Channel Selection Enhancement in High-Density Wi-Fi 6 Network Using Machine Learning

Rasha Qays Aswad

[1]Department of Mathematics, Al-Muqdad College of Education, Diyala University – Diyala, Iraq

Abstract: Despite the progress made by Wi-Fi 6 network ,they still suffer from interference in dense environments. This research aims to enhance the channel selection process using machine learning algorithms to reduce the interference. Synthetic dataset generated to simulate dense wireless network using the network properties such as RSSI, noise level, channel width, channel utilization, the number of adjacent access points, and interference level. The Random Forest (RF), Support Vector Machine (SVM) and K-Nearest Neighbors (KNN) as machine learning algorithms used to classify the channels and select the optimal channel. The evaluation results of algorithms performance show that Random Forest outperformed the other algorithms. Random Forest achieved high accuracy reached to 97.67%, also achieved high network performance thereby contributed to increase throughput by approximately 92% and reduce the delay by 41%, in addition to reduce packet loss rate by 78% compared with random channel selection method (Rand). The results also showed a direct relationship between improving classification accuracy and improving network performance metrics, confirming the effectiveness of using machine learning techniques in channel selection within dense wireless environments.

Keywords: *Wi-Fi 6, channel Selection, Machine Learning, Random Forest, SVM, KNN, interference*

1. Introduction

In recent years, wireless communication networks have been widely used in hospitals, homes, institutions, and public facilities to provide internet connectivity for devices. To provide high data rate and efficient communication in high-density environments, the IEEE 802.11 ax standard, also known as Wi-Fi 6, was developed and became widespread in May 2014 [1][2]. Wi-Fi 6 provided significant technical improvements in throughput, capacity, and spectrum utilization through technologies such as Mu-MIMO, OFDMA and Spatial Reuse [3].

Despite these improvements in Wi-Fi 6, it still faces challenges in dense environments with a number of connected devices, such as hospitals, universities, and residential areas. Among these challenges is that as the number of devices increases, interference between channels increases, making it difficult to manage and preventing better performance. This is because interference resulting from the entanglement of signals from close or overlapping channels can lead to a decrease in throughput, an increase in latency, and a rise in packet loss, all of which affect the quality of service [2,4].

Channel selection is one of the most significant variables impacting wireless network performance. Traditional methods for channel selection and network resource control rely on fixed or random selection method. However, these methods are unable to adapt to the dynamic nature of wireless networks and the interference and channel congestion that occur in real-world environments, resulting in low network performance. [5]

Due to the weakness of these methods, artificial intelligence and machine learning were used because they are considered powerful tools for improving network performance. Many previous studies have shown that machine learning is capable of selecting the optimal channel, predicting interference conditions, and dynamically optimizing channel access settings, which has led to improved network performance compared to traditional methods and techniques. [5,6]. The main contributions of this paper are as outlined below:

1- Enhance the channel selection in Wi-Fi 6 environment using MATLAB 2023.
2- Synthetic dataset generated to simulate the wireless network conditions based on set of network properties
3- Three machine learning algorithms were used: Random Forest, Support Vector Machine and K-Nearest Neighbor.
4- Evaluate and compare the performance of machine learning algorithms with the traditional methos of channel selection

The rest of this paper are structured as follows: Section 2 analyze the related and previous studies, section 3 details the methodology of this work, section 4 present the results, section 5 concludes the paper and outlines the future work.

2- Related works

Wireless networks have recently witnessed remarkable development with the emergence of the IEEE 802.11 ax standard, also known as Wi-Fi 6. This standard offers several technologies that mitigate interference and improve spectrum efficiency. Among the most important of these technologies are: Basic Service Set (BSS) coloring, orthogonal frequency-division multiple access (OFDMA), Target Wake Time (TWT), and Overlapping BSS. [2]

Despite this development and the emergence of these new technologies, the mechanisms used for channel selection still rely on fixed indicators such as RSSI and real-time occupancy indicators. This makes the network performance poor in terms of throughput and leads to an increase in latency and packet loss rate because the network is unable to adapt to dynamic changes, especially in high-density environments. [3][4][7]

