Real-Time Performance Analysis of Multi-Armed Bandit Solution for OBSS and Dynamic Channel Bonding in WLAN

Remya P.R
Department of Computer Science and Engineering
Amrita School of Computing, Coimbatore
Amrita Vishwa Vidyapeetham, India
pr_remya@cb.students.amrita.edu

J.Govindarajan
Department of Computer Science and Engineering
Amrita School of Computing, Coimbatore
Amrita Vishwa Vidyapeetham, India
j_govindarajan@cb.amrita.edu

***Abstract*—Rapid expansion of dense Wi-Fi network deployments creates challenges in optimizing performance. Existing multi-armed bandit (MAB) solutions for dynamic channel bonding (DCB) and overlapping basic service set (OBSS) issues are tested using simulations. This research aims to identify the gaps in current solutions through real-time performance analysis. Alongside this primary objective, we proposed a modified exploration-first algorithm to minimize convergence time. This secondary objective helped us achieve the primary objective. The results demonstrate that our algorithm can adapt to dynamic Wi-Fi environments and improve the user experience. We also present a summary of the inferences from the primary objective study.**

***Index Terms*—WLAN, Dynamic Channel Bonding, OBSS, MAB, Exploration-First**

I. Introduction

Cellular networks and WiFi are among the most widely used wireless technologies in the age of ubiquitous internet access. Every home, office and public place now uses WiFi access points due to technological advancements and their affordable cost. This has resulted in a denser wireless local area network (WLAN) environment [1] in all locations. On the other hand, this large-scale expansion introduces new challenges, such as high collision among the access points (APs), and affects the expected performance from these APs. Various IEEE 802.11 standards came at different times to the industry to provide better WiFi performance. The IEEE 802.11ac standard [2] was launched in 2013 to improve spectral efficiency and therefore increase the overall capacity of the network. However, controlling all these features in a dense WLAN still requires adaptive optimization techniques. Dynamic channel bonding (DCB) is a feature that enhances data transmission rates and network performance by merging multiple channels into a single broader channel. DCB works efficiently in a dynamic and challenging environment with varying levels of wireless interference and traffic loads.

A viable strategy to tackle the problems of a dynamic environment is the utilization of multi-armed bandit (MAB) [3] algorithms, which are a subset of Reinforcement Learning (RL) algorithms. They offer a comprehensive framework for decision-making under uncertain conditions. MAB algorithms maximize network performance in real time by balancing exploration and exploitation to adjust WiFi characteristics, including transmission power and bandwidth. This research examines the efficacy of the modified exploration-first method to enhance WLAN performance in densely populated WLANs. This study aims to analyze the algorithm to improve throughput and guarantee equitable access for all users to the wireless medium. We implemented the deployment in actual network settings, encompassing diverse node densities and traffic patterns. The findings indicate that the modified exploration-first approach can substantially improve WLAN performance by adjusting to the fluctuating characteristics of congested wireless environments.

This research underscores the ability of MAB algorithms to transform WLAN management in congested settings, providing a pathway to more intelligent and adaptive wireless networks. By addressing the challenges posed by high-density deployments, MAB-based approaches can play a crucial role in meeting the growing demands for reliable and high-performance wireless communication. The key contribution of this paper lies in the algorithm's ability to converge to a specified percentage of maximum throughput.

Contributions of this paper: 1. Real-time study of the MAB algorithm for the OBSS and DCB problem in WLAN since most of the existing solutions and results are based on the simulation environment. The main aim of this work is to show the various aspects to be considered in the design of the MAB algorithm for dynamic channel bonding rather than to compare different MAB algorithms. 2. This paper proposes a modified exploration-first algorithm and demonstrates its performance through a real-time experiment. Dynamically adjusting these parameters will enable the algorithm to adapt its behavior in response to changing environmental conditions.

II. Various aspects of WLAN performance

A wireless local area network is a network that enables devices to connect and transmit wireless signals within a limited region, typically a building or campus. WLANs provide adaptable and mobile access to network resources through the transmission of data between devices via radio frequency signals. Unlike traditional wired networks, they are characterized by their cost effectiveness, scalability, and ease of implementation. Security protocols such as encryption and authentication techniques protect data transmissions over WLANs, preserving the confidentiality and integrity of the data. WLANs provide interoperability between various devices and manufacturers by accommodating various configurations and standards, such as IEEE 802.11, that govern their operation. WLANs are increasingly sophisticated in terms of speed, performance, and coverage due to technological advancements, making them indispensable for modern connectivity solutions in companies, residences, and public spaces globally.

