A_S})$
 For $each\ attention\ layer\ l$,$l = 1, ...\ in\ (i,j) \in G$**do**
  $\mathbf{D}^{(l)} \leftarrow \mathbf{D}^{(\mathbf{l})} W^{(l)}$
  $\mathbf{S}^{(l)} \leftarrow \mathbf{S}^{(\mathbf{l})} W^{(l)}$
  COMPUTE ATTENTION VECTOR
  $\mathbf{A}^D_{i,j} \leftarrow Att(\psi D^{(l)}_i, \psi D^{(l)}_j)$
  $\mathbf{A}^S_{i,j} \leftarrow Att(\psi S^{(l)}_i, \psi S^{(l)}_j)$
  $\theta_{D,S} \leftarrow softmax(A_{D,S})$
  LEARN INFORMATIVE NODE FEATURE REPRESENTATIONS
  $\hat{\mathbf{D}} \leftarrow D^{(l)} \theta^D$
  $\hat{\mathbf{S}} \leftarrow W S^{(l)} \theta^S$
  $\gamma \leftarrow FusionGate(\hat{\mathbf{S}}, \hat{\mathbf{D}})_k$
  End
 POINTWISE FEEDFORWARD
 $P_w = \sigma(\gamma W + \beta) \times W' + \beta'$
 $z = softmax(W^* P_w + b^*)$
 return z

V. RESULT ANALYSIS

Here, we compare and discuss the results achieved by the proposed model against those obtained with the baseline model by demonstrating the robustness of the results obtained. For each node in the uplink transmission, the goal is tois to optimize the possible throughput performance with an airtime fair constraint, while retaining the frame error rate and reducing the delay of every nodes [25]. It can be expressed in terms of a single objective function ν

$$\nu = \sigma[\frac{throughput_\alpha \times \beta}{delay_\delta \times f}] \tag{19}$$

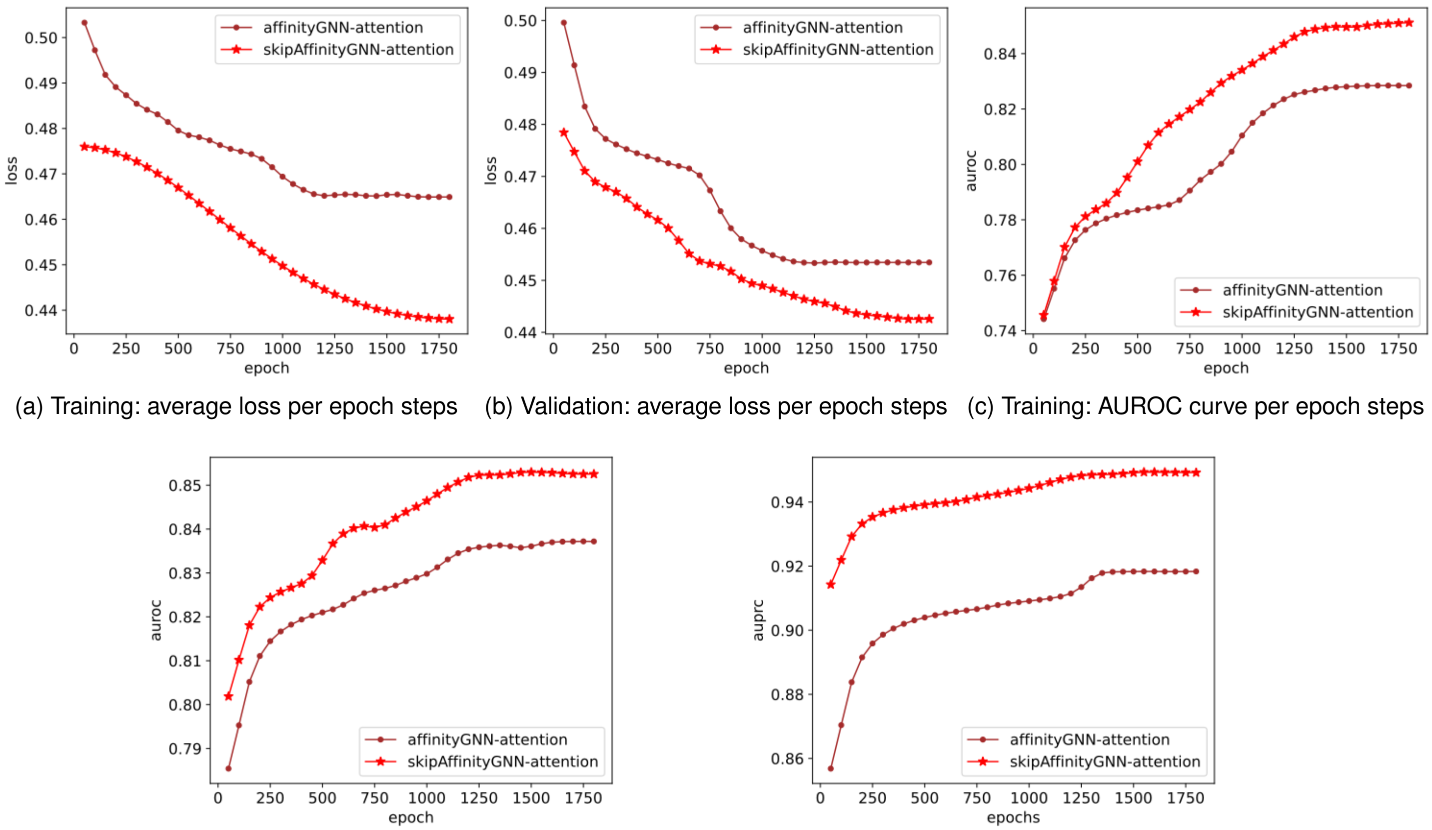
where f is the frame error rate, δ denotes average delay, α as aggregated network and β representing the fairness of the network.