Recent survey studies have shown that the use of machine learning techniques in managing wireless network resources is a promising trend for overcoming the problems of traditional methods. For example, a survey study [5] conducted by Szott et al., showed that the use of machine learning algorithms outperforms traditional methods by 15-25 %. While Baccouche and Andre [4] emphasized in their study that lightweight models must be used because of their adaptability to changing crowding patterns. In another study [2] by Mozaffariahrar et al., the problem of intelligent channel adaptation in Wi-Fi 6 was highlighted, where the use of dynamic predictive algorithms based on multiple channel criteria was emphasized.

From a practical perspective, Xin et al [8] designed the MAC protocol using deep neural networks (DNN) and recurrent neural networks (RNN) to integrate transmission rate adaptation, channel access mechanisms, and channel switching within a unified framework to select an optimal channel. This model achieved a clear improvement in throughput and reduced collisions. This model is based on limited-scale simulation and lacks application in a high-density environment. Natkaniec and Bieryt analyzed the effect of BSS coloring technology on the performance of a hybrid Wi-Fi 6 network [8]. The results showed that activating BSS coloring improves spatial spectrum reuse, leading to reduced latency and increased throughput compared to network performance without BSS coloring.

In the experimental perspective, Havinga et al. [7] have detected the interference between cross technology, and they used machine learning algorithms such as Random Forest and SVM with OFDMA, the classification accuracy reached 99.2%. However, the study focuses on cross-technology interference and not on the optimal internal channel selection within a high-density Wi-Fi 6 network. Day et al.[4] compared the traditional OFDM technology used in Wi-Fi 5 with the OFDMA and TWT technologies used in Wi-Fi 6. The comparison was conducted using NS3, and the results showed that OFDMA improved QoS and reduced power consumption using TWT. However, this study did not focus on using machine learning for smart channel selection. In a comparative study of machine learning algorithms by Chau et al. [9], the performance of classification algorithms (decision tree, random forest, and neural networks) was analyzed

Reference	**Method**	**Dataset**	**Limitation**	**Performance metrics**	**Results**
[4]	Analyzing Wi-Fi 6 mechanisms (TWT and OFDMA)	Simulation based on IEEE 802.11 ax standards	Focus only on QoS without treating interference	Throughput, Delay and Packet Delivery ratio	Improved QoS and latency reduced
[5]	Survey on machine learning application to enhance IEEE 802.11	N/A	N/A	N/A	Machine learning techniques outperform traditional methods
[6]	Analyzing the role of AI in Wi-Fi enhancement	Network simulation of real data	They relied on part of performance criteria without multi criteria integration	Throughput and Latency	Spectral interference reduced
[7]	Integrating OFDMA with Machine learning (RF and SVM) to reduce interference	Generating synthetic dataset	Focus on cross technology interference	Accuracy, scheduling efficiency and bit error rate	The accuracy reached 99.2% and improved scheduling efficiency
[8]	MAC protocol based deep learning	real-world data collection within the 2.4 GHz frequency band	Lacks adaptive channel selection in high density	Throughput and Delay	Improved throughput and reduced latency
[9]	Simulation Wi-Fi 6 with 5 GHz band (focus on BSS coloring mechanism)	Simulation of IEEE 802.11 ax standard	Lacks intelligent channel selection in high density	Latency, throughput, spatial spectrum reuse efficiency	Improved throughput and reduced latency but the interference still persists
[10]	Channel selection in wireless network using machine learning algorithms (RF, SVM and NN)	Synthetic data generated	IEEE 802.11a simulation. Does not emulate IEEE 802.11ax	SINR Improvement, Accuracy , F1-Score, Precision, Recall	The performance of Random Forest and Neural network is better than Decision Tree, where F1 > 95%. this led to network stability

for dynamic channel selection in a network using the IEEE 802.11a standard. Four criteria were used: RSSI, channel utilization, access point, and signal-to-interference-and-plus-noise-ratio (SINR). Data were generated during simulations, comprising 5000 samples collected from the dynamic changes of the network for various scenarios. The simulation results and experimental tests showed that the random forest algorithm achieved high predictive accuracy (98.8%) and improved the mean SINR by 12.65–14.21%. The summary of previous studies is presented in Table 1, which includes the methods used, dataset, study limitations, performance measures, and findings.