WiFi 5 (802.11ac) improves network efficiency in multiple important ways. It can handle channel bandwidths of up to 160 MHz, which enables faster data transfer and less congestion. With the advent of Multi-User Multiple Input Multiple Output (MU-MIMO), several devices can communicate simultaneously, increasing network efficiency overall by decreasing data transmission wait times. The most recent version of wireless networking standards, IEEE 802.11ax, also known as WiFi 6, is intended to improve WLAN performance and efficiency in high-density areas. It expands on 802.11ac (WiFi 5) by adding more devices and enhancing network efficiency by decreasing wait times for major advancements. WiFi 6 is capable of faster data transfer. By encoding more bits per symbol, 256-QAM optimizes throughput, hence increasing transmission rate. Orthogonal Frequency Division Multiple Access (OFDMA), which enables numerous devices to share the same channel simultaneously, lowering latency and increasing efficiency in busy conditions, is one of the most notable enhancements. Moreover, 802.11ax offers MU-MIMO in both the uplink and downlink directions, allowing simultaneous communication with several devices and boosting performance overall. Target Wake Time (TWT), which extends battery life, especially for IoT devices, is another feature of the standard that minimizes power usage by allowing devices to arrange their connection periods with the access point. Together, these advancements ensure better spectrum utilization, higher data rates, and improved performance, making WiFi 6 particularly effective in environments with many connected devices. 802.11ac and 802.11ax specification [2] [4] gives details of each of these features.

A. *Channel Bonding and Dynamic Channel Bonding*

Channel bonding refers to the technique of combining multiple adjacent radio frequency channels to increase the overall data throughput and bandwidth available for data transmission. WLANs with dynamic channel bonding can improve overall network performance and provide a more stable user experience by adapting channel widths as needed. However, in congested areas with many competing devices, this can lead to increased interference and reduced performance, making proper management of DCB in WLAN critical to fully achieving its advantages.

B. *Beamforming*

Beamforming is a wireless communication technique that improves signal strength and data transmission efficiency. Usual wireless signals are broadcast in all directions. Beamforming focuses the wireless signal in a specific direction. This process enhances and strengthens the signal, which improves coverage, speeds up data transfer, and reduces interference from other devices. An AP that supports the 802.11ac standard explicitly performs beamforming by sending sound frames. AP decides the phase and amplitude of the signals from the feedback from the client device. Network performance is increased when beamforming is used in combination with downlink MIMO. 802.11ax supports uplink and downlink MU-MIMO. 802.11ax also uses cutting-edge technologies such as Basic Service Set Coloring (BSS) and OFDMA. These technologies make the best use of spectrum and lower interference to create a more reliable and effective communication infrastructure. Taking everything into account, 802.11ax significantly boosts the effectiveness of beamforming to improve WiFi performance, addressing the challenges posed by contemporary network environments, even when both protocols utilize it.

C. *Transmission Power*

In wireless communication systems, such as WLANs, transmission power is one of the most important characteristics. Transmission power refers to the strength with which wireless equipment transmits its signals. This includes devices such as access points or client devices sending signals into the surrounding environment. While a signal's range is usually increased with higher transmission power, interference and power consumption may also rise. In order to maximize WLAN coverage, dependability, and performance and to ensure efficient communication while reducing interference and power consumption in a variety of deployment scenarios, transmission power and bandwidth are crucial.

D. *Collisions in OBSS*

Collisions and OBSS problem interference have a significant impact on WLAN performance. OBSS interference occurs when nearby APs use the same channel or adjacent channel for transmission. On the other hand, a collision occurs when several devices try to send data on the same channel at the same time. This causes data packets to get corrupted and needs re-transmission. As a result of these problems, the overall network performance is reduced.

WLANs use techniques like Request-to-Send (RTS) and Clear-to-Send (CTS) frames to manage contention and Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) to restrict access to the medium to reduce collisions. Effective channel planning, in which APs are set up to use non-overlapping channels, or sophisticated strategies like dynamic frequency selection (DFS) to avoid channels from being used by radar systems, can be employed to mitigate OBSS interference. Maintaining reliable and high-performance WLANs requires effective management of collisions and OBSS interference, particularly in crowded deployment scenarios.