A. Model Evaluation

In this setting, we first correlate the experiment's evaluation metrics attained by AffinityGNN-attention performed on a weighted IEEE 802.11ax-compliant network effected by variable transmission power operating in the 5 GHz band with skipAffinityGNN-attention model executed on the weighted network coexistence of IEEE 802.11ax network operating in the 5 GHz band and legacy devices operating in the 2.4 GHz band inherent of varying transmission power. We then validate the effectiveness and efficiency of AffinityGNN-attention by comparing it the baseline AffinityGNN model. In order to be extensible through modularity and to support the incorporation of existing information, a hyperparameter optimization framework should be created that is independent of the environment. Likewise, since evaluating hyperparameter configurations is frequently computationally intensive, flexible distribution of the computation should be enabled. To shed light on the optimization for affinityGNN-attention, The learning rate is set to $5e^4$; minibatch size of 256; the Adam optimizer. dropout rate of 0.1; and the hidden size is set to [16,64] in thefist and second layer respectively.

In our experiment, 1800 is taken as the total number of epochs in the experiment. Each epochs contains 50 steps. That is, within 3000 steps, 36 epochs will be executed. Figure 4 shows the learning curves for our proposed approach which depicts that performance level of skipAffintyGNN-attention over the just affinityGNN-attention for which across all the evaluation metrics, skip-AffinityGNN-attention outperforms AffinityGNN-attention. Fig 4a&b show lower loss values for both training and validation losses for skipAffinityGNN w.r.t to epochs compared to AffinityGNN-attention model, resulting in higher training and validation accuracies according Fig 4c,4d,&4e. The AffinityGNN-attention curve reached the convergence zone at about 950-1100 epochs across all learning cuvre. In contrast, the converging speed of skipAffinityGNN-attention is much slower, taking a much longer steps and time to reach the converging zone at about 1150 - 1250 epochs. For the skipAffinityGNN-attention model, the losses curve keep a certain distance above the AffintyGNN-attention losses explaining the larger difference between validation and training accuracies reported for this model. Referring to Fig4e for instance, the learning curve of skipAffinityGNN-attention method converges to about 94.5% AUPRC accuracy of it validation while the curve of AffinityGNN-attention increases up to 91.8%, that is, skipAffinityGNN-attention is 2.9% higher than AffinityGNN-attention over the AUPRC metric. All in all, the performance of skipAffinityGNN-attention is much better than its baseline direct AffinityGNN-attention.

Furthermore, a direct correlation of the proposed AffinityGNN-attention model with the baseline AffinityGNN model is another important criterion for evaluating the model in evaluation of the quality of the performance outcomes. As illustrated in fig. 5, the AffinityGNN model converges substantially quicker than the AffinityGNN-attention model. As such, AffinityGNN's convergence zone is reached before 250 steps, whereas the AffinityGNN-attention loss curve isn't reached till 1200 epochs, and



(a) Training: average loss per epoch steps (b) Validation: average loss per epoch steps (c) Training: AUROC curve per epoch steps

(d) Validation: AUROC curve per epoch steps (e) Validation: AUPRC curve loss per epoch steps.