As shown in Table (1) above, most studies lack channel selection using machine learning algorithms in Wi-Fi 6networks. This study uses machine learning algorithms (Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbor Number (KNN) to improve network performance and reduce interference by enhancing channel selection in high-density Wi-Fi 6 networks.

3- Methodology

This study aims to improve the performance of Wi-Fi 6 networks and reduce interference by enhancing the channel selecting using machine learning algorithms in dense Wi-Fi 6 networks. A Wi-Fi 6 network simulation was performed using MATLAB 2023 and Communication Toolbox. Dataset was collected from the network's basic parameters and then trained using machine learning algorithms (Random Forest, KNN, and SVM). These algorithms classify channels to select the optimal one. Finally, the performance of the model and the network was evaluated using several performance metrics. Figure (1) illustrates the system's workflow, which starts with data generation and ends with performance evaluation.

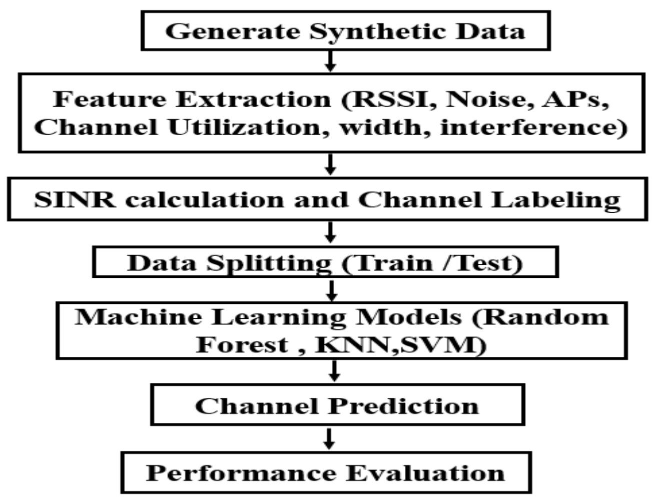

Figure (1) system workflow

3-1 Synthetic data generation

Due to the unavailability of databases that simulate real-world Wi-Fi 6 network conditions, synthetic data was generated in this study using MATLAB. The generating of Synthetic data is common in many studies [9][10] . 1500 samples were generated, each representing different scenarios simulating that simulate various wireless network conditions. Each sample contained a number of features that represented the behavior of a Wi-Fi6 network in a real-world environment as shown in Table 2.

Table 2 network features

Feature	Range
RSSI	-30 to -70dBm
Noise	-70 to -95dBm
Adjacent APs	1 – 20
Channel Utilization	0 – 1
Channel width	20MHz, 40 MHz, 80MHz
Interference	Adjacent APs × Channel utilization

a) Received signal strength indicator (RSSI): quantifies the received power level at a receiver from access point.[11]. Random values were generated in the range between -30dBm to -70dBm.
b) Noise level: represents the power of unwanted background signal in wireless channel, the noise values were generated within the range -95 dBm to -70 dBm.
c) Adjacent Access points (APs): the number of adjacent access points that cause the interference, the number of APs range between 1 to 20.
d) Channel Utilization: represents the percentage of channel utilization in the network, the continuous values were generated between 0 to 1
e) Channel width: to simulate the different Wi-Fi 6 scenarios, three levels of channel widths were defined (20 MHz, 40 MHz and 80 MHz).
f) Interference: to simulate the effect of interference in dense network, the interference was calculated based on channel utilization and adjacent APs.