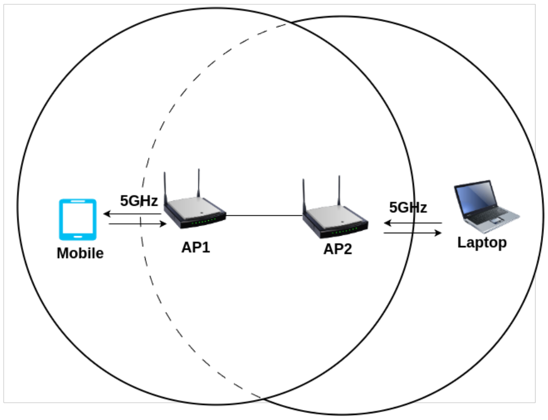

Fig. 1. Real-time set-up to demonstrate OBSS problem

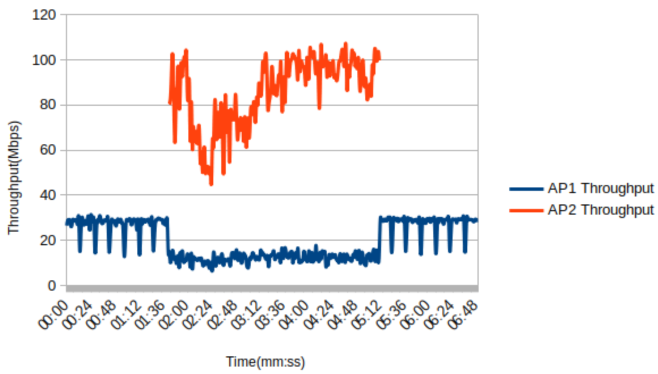

Fig. 2. Throughput measured with OBSS problem

E. Real-time demonstration of OBSS problem

We conducted real-time experiments using two APs with connected devices, testing them on the same channel with user-datagram protocol (UDP) traffic. A noticeable drop in throughput was observed when both access points operated on the same channel and their transmission ranges overlapped. This performance reduction aligns with the OBSS problem. The experiments and corresponding graphs are presented in Fig. 1 and Fig. 2 to illustrate this OBSS problem. This experiment was carried out using the Asus AX1800 dual-band WiFi 6 (802.11ax) router and the TP-Link Archer C20 AC750 wireless dual-band router. The experiment started with AP1, where the measured throughput was 30 Mbps. After 2 minutes, the Asus router started to operate in the same channel, resulting in a throughput of approximately 10 Mbps. This indicates a reduction of 20 Mbps. This reduction is attributed to the OBSS problem.

III. Existing MAB Solutions: A Summary

Multi-armed bandit (MAB) solutions are algorithms used in WLANs to improve network performance by dynamically allocating resources, such as transmission power or channel selection, based on feedback received in real time.

Recent advances in wireless networking have led to significant research on optimizing network performance using MAB algorithms. Addresses the issue of resource starvation by dynamically adjusting channel selection based on performance observations. This approach improves the user experience and enhances overall throughput, showcasing MAB's practical applicability in real-world scenarios.

In this paper [5], they describe a centralized technique that uses the MAB architecture to dynamically alter the transmission power and the sensitivity threshold of the Clear Channel Assessment (CCA). They propose a new technique for sampling novel network configurations using a Gaussian mixture based on Thompson Sampling and a reward function. Maximizing reward function prevents starvation. By maximizing transmission power and sensitivity threshold parameters using an MAB framework, this paper [6] offers a way to improve spatial reuse in WLANs.

In order to improve the spatial reuse of radio channels, [7] focuses on optimizing the sensitivity threshold and the transmission power. Through the use of Gaussian processes (GPs) to carry out local Bayesian optimizations, INSPIRE's distributed online learning methodology enables each access point to modify its settings in response to local conditions independently. The paper [8] examines parameters related to network performance, specifically in WiFi settings, including transmission power and CCA thresholds, which are crucial to optimizing performance in dynamic environments. The primary algorithm used in the study is the multi-armed bandit approach, specifically the multi-armed contextual multi-armed bandit (MA-CMAB) solution. The paper also discusses various MAB algorithms, including ϵ-greedy, Upper Confidence Bound (UCB), and Thompson Sampling, comparing their performance under different action sets. This paper [9] proposes a Bayesian optimization on-line learning algorithm that uses GPs to optimize CCA thresholds and transmit power. This approach formulates the spatial reuse problem as a multi-armed bandit problem. The algorithm aims to maximize the reward function, improve total throughput, reduce starvation between stations, and enhance network fairness. The difficulties with DCB in WLANs are discussed in this work [10], with a focus on high-density settings where there may be overlaps between several BSSs. It highlights how crucial spectrum management and allocation are, taking into account variables such as throughput, latency, and the impact of dynamic traffic loads. The study criticizes the intricacy of conventional RL algorithms and promotes stateless methods, contending that MABs enable quick adaptation without requiring substantial state definitions. In all these contributions, the algorithms are tested in a simulated environment, and a maximum of only two parameters are considered for optimization. In addition to MAB, many other approaches, such as deep learning or deep neural networks, are being applied to solve these types of problem [11] [12] [13]. Reinforcement learning is being applied to other problems related to wireless, like routing and connectivity [14] [15].