Fig. 4. Experimental result showing the learning cuves of AffinityGNN-attention and skipAffinityGNN-attention

there's an noticeable downward trend between zero and 1200 epochs. AffinityGNN-attention and AffinityGNN validation PRC-AU curves converge at approximately 91.7% and 90.8%, respectively. Thus, the AffinityGNN-attention method achieves greater improvements on the performance metrics index than the baseline AffinityGNN method. Consequently, while the skipAffinityGNN curve is consistently more stable in the early on and retains its quick convergence features, the skipAffinityGNN-attention curve continues to improve over time. In terms of the PRC-AU and ROC-AU trends depicted in curve, skipAffinityGNN-attention model exhibit the highest performance among the others. According to the PRC-AU curve displayed in Figure 5e&6e, our proposed models, were successful in prediction classification with accuracy above 91% point.

B. Impact of Missing Edges

As wireless networks connectivity evolves constantly, one of the critical challenges of network analysis is classification predictions in the presence of missing edges. Fresh APs and edges are added or disconnected over time, making wireless networks extremely adaptable. Using the current topological structure of the wireless network scenario, the prediction aims to approximate the likelihood of existence between a pair of APs. Using the current topological structure of wireless networks, the prediction aims to approximate the likelihood of existence between a pair of APs. With the wireless network dynamics taken into account, we estimated the proposed model's performance as it neutralizes the missing edges during the network evolution. Our hypothesis is that the proposed model, being able to handle the dynamics of wireless networks, would be less penalized by missing edges compared to baseline algorithms.

For further analysis, we evaluate systematically the network completeness in the studied scenario by evaluating the robustness of the proposed model in examining the overall topology of the network over the missing edges as a way to evaluate each model's performance as well as its robustness. In the experiments, we randomly excluded 20 %, 40%, 60% & 80% of the network edges to evaluate each algorithm on its classification prediction performance ability of the overall network's missing edges. Based on the capability that each method, affinityGNN-attention and skipAffinityGNN-attention, has to compute the probability based on completeness of the network each, we quantify the model prediction by using AUROC and AUPRC compared to the baseline approaches. Figure 7 shows the results of our proposed approach as skipAffinityGNN-attention gives a significant boost in performance in capturing true communicating network structure compared to affinityGNN-attention in both AUROC and AUPRC score in Figure 7a&7b respectively. Relative to skipAffinityGNN