These features were merged in Feature matrix to use in machine learning models training

3-2 channel labeling

The Signal-to-Interference-plus-Noise Ratio (SINR) was calculated to determine the optimal channel. SINR calculated for each sample according to equation 1, If SINR was low, this meant the interference is high and vice versa. The channel was classified into three classes using a scientifically defined threshold as shown in table 3.

$$SINR = \frac{RSSI}{Noise + Interference}$$

Table 3 channel labeling

classes	**Threshold**
Clean or Optimal channel	SINR > -0.5, channel utilization < 0.4
Moderate channel	Channel utilization < 0.7
Congested channel	Channel utilization ≥ 0.7 or SINR ≤ -0.5

The SINR is the most important indicator which is used to estimate the channel quality in wireless network. If SINR value reduces, the network performance decreases. The threshold (-0.5) was chosen to recognize the low-quality channel, this value represents the situation of weak signal compared to the interference, which indicate instable network condition in high density environment. In addition, channel utilization was used as the second indicator to represent the channel congestion, the threshold (0.4) indicates low utilization whereas the value (0.7) indicated to high congestion

This method of channel labeling simulates the real-world scenarios, where the clean channel achieves best performance and the congested channel reduced the throughput and increased latency and packet loss. To make the generated data more realistic and to simulate the conditions of wireless network environments, 8% random noise labeling was introduced. This addition aims to reduce overfitting and provide more realistic results for evaluating machine learning models.

3-3 Data splitting:

two techniques used to divide and evaluate the model's performance. The first is Hold – out method which determine the ratios of training and testing sets, where 70% of data used as training and 30% used as testing data. the second technique is K-fold cross validation with 5-Folds, this method reduces the overfitting and boosts the evaluation reliability, the dataset divided into 5 equally sets and the average of final results calculated.In this section, the data cleaning was not required because it's synthetic and was generated within specific ranges, thus free of incorrect or missing values. It also doesn't require normalization because it was generated within a convergent range.

3-4 Machine learning algorithms Parameter

To determine the optimal channel, three supervised machine learning algorithms were used in this research as follows:

3-4-1 Random Forest (RF): it is an ensemble learning algorithm. It creates a number of decisions trees and then combines their results to achieve high classification accuracy [10]. in this research, 50 trees were used, and the out-of-Bag Error used to estimate evaluate the performance through training.

3-4-2 Support Vector Machine (SVM):it is based on obtaining the best level of hyperplane between the different classes. It considers one of the strongest classification algorithms [11]. In this research, multi-class SVM was used, and Error Correcting output code (ECOC) method applied to support multi classes classification.

3-4-3 K-Nearest Neighbors (KNN): this algorithm depends on sample distance, where the new sample classified based on its nearest neighbor [12] . In this research the number of neighbors is 5.

The architecture of training data using the machine learning algorithms is shown in figure 2.

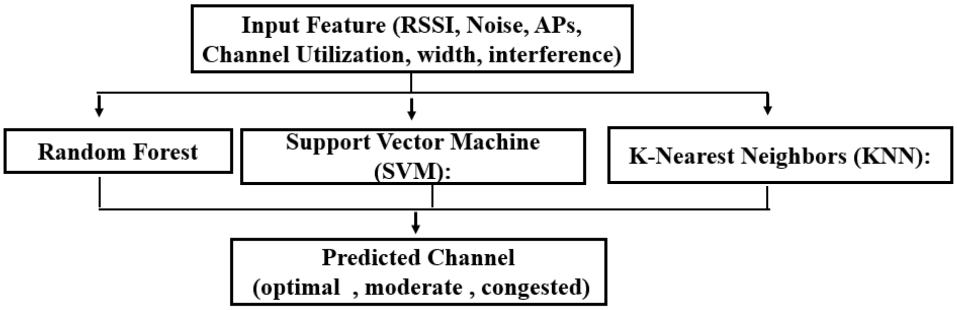

Figure2: machine learning algorithms architecture

3-5 performance metrics

3-5-1 network performance metrics: To evaluate the network performance, this research used the following metrics:

1- Throughput: It represents the amount of data that is transmitted successfully per unit time. It usually measured by Mbps. The network efficiency increases when the throughput increases, and also the high throughput causes low interference.
2- Delay: represent the required time that used to transfer the data through the network. It is measured by millisecond (ms). The network performance enhances as the delay decrease.
3- Packet loss: it represents the percentage packet that missed through the data transmission. It calculates based on the lost packet and sent packet.