These MAB solutions leverage machine learning algorithms and adaptive control techniques to make intelligent decisions in real time, continuously learning from feedback to optimize WLAN performance in terms of throughput, reliability, coverage, and user experience. As WLAN environments become more complex and dynamic, these adaptive MAB approaches play a crucial role in maximizing the efficiency and effectiveness of wireless network operations.

IV. Need for the real-time study of Dynamic Channel Bonding and OBSS problems using MAB algorithm

The real-time study of DCB and the OBSS problem in WiFi networks using a multi-armed bandit algorithm is essential for several reasons:

Dynamic Channel Bonding: To boost throughput, DCB enables WiFi devices to bond several nearby channels. However, interference or competition from other devices can cause channel availability to fluctuate in real-time, which might result in ineffective bonding decisions. Through adaptive channel selection that balances the trade-off between bonding for increased throughput and avoiding interference, real-time research employing the MAB algorithm can help improve overall network performance.

Overlapping Basic Service Set Problem: This type of interference occurs when several networks use adjacent or overlapping channels, particularly in densely populated WiFi areas. By choosing channels with less interference and modifying transmission parameters to reduce collisions, MAB algorithms, which are made to dynamically explore and exploit configurations based on the real-time environment, can help mitigate this problem.

Real-Time Adaptation: High-dynamic wireless environments, such as those involving DCB and OBSS, are subject to sudden changes in user density, channel availability, and interference. Learning from continuous network performance, the MAB algorithm allows real-time adaptation, enabling the WiFi system to continuously modify its configuration to reduce interference, increase spatial reuse, and improve user experience.

Complexity handling: Real-world WLAN environments can be highly dynamic and complex, with interactions between multiple factors (e.g. user mobility, interference patterns, network topology changes). Real-time studies allow MAB algorithms to take decisions based on observations. This allows them to work more effectively by improving the strategies in every iteration.

Finally, the MAB algorithm's real-time analysis of DCB and OBSS is important for adapting to changing WiFi conditions, making the best use of spectrum, and lowering interference. All of these make the network work better and more fairly in dense installations.

V. Proposed Modified Exploration-First Algorithm

One of the most effective multiarmed bandit (MAB) algorithms, the exploration-first algorithm optimizes WiFi parameters to enhance performance. In this approach, exploration is conducted exclusively at the beginning, followed by the exploitation phase. Our initial study on the standard exploration-first algorithm for this problem reveals that the initial exploration itself will take more than an hour. From this understanding of the exploration-first algorithm, we proposed the modified exploration-first algorithm to minimize the delay introduced by exploration. The challenge arises when the number of configurations to explore is large, as the exploration phase becomes time-consuming.

Convergence time of Exploration-First Algorithm: If n is the total number of configurations and t is the testing time per configuration in minutes, the total time required to explore all configurations (convergence time) is $n \times t$ minutes.

In our experiment, we have taken 4 channels, 5 transmission powers, and 3 bandwidths. Hence n is equal to 60 as given below.

$$\begin{aligned}\text{Total number of configurations to explore (n)}\\ &= 4 \times 5 \times 3\\ &= 60\end{aligned}$$

Each configuration is tested for two minutes. Hence, the convergence time is 120 minutes, as is given below.

$$\begin{aligned}\text{Convergence time of Standard Exploration-First}\\ &= 2 \times 60\\ &= 120\,\text{minutes}\end{aligned}$$

To reduce the exploration time, it is essential to optimize the number of configurations to be evaluated.