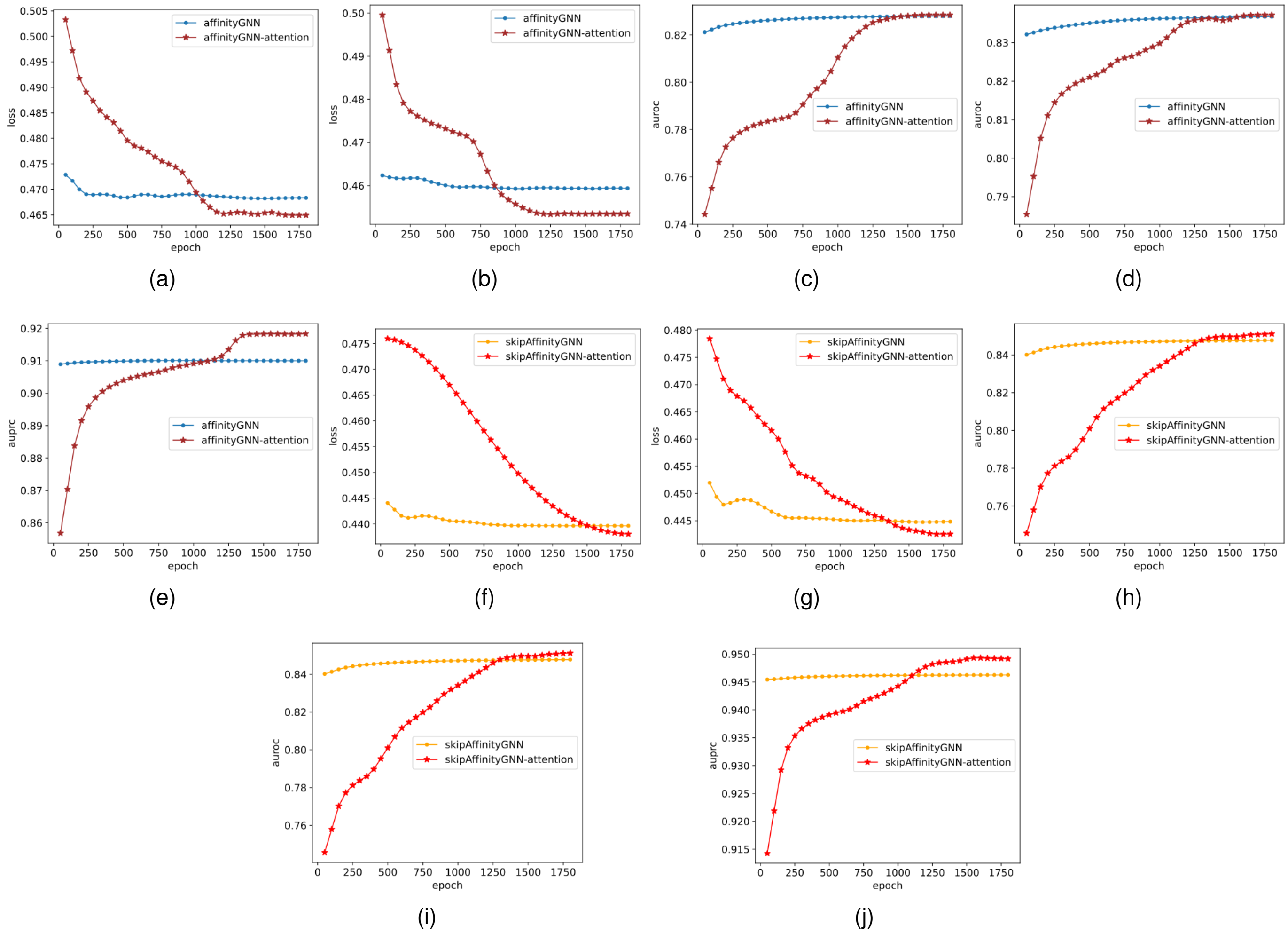


Fig. 5. Experimental result showing the learning cuves of (i)AffinityGNN-attention vs AffinityGNN of (a)Training: average loss per epoch steps; (b)Validation: average loss per epoch steps; (c)Training: AUROC curve per epoch steps; (d)Validation: AUROC curve per epoch steps (e) Validation: AUPRC curve loss per epoch steps. (ii)skipAffinityGNN-attention vs skipAffinityGNN of (f)Training: average loss per epoch steps; (g)Validation: average loss per epoch steps; (h)Training: AUROC curve per epoch steps; (i)Validation: AUROC curve per epoch steps (j) Validation: AUPRC curve loss per epoch steps

shows better performance in AUROC but a less score in AUPRC. In other respects, affinityGNN-attention performs particularly well when edges are incomplete over affinityGNN. Accross AUROC measure, affinityGNN-attention outperforms the baseline but starts to decrease for higher missing edges ratios.

In general, we find that our algorithm performs better than other algorithms for all indices considered. A direct illustration of its effectiveness is in the metric index, which indicates the probability of (correctly) assigning a higher score despite the existence of missing edges. Additionally, this explains why our algorithm is less sensitive to the densities of edges at the start: any edges observed are accounted for as soon as our weighted-based edge structures identify the network-generating process. Additionally, our comparison shows that the usage of information in the given classification prediction algorithm is an important factor in determining the goodness of the given algorithm. According to the baseline methods, the assumption that nodes establish an affinity with a probability that is proportional without regard to their weighted feature fails to capture the processes influencing the wireless network topology; by preserving the edges' transmission power information, the proposed model makes better use of available information.