3-5-1 Classification performance metrics: four metrics used to evaluate the correctness of the model results. These four metrics are: accuracy, Precision, Recall and F1 score, they are based on the confusion matrix which contain: True positives (TP), True Negatives (TN), False Positive (FP) and False Negative (FN).

1- Accuracy: is the most important metric that used to evaluate the correctness of ML algorithms. It is calculated as follow

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

2- Precision: it is the proportion of optimal channels for each class compared with all predictions for that class. It is calculated as follow:

$$Precision = \frac{TP}{TP + FP}$$

3- Recall: the ability of the model to detect all the correct samples in each class. It is calculated as follow:

$$Recall = \frac{TP}{TP + FN}$$

4- F1 score: it is a harmonic mean between Recall and Precision. It is calculated as follow :

$$F1\ score = 2\ x\ \frac{Precision\ x\ Recall}{Precision + Recall}$$

4-Results and discussion

4-1 classification results

Table (4) represents the evaluating machine learning models results using K-Fold Cross Validation. The results shows that the Random Forest (RF) algorithm outperformed the others algorithms in all classification metrics, the superiority of RF algorithm attributed to its efficiency in dealing with non- linear data. The results also show that the precision, F1-score, and Recall values are relatively close, this indicates that the RF algorithm is capable of classification with minimal error. This performance is attributed to the algorithm's reliance on a set of decision trees for prediction, thus reducing the impact of noise in the data.

In contrast, the performance of Support Vector Machine (SVM) algorithm performed fairly well, but still lower than the RF algorithm, this is due to its sensitivity to class interference and complex data properties within the wireless network environment. The performance of K-Nearest Neighbors (KNN) model was the lowest among the models used. The KNN algorithm depends on distance calculation between samples, which makes it suboptimal in dense environments, and it is also sensitive to data interference

Table 4 K-Fold Cross Validation Results

ML algorithms	**Accuracy**	**Precision**	**Recall**	**F1**
RF	94.7%	94.2%	93.3%	93.7%
KNN	73.0%	70.5%	71.3%	70.2%
SVM	89.5%	90.9%	84.9%	87.2%

Table 5 shows the results of algorithms evaluation using Hold-out method. The results are very similar to the K-Fold cross validation results, this indicates the stability of models and overfitting reduction. The results in Table 2 show that the Rf model still outperforms the other algorithms used. The SVM algorithm performed well but less than RF, while KNN performed worse than the other algorithms. These results demonstrate the reliability of the models and their ability to generalization in high-density Wi-Fi 6 networks.

Table 5 Hold Out Results

ML algorithms	Accuracy	Precision	Recall	F1
RF	93.78%	92.9%	92.3%	92.6%
KNN	71.78%	69.4%	69.0%	68.2%
SVM	90.44%	90.8%	87.4%	88.9%

4-2 network performance metrics results

This section analyzes the network performance metrics of the used machine learning algorithms compared with the traditional channel selection method, which is labeled (Rand) in the following figures.

Figure (3) illustrates the throughput results of the three algorithms compared with Rand. The highest throughput is 27.61 Mbps which achieved by RF algorithm, followed by the SVM at 27.23 Mbps. The KNN throughput reached 25.57 Mbps which lower than RF and SVM, while the lowest throughput achieved by the tradition method (Rand). The enhancement of the machine learning algorithms' performance is attributed to their ability to choose the optimal channel in a congested environment, this led to an increase in the efficiency of data transmission.