Convergence time of Modified Exploration-First Algorithm: In our algorithm, at the beginning of every iteration, a subset of configurations is filtered according to the channels available. This reduces the number of configurations to be explored; therefore, the convergence time will be minimized proportionally. Let m represent the reduced number of configurations after applying the algorithm, where $m < n$. Consequently, $m \times t \; < \; n \times t$

In our experiment, we observed that the maximum subset size m is 11. Hence, the maximum convergence time of our modified algorithm is 22 minutes.

$$\begin{aligned}\text{Convergence time of Modified Exploration First}\\ &= 2 \times 11\\ &= 22\,\text{minutes}\end{aligned}$$

Optimization is applied across all WLAN parameters. The channels are optimized by scanning the channels available for that interface each time before making a selection. The most suitable combination of bandwidth and transmission power will be selected to optimize the throughput continuously, starting from the higher bandwidth and low transmission power. At every step, the next lowest possible configuration (next lowest bandwidth, next highest transmission power) will be explored, and the rewards (R_i and R_{i+1}) will be compared between the current configuration and the previous configuration. If R_{i+1} is less than R_i, then all combinations of configurations related to the lower bandwidths will not be considered for learning. Otherwise, the exploration will continue with the next lower possible configurations [refer to Algorithm 1 in Fig. 3].

In our experimental setup, we utilized channels 36, 40, 44, and 48, with transmission powers of 5, 10, 15, 20, and 24 dBm and bandwidths of 20, 40, and 80 MHz.

Additional details about our experimental setup include the fact that the experiments were conducted in a congested environment. We intentionally created a dense setting by placing the routers just 1 meter apart. This proximity ensured that even with the lowest transmission power settings, the routers stayed within each other's range, allowing for a realistic evaluation of their performance in an overlapping WiFi environment. The reward for each configuration was the throughput measured during the experiment for that specific configuration. It was crucial to ensure that the reduction in transmission power did not exclude any client from the coverage area, so careful consideration was taken when selecting the highest transmission power. In our setup, client devices were placed at a distance from the router to ensure they remained consistently within range.

Algorithm 1: Exploration phase of Modified Exploration-First Algorithm

Input:
  C: Array of channels
  AC: Array of available channels by scanning the medium (subset of C)
  P: Array of Transmission powers
  B: Array of bandwidths
  R: Array of Rewards for each configuration
Output: Selection of subset of configurations for exploration
Initialize array *AC* $\leftarrow [c_1, c_2, \ldots, c_i]$
Initialize array *P* $\leftarrow [p_1, p_2, \ldots, p_j]$
Initialize array *B* $\leftarrow [b_1, b_2, \ldots, b_k]$
Initialize array *R* $\leftarrow [0_1, 0_2, \ldots, 0_{i*j*k}]$
Total Configurations $n = i \times j \times k$
Channels in usage *UC* = list of channels which are used by the other access points
Available Channels *AC* = *C* - *UC*
for each c in *AC* **do**
  for each p in *P* **do**
    for each b in *B* **do**
      Apply configuration with values c, p and b
      Measure Throughput and assign to *CurrentThroughput*
      Update Throughput in the Array R for the configuration
      if *CurrentThroughput* > *PreviousThroughput* **then**
        PreviousThroughput = *CurrentThroughput*
      else
        Break
      end if
    end for
  end for
end for
Return *R*

Fig. 3. Exploration phase of Modified Exploration-First Algorithm

VI. Experimental Design

We have conducted all our experiments using exploration-first as a representative of various MAB algorithms, since our objective is to understand the various aspects of designing MAB-based solutions for OBSS and DCB problems.

We formulated our experimental design to reflect the congested urban environment. In densely populated urban environments, such as busy marketplaces and residential areas, almost everyone uses their access points to meet daily multimedia needs such as video downloading, uploading, streaming news, watching movies, and sharing content on social media. Today, much of this activity is conducted on smartphones. In these settings, proper planning and coordination of the deployment of access points among users is often impractical. Consequently, access points are frequently placed close, with distances as short as 2 to 3 meters. In addition, access points, hotspots, and dongles also contribute significantly to the creation of this dense network environment.

Our experiment setup is presented in Fig. 4. The access points were the Archer C20 AC750 wireless dual-band router (AP1) and the Archer C60 AC1350 dual-band wireless router (AP2). Both access points are working in the 5GHz band in 802.11ac. To emulate the behavior of the OBSS problem throughout the experiment, the configuration changes made to AP1 were transmitted to AP2 through the wired connection, ensuring that both access points compete with each other when they operate on the same channel. This also helps to study the implication of getting the same configuration when the optimized algorithm runs at all the access points.