C. *Impact on Throughput Performance Estimation*

The results presented in Figure 7 demonstrate that, as the network size of legacy devices increases, the overall performance continuously degrades. As a result, for an increase in diameter of the network, it cost IEEE802.11ax networks more transmission power per node interacting with legacy nodes. To compare the efficiency of the proposed method for networks of various sizes, it is concluded that the proposed attention method offers greater precision than the baseline method inherent of an optimal transmission power that maximizes the network throughput. According to figure 8, the higher access throughput indicates that APs are assigned relatively frequently, thus, increasing access frequency exerts a similar effect as increasing the number of APs. The network throughput increases with the transmission range of any affiliated

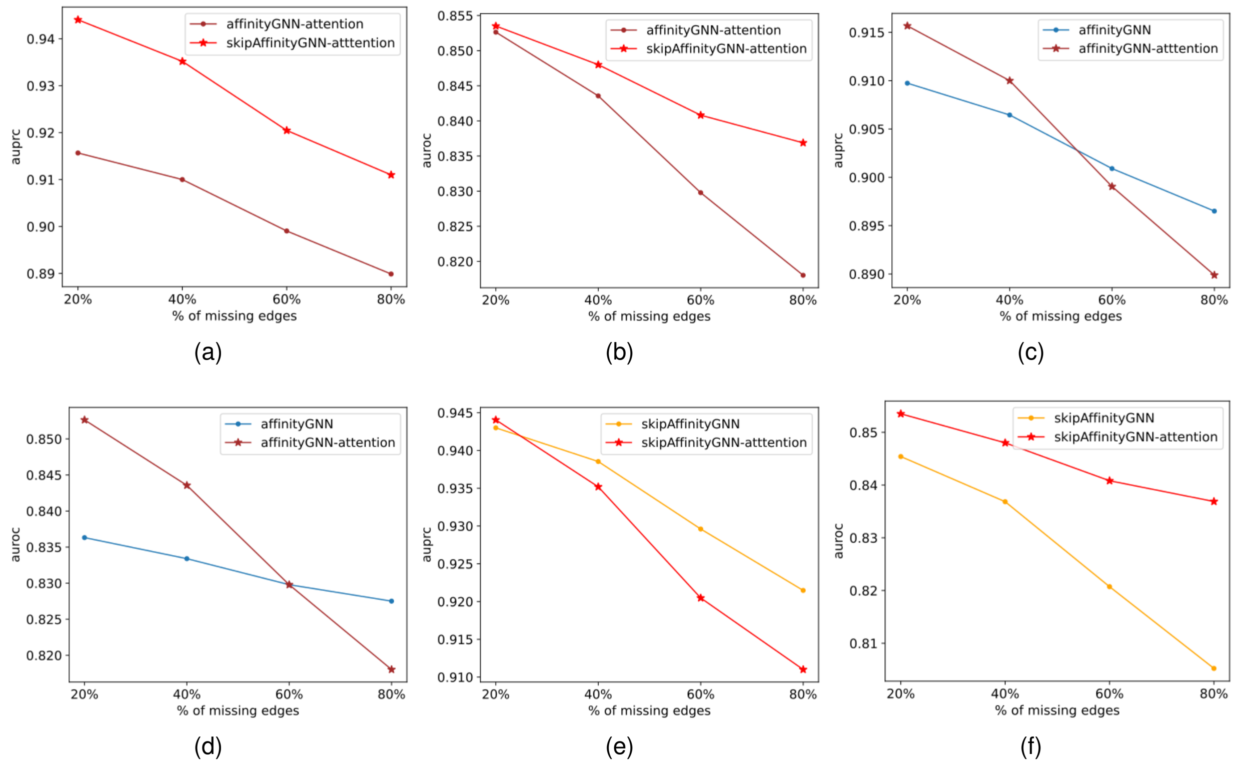


Fig. 6. Experimental results showing the predictive robustness peformance evaluation in the presence of incomplete network