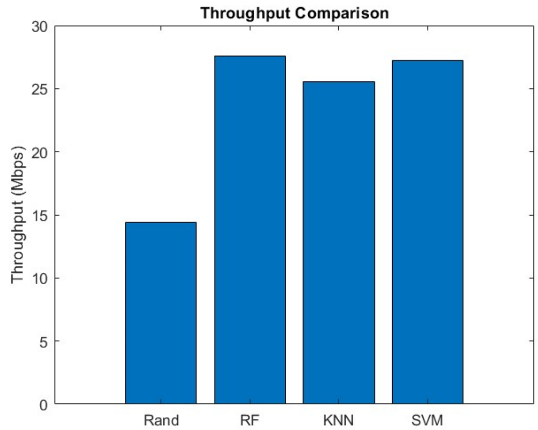

Figure 3 Throughput Results comparison

The delay results shown in Figure 4 indicate that the RF model achieved a significant improvement in delay compared to other models, with a delay value of 189.02 ms. This was followed by the SVM algorithm, which achieved a delay of 190.75 ms, while the KNN model recorded a higher delay of 253.80 ms. The delay performance of these algorithms was better than the traditional Rand method, with a delay value of 319.80 ms.

These results demonstrate that optimizing channel selection leads to reduced congestion and packet retransmission, which directly translates to lower network latency.

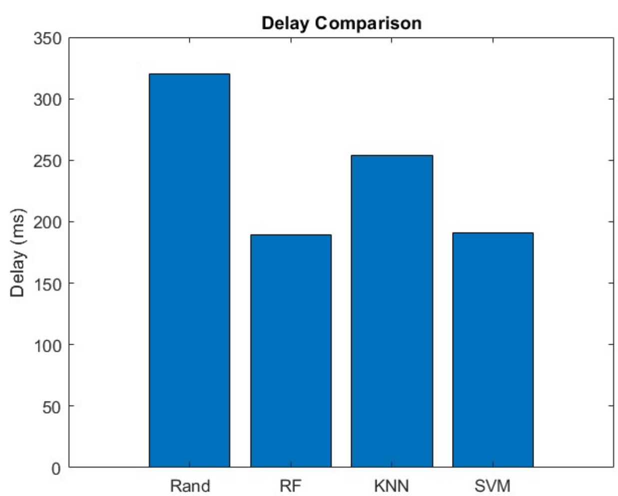

Figure 4: Delay Results comparison

Figure 6 represent the packet loss results comparison between three ML algorithms and Rand. The RF model achieved less percentage of packet loss which reached 3.40%, then the SVM packet loss with 5.22%, while KNN recorded high packet loss percentage of 15.44% which was close to Rand packet loss amounted 15.42%. These results confirm a direct relationship between channel classification accuracy and network reliability enhancement, as higher-accuracy models contributed to reducing packet loss and improving communication quality in dense Wi-Fi 6 environments.

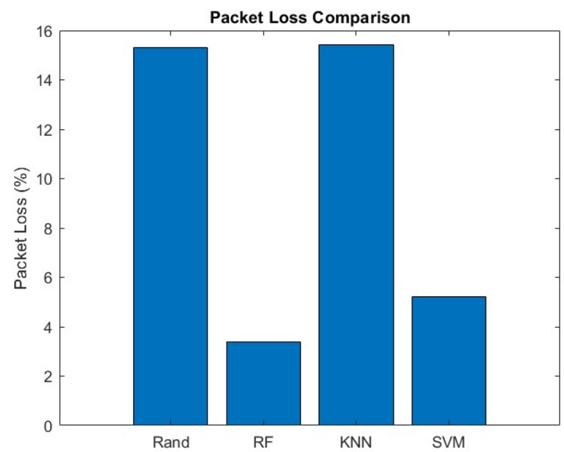

Figure 6 Packet Loss Results comparison

In general, the performance of Random Forest (RF) was the best compared with the SVM and KNN. The model achieved an increase in throughput of approximately 92%, along with a reduction in delay of approximately 41%, as well as a reduction in packet loss of approximately 78%. The results also demonstrate to direct relationship between the classification accuracy result and network performance result, the RF achieved high accuracy and also achieved high network performance metrics. These results indicate that relying on machine learning algorithms in channel selection can effectively contribute to improving the efficiency and reliability of Wi-Fi 6 networks, especially in high-density, high-interference wireless environments.