Initially, our goal was to understand the performance difference between two different routers. The results, plotted over time in Fig. 5, compared the performance between the two access points. The analysis revealed that AP2 consistently exhibited higher throughput than AP1. In the same scenario, different access points exhibit varying performance, which is influenced by the capabilities of each device.

The AP1 (Archer C20) demonstrates only minor variations in throughput across different configurations of transmission power and bandwidth. In addition, its throughput is quite volatile, responding dramatically to small changes in the WiFi environment. This variability complicates the analysis of performance differences in real-time scenarios.

In contrast, AP2 (Archer C60) shows a clear and significant

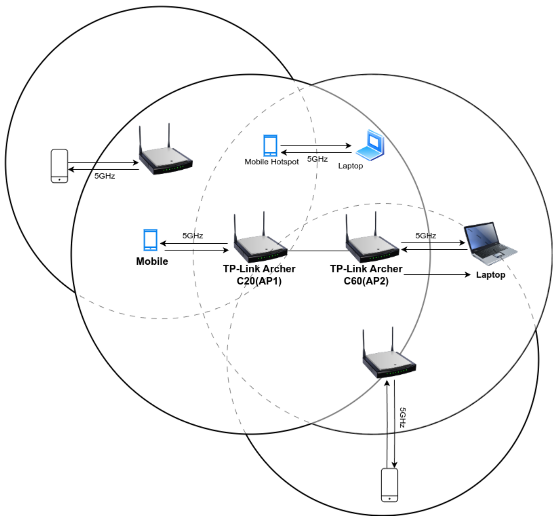

Fig. 4. Real-time setup for conducting the Modified Explore-First experiment with background traffic and two routers

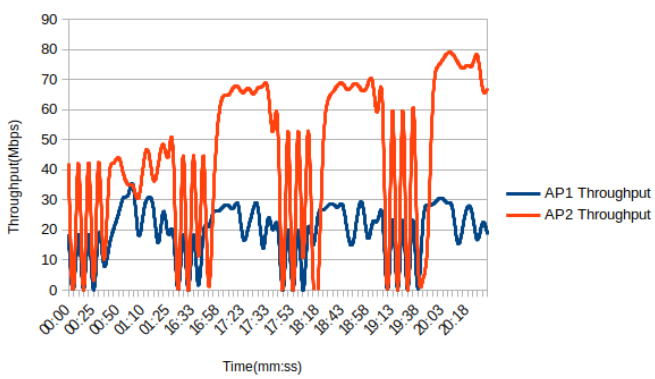

Fig. 5. Comparison of throughput of access points in the same configuration to understand the performance difference between two different routers

difference in throughput between 20- and 40-MHz bandwidth configurations, making it easier to analyze results and draw conclusions. Therefore, while we will present the results of both routers, interpreting the performance of the Archer C20 will pose challenges due to its less stable throughput.

Based on the observations mentioned above on different routers, we decided to compare the performance of each access point with their performance in different cases rather than comparing with other access points. Hence, throughout our analysis, the performance of each access point is compared with their performance rather than compared with the other access points.

VII. Experimental Results and Analysis

The experiments to study the performance of the modified exploration-first algorithm for OBSS and DCB problems have been formulated in four stages. Stage 1: An experiment was conducted to measure the throughput of both routers for 30 minutes. Here, throughput refers to the maximum throughput they can achieve without any traffic from the peer router and without the algorithm (that is, without optimal configuration). Stage 2: An experiment with traffic generated on the same channel to observe the effects of high traffic on throughput without the algorithm having been conducted for 30 minutes. Stage 3: An algorithm experiment was executed on the routers while simultaneously running traffic for 30 minutes. Stage 4: Finally, the algorithm was run on routers with traffic simultaneously for a longer period (one hour) to study the convergence aspect of the algorithm. The details are summarized in Table I.

TABLE I
Details of the Experiments Conducted in Four Stages

Stage	**with OBSS problem**	**with DCB**	**with Algorithm**
Stage 1 (for 30 minutes)	No	No	No
Stage 2 (for 30 minutes)	Yes	No	No
Stage 3 (for 30 minutes)	Yes	Yes	Yes
Stage 4 (for 1 hour)	Yes	Yes	Yes

Note:

1) Both AP1 and AP2 were running OpenWrt 22.03.3.
2) The OBSS problem has been created by invoking the traffic from other access points as mentioned in Fig. 4.
3) To enable the DCB, the UCI (Unified Configuration Interface) command of OpenWrt has been used.