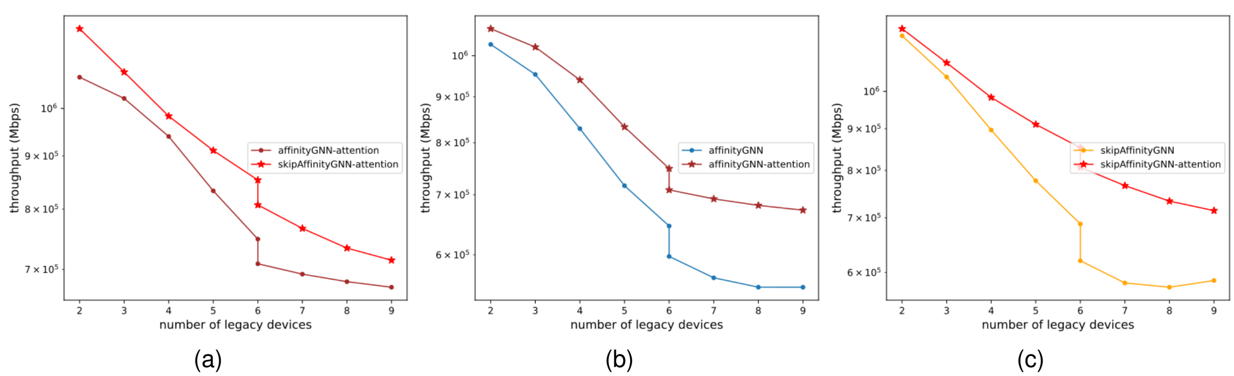


Fig. 7. Experimental results showing the throughput estimation in presence of legacy devices

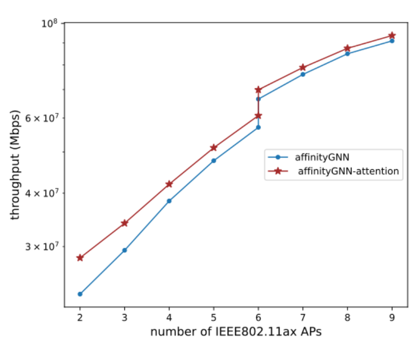


Fig. 8. Throughput estimation of IEEE802.11ax versus the number of connecting IEEE802.11ax APs

AP; it eventually reaches a maximum at the point where the transmission distance is increased to a certain extent. Interacting rate dominates network performance in networks with small transmission ranges, and it increases with increasing transmission range. As transmission range increases, a node's neighborhood size increases and the number of nodes competing to transmit causes more packet collisions. Consequently, as in Fig. 7&8, using various estimating methodologies, network throughput is charted against IEEE802ax and Legacy. The curves expresses the results of our improved affinityGNN algorithm, which demonstrates the optimal attributes in overall throughput estimation.

VI. Conclusion

The effect of variable transmission power in heterogeneous networks coexistence is relatively proportional to the node and network existence as the appropriate transmission power rate is dependent on the interaction range. While high transmission power will double the number of simultaneous transmission APs, it as well increases interference from prospective transmitting APs, resulting in low spatial reusing. In order to achieve maximum network throughput, it is therefore necessary to find a balance between transmission rate and interference levels effected by transmitted power. As a result, a model integrating attention structure with a deep learning approach, affinityGNN, is proposed in this study. To comprehend the transformation mechanism of the wireless network, a fusion of affinityGNN and an attention structure is utilized to capture the relative importance of each edge in terms of contribution to total network throughput. In the future, considerably introducing other factors asides transmission power that affect the wireless network, such as the impact of variable data rate and fixed data rate generated in the network to significantly improve the model's predictive capabilities.

References

[1] M. S. Afaqui, E. Garcia-Villegas, E. Lopez-Aguilera, G. Smith and D. Camps, "Evaluation of dynamic sensitivity control algorithm for IEEE 802.11ax", *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, pp. 1060-1065, Mar. 2015.

[2] M. Krunz, A. Muqattash, and S.-J. Lee, "Transmission power control in wireless ad hoc networks: challenges, solutions and open issues," *IEEE Network*, vol. 18, no. 5, pp. 8–14, Sept 2004.

[3] Stephen Azeez, Shagufta Henna and Mohamed Ibrahim, "Graph-based Deep Learning for Spatial Reuse Optimization in Dense WLAN Deployments", *IEEE Transactions and Communications*, 2022

[4] K. P. Shih, Y. D. Chen, and C. C. Chang, "A physical/virtual carriersense-based power control mac protocol for collision avoidance in wireless ad hoc networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 22, no. 2, pp. 193–207, Feb 2011.

[5] D. Dang, C. Hong, S. Lee, and J. Lee, "A sinr-based mac protocol for wireless ad hoc networks," *IEEE Communications Letters*, vol. 16, no. 12, pp. 2016–2019, December 2012.

[6] Afaqui, M. S., Brown, S., and Farrell, R. (2018)." Uplink performance optimization of ultra dense Wi-Fi networks using AP-managed TPC", *Wireless Days (WD)*. doi:10.1109/wd.2018.8361703, 2018.