5- Conclusion and Future works

This paper focused on channel selection problem in high density Wi-Fi network which suffers from interference. In order to enhance the channel selection in dense environment, three machine learning algorithms evaluated using synthetic dataset that simulate the wireless environment based on wireless properties such as RSSI, noise level and access point number. The results showed that the Random Forest (RF) algorithm achieved the best performance compared to the rest of the algorithms in terms of classification accuracy and improvement of network performance metrics, such as increasing throughput, reducing delay, and packet loss. The results confirmed that the machine learning algorithms enhance the communication quality and reduce the interference in dense Wi-Fi 6 network. For future work of this research, we plan to use deep learning models to enhance channel selection in dynamic wireless network.

References

[1] Official IEEE 802.11 Working Group Project Timelines. Available online:https://www.ieee802.org/11/Reports/802.11_Timelines.htm(accessed on 24 February 2026).

[2] E. Mozaffariahrar, F. Theoleyre, and M. Menth, "A Survey of Wi-Fi 6: Technologies, Advances, and Challenges," *Future Internet*, vol. 14, no. 10, p. 293, Oct. 2022, doi: 10.3390/fi14100293.

[3] E. Khorov, A. Kiryanov., A. Lyakhov, and G. Bianchi.A tutorial on IEEE 802.11 ax high efficiency WLANs. *IEEE Communications Surveys & Tutorials*, *21*(1), 197-216.2018.

[4] C. T. Day, A. Salama, R. Saatchi, M. Bagheri, N. Ul Hasan, and S. Betts, "Evaluating OFDMA and TWT in Wi-Fi 6/7 for QoS Assurance in IoMT Networks," *Electronics (Basel).*, vol. 15, no. 5, p. 911, Feb. 2026, doi: 10.3390/electronics15050911.

[5] Szymon Szott; Katarzyna Kosek-Szott; Piotr Gawłowicz; Jorge Torres Gómez; Boris Bellalta; Anatolij Zubow, "Wi-Fi Meets ML: A Survey on Improving IEEE 802.11 Performance With Machine Learning," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 3, pp. 1843–1893, 2022.

[6] Y. Baccouche and J. Andre, "AI and Machine Learning for Wi-Fi Optimization," *In proceedings of the seminar Innovative Internet Technologies and Mobile Communications (IITM)*, vol. NET-2025-05-1, pp. 55–59, 2025.

[7] T. Havinga, X. Jiao, W. Liu, B. Chen, A. Shahid, and I. Moerman, "Wi-Fi 6 Cross-Technology Interference Detection and Mitigation by Ofdma: an Experimental Study," in *2025 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*, IEEE, Jun. 2025, pp. 1–6. doi: 10.1109/EuCNC/6GSummit63408.2025.11037177.

[8] J. Xin, W. Xu, B. Cao, T. Wang, and S. Zhang, "A deep-learning-based MAC for integrating channel access, rate adaptation, and channel switch," *Digital Communications and Networks*, vol. 11, no. 4, pp. 1042–1054, Aug. 2025, doi: 10.1016/j.dcan.2024.10.010.

[9] M. Natkaniec and N. Bieryt, "An Analysis of the Mixed IEEE 802.11ax Wireless Networks in the 5 GHz Band," *Sensors*, vol. 23, no. 10, p. 4964, May 2023, doi: 10.3390/s23104964.

[10] A. Chau, H. Chan and Geraghty. Channel Selection using Machine Learning Algorithms for Wireless Access Points. In *2025 IEEE International Conference on Technology Management, Operations and Decisions (ICTMOD)* (pp. 1-6), October 2025 IEEE.

[11] C. Cortes and V. Vapnik, "Support-Vector Networks," Machine Learning, vol. 20, no. 3, pp. 273–297, 1995.

[12] T. Cover and P. Hart, "Nearest Neighbor Pattern Classification," IEEE Transactions on Information Theory, vol. 13, no. 1, pp. 21–27, 196.