The measured throughput (in Mbps) of AP1 and AP2 for the first three stages is presented in Fig. 6 and Fig. 7, respectively. The results of stage 4 are presented in Fig. 8 and Fig. 9, respectively.

As mentioned above, the throughput differences for various configurations of AP1 are minimal, as shown in Fig. 6, making it difficult to analyze any improvements. However, for AP2, the changes are noticeable, as in Fig. 7, which allows us to conclude that after 20 minutes of running the algorithm, it attempts to converge. For both AP1 and AP2, the algorithm (in stage 3) has increased throughput compared to throughput with the OBSS problem and without the algorithm (in stage 2). In addition, the algorithm has yielded higher throughput compared to the throughput of APs without the OBSS problem (in Stage 1). That means that the algorithm had suggested an optimal configuration (primary channel, transmission power, and bandwidth) continuously throughout the duration of the experiment according to the traffic and reward of the peers.

As depicted in Fig. 6, under ideal conditions without OBSS interference, AP1 maintains a stable throughput of 28 Mbps (in Stage 1) with minimal variation. However, under OBSS conditions, the throughput decreases to an average of 15 Mbps (in Stage 2), fluctuating between 7 Mbps and just below 28 Mbps. When the OBSS scenario is repeated with the algorithm applied on AP1, averaging 20 Mbps (in stage 3) with significant fluctuations and reaching as high as 35 Mbps.

For AP2, the ideal throughput averages 70 Mbps (in Stage 1), but under OBSS conditions, it decreases to an average of 60 Mbps (in Stage 2) and fluctuates between 70 and 5 Mbps. When the algorithm was applied, the average throughput increased to 62 (stage 3) Mbps, fluctuating between 25 and

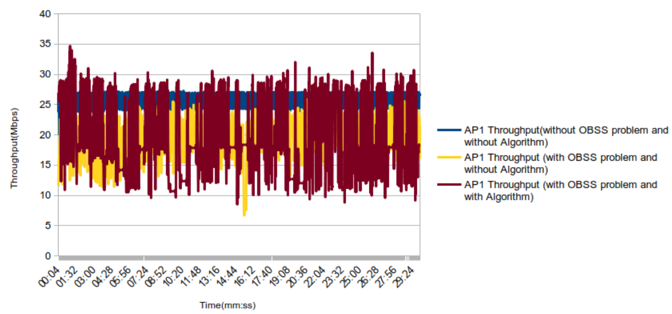

Fig. 6. Throughput measured during Modified Exploration-First algorithm for AP1

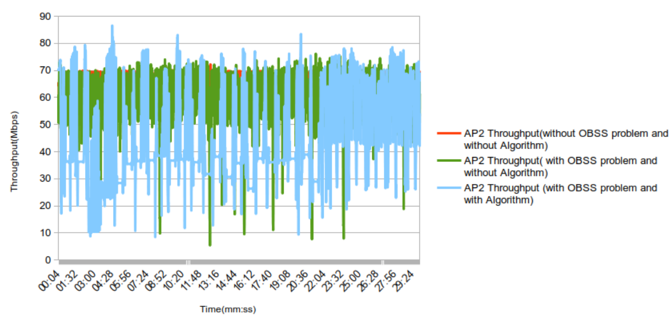

Fig. 7. Throughput measured during Modified Exploration-First algorithm for AP2

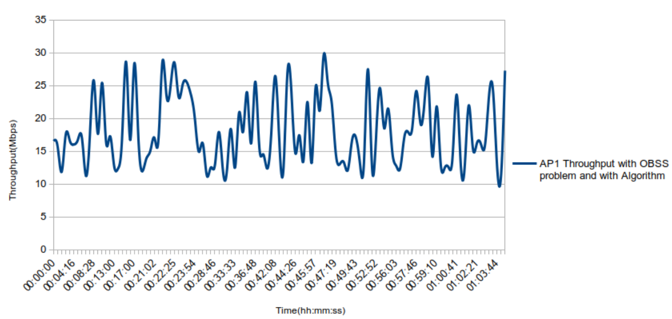

Fig. 8. Throughput measured during configuration changes with Modified Exploration-First algorithm for AP1

78 Mbps after the exploration phase. Fig. 7, shows a clear improvement in performance compared to the OBSS scenario.

We extended our experiments to approximately one hour to observe the convergence behavior of the algorithm as shown in Fig. 8 and Fig. 9 . After the exploration phase, we observed a trend towards convergence. AP2 consistently converges to a throughput of around 70 Mbps, with a few exceptions, while AP1 converges within the range of 10 to 25 Mbps. A convergence to a specific value is not possible in a dynamic environment. Hence, we focus on the convergence to a range of values. So we have achieved the convergence to a smaller range of throughput rather than the convergence to a specific value.