[7] D. Qiao, S. Choi, A. Jain, and K. G. Shin, "Adaptive transmit power control in IEEE 802.11a wireless LANs," *VTC Spring.*, vol. 1, pp. 433–437 vol.1,April 2003.

[8] S. Gray and V. Vadde, "Throughput and Loss Packet Perfomnee of DCF with Variable Transmit Power.", *IEEE 802.1 I-011227*. May 2001.

[9] Khorov, E., Kiryanov, A., Lyakhov, A., & Bianchi, G., "A Tutorial on IEEE 802.11ax High Efficiency WLANs", *IEEE Communications Surveys & Tutorials*, doi:10.1109/comst.2018.2871099, 2018

[10] V. P. Mhatre, K. Papagiannaki, and F. Baccelli, "Interference mitigation through power control in high density 802.11 wlans," *IEEE INFOCOM*, pp. 535–543., May 2007.

[11] O. Oteri, P. Xia, F. LaSita, and R. Olesen, "Advanced power control techniques for interference mitigation in dense 802.11 networks," *WPMC*, pp. 1–7, June 2013.

[12] T. ElBatt, S.V. Krishnamurthy, D. Connors, and S. Dao, "Power management for throughput enhancement in wireless ad-hoc networks," IEEE International Conference on Communications (ICC), June 2000.

[13] J. Gomez and A. Campbell, "Variable-Range Transmission Power Control in Wireless Ad Hoc Networks," *IEEE Trans. Mob. Comput.*, vol. 6, no. 1, pp. 87–99, Jan. 2007.

[14] Schmidt, M., Block, D., and Meier, U.,"Wireless interference identification with convolutional neural networks". *IEEE 15th International Conference on Industrial Informatics (INDIN)*. doi:10.1109/indin.2017.8104767,2017

[15] Grunau S., Block D. and Meier U., "Multi-Label Wireless Interference Classification with Convolutional Neural Networks," *IEEE 16th International Conference on Industrial Informatics (INDIN)* pp. 187-192, doi: 10.1109/INDIN.2018.8471956, 2018.

[16] Y. Junfei, L. Jingwen, S. Bing and J. Yuming, "Barrage Jamming Detection and Classification Based on Convolutional Neural Network for Synthetic Aperture Radar," *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, 2018, pp. 4583-4586, doi: 10.1109/IGARSS.2018.8519373.

[17] Shen, Y., Shi, Y., Zhang, J., Letaief, K.B." A graph neural network approach for scalable wireless power control", *IEEE Globecom Workshops (GC Wkshps)*. pp. 1–6 (2019)

[18] Naderializadeh, N., Eisen, M., Ribeiro, A.," Wireless power control via counterfactual optimization of graph neural networks" *2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. pp. 1–5 (2020)

[19] Naderializadeh, Navid. "Wireless Link Scheduling via Graph Representation Learning: A Comparative Study of Different Supervision Levels." *arXiv preprint arXiv:2110.01722*, 2021.

[20] Z. Zhao, G. Verma, C. Rao, A. Swami, and S. Segarra, "Distributed scheduling using graph neural networks," *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4720–4724, 2021.

[21] Sangwoo M., Sumyeong A., Kyunghwan S., Jinwoo P., and Yung Yi. "Neuro-DCF: Design of Wireless MAC via Multi-Agent Reinforcement Learning Approach", *Proceedings of the Twenty-second International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing*. Association for Computing Machinery, New York, NY, USA, 141–150. DOI:https://doi.org/10.1145/3466772.3467043, 2021

[22] Mota, M.P., Valcarce, A., Gorce, J.M. and Hoydis, J.,"The Emergence of Wireless MAC Protocols with Multi-Agent Reinforcement Learning", *arXiv preprint arXiv:2108.07144*, 2021

[23] Jin, Y., Vannella, F., Bouton, M., Jeong, J. and Hakim, E.A., "A Graph Attention Learning Approach to Antenna Tilt Optimization", *arXiv preprint arXiv:2112.14843*, 2021.

[24] Li, M., Yuewen, W., Zhaowen, W., and Huiying, Z., "A deep learning method based on an attention mechanism for wireless network traffic prediction". *Ad Hoc Networks, 102258*, doi:10.1016/j.adhoc.2020.102258, 2020.

[25] Afaqui, M. S., Brown, S., and Farrell, R.,"Uplink performance optimization of ultra dense Wi-Fi networks using AP-managed TPC" *Wireless Days (WD)*, doi:10.1109/wd.2018.8361703, 2018.

[26] Ifedayo, A. O., and Dlodlo, M., "Variable transmission power control in Wireless Ad-Hoc Networks", *AFRICON*, doi 10.1109/afrcon.2015.7331980, 2015.

[27] M. Iwata, K. Yamamoto, B. Yin, T. Nishio, M. Morikura and H. Abeysekera, "Stochastic Geometry Analysis of Individual Carrier Sense Threshold Adaptation in IEEE 802.11ax WLANs," in IEEE Access, vol. 7, pp. 161916-161927, 2019, doi: 10.1109/ACCESS.2019.2951608.

[28] Doina B., Ajoy D., Prathima S. , and Zheng. S, "Impact of Variable Transmission Range in All-Wireless Networks", *42nd Hawaii International Conference on System Sciences*. doi:10.1109/hicss.2009.246, 2009.

[29] H. ElSawy and E. Hossain, "On stochastic geometry modeling of cellular uplink transmission with truncated channel inversion power control", *IEEE Trans. Wireless Commun.*, vol. 13, no. 8, pp. 4454-4469, Aug. 2014

[30] O. G. Adewumi, K. Djouani, and A. M. Kurien, "RSSI based indoor and outdoor distance estimation for localization in WSN," *IEEE Int. Conf. Ind. Technol.—*, pp. 1534–1539, Feb. 2013

[31] Li, M., Yuewen, W., Zhaowen, W., and Huiying, Z., "A deep learning method based on an attention mechanism for wireless network traffic prediction". *Ad Hoc Networks,102258*. doi:10.1016/j.adhoc.2020.102258, 2020

[32] K. Zen, M. N. Jambli, S. M. Suhail, H. Lenando, and A. Tully, "Transmission Power Control in Mobile Wireless Sensor Networks: Simulation-based Approach," *IET International Conference on Wireless Communications and Applications (ICWCA 2012)*, pp. 095–101.,2021.

[33] X. Zhang et al., "Deep learning for interference identification: band training SNR and sample selection" *SPAWC*, Cannes, France, pp. 1-5, 2019.

[34] Deng D.J., Lien S.Y., Lee J., Chen K.C, "On Quality-of-Service Provisioning in IEEE 802.11 ax WLAN", *IEEEAccess. 2016;4:6086–6104. doi: 10.1109/ACCESS.2016.2602281*

[35] Shen, Y., Shi, Y., Zhang, J. and Letaief, K.B., 2019, December. A graph neural network approach for scalable wireless power control. *IEEE Globecom Workshops (GC Wkshps)* (pp. 1-6). IEEE., 2019

[36] A. Ajami and H. Artail, "On the modeling and analysis of uplink and downlink ieee 802.11ax wi-fi with ltein unlicensed spectrum", *IEEE Transactions on Wireless Communications*, 16(9):5779–5795, Sept 2017.

[37] P. de Kerret, D. Gesbert, and M. Filippone, “Team deep neural networks for interference channels,”. *IEEE International Conference on Communications Workshops (ICC Workshops)*. IEEE, 2018, pp. 1–6.,2018

[38] D. Xu, X. Che, C. Wu, S. Zhang, S. Xu, and S. Cao, “Energy-efficient subchannel and power allocation for hetnets based on convolutional neural network,” *arXiv preprint arXiv:1903.00165*, 2019.

[39] Cheng H, Zhou JT, Tay WP, Wen B.,” Attentive graph neural networks for few-shot learning”, *arXiv preprint* arXiv:2007.06878., Jul 2020.

[40] Zheng C, Fan X, Wang C, Qi J., ”Gman: A graph multi-attention network for traffic prediction”, *Proceedings of the AAAI Conference on Artificial Intelligence*, (Vol. 34, No. 01, pp. 1234-1241), Apr 2020.

[41] Zhang, C., Yu, J. J. Q., and Liu, Y., ”Spatial-Temporal Graph Attention Networks: A Deep Learning Approach for Traffic Forecasting”, *IEEE Access, 7, 166246–166256*. doi:10.1109/access.2019.2953888, 2019.

[42] Veličković P, Cucurull G, Casanova A, Romero A, Lio P, Bengio Y., ”Graph attention networks”, *arXiv preprint* arXiv:1710.10903, Oct 2017.