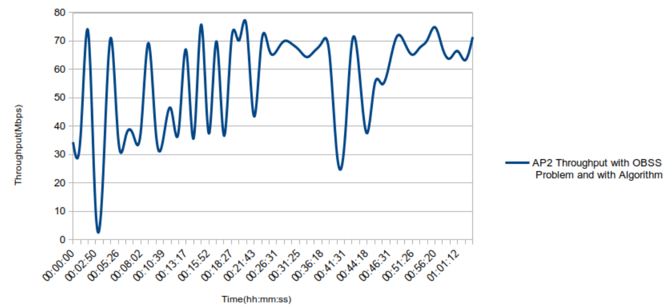

Fig. 9. Throughput measured during configuration changes with Modified Exploration-First algorithm for AP2

VIII. A Summary on Our Inferences on Design and Implementation Aspects of MAB Algorithm for OBSS and DCB Problem

In this section, we consolidate various design and implementation aspects that need to be considered in the design of the RL algorithm for the dynamic channel bonding problem.

1) The large number of possible configurations resulting from combinations of channel, transmission power, and bandwidth make selecting the best configuration very challenging. Therefore, we need an optimized methodology to apply the MAB algorithm to a subset of configurations rather than a large set of configurations.
2) Frequent configuration changes in access points can disrupt both the clients and the AP itself. A minimum interval between consecutive changes is necessary, as frequent adjustments can make it difficult for clients to stay connected and can exceed the ability of the AP to adapt.
3) Another challenge lies in the tendency of devices to rely on the same channels, leading to congestion while other channels remain underutilized. This makes it crucial to find optimal configuration settings that are suited to the environment, reducing the need for frequent changes.
4) The unpredictability of the radio environment presents another challenge when applying algorithms in WiFi, as the medium is highly unstable.
5) Maximizing the performance of the MAB algorithm without compromising client coverage is a significant challenge.
6) Reducing transmission power randomly can cause clients to drop out of range, making it essential to maintain their connection.
7) Technical problems can also prevent configuration changes from being implemented as expected, with access points sometimes failing to modify channel, transmission power, or bandwidth.
8) Access point limitations, such as restricted processing capabilities or a limited number of SSIDs that can be scanned, further affect the algorithm's performance.
9) During the exploration phase, fluctuations in channel, transmission power, and bandwidth can create a dynamic

environment, making it difficult for other access points to adjust and stabilize, contributing to overall system instability.

Among the list of points described above, the first three points have been addressed in the modified Exploration First algorithm. The rest of all other points will be addressed in our future work.

IX. Conclusion

We conducted a real-time investigation on the OBSS problem and DCB utilizing MAB algorithms. Most existing studies on this problem are being carried out in a simulated setting. Our proposed optimized exploration-first algorithm demonstrates better performance with reduced learning time. The results of the real-time experiments indicate substantial enhancements, underscoring the benefit of minimizing the duration of the exploration time, which theoretically reduces the convergence time. This work is not limited to the 802.11ac standard. This study will help us design an optimized algorithm for the 802.11ax standard. The logical next step is to validate our proposed algorithm in a large-scale WLAN. Future work will focus on resolving the identified unaddressed design aspects, such as the unpredictability of the radio environment, ensuring client coverage, the limited processing capabilities of access points, and the instability experienced by neighboring access points during the exploration phase.

References

[1] IEEE Standard for Information Technology—Telecommunications and Information Exchange Between Systems—Local and Metropolitan Area Networks—Specific Requirements—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, IEEE Std 802.11-2020 (Revision of IEEE Std 802.11-2016), 2021.

[2] IEEE Standard for Information Technology–Telecommunications and Information Exchange between Systems–Local and Metropolitan Area Networks–Specific Requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 4: Enhancements for Very High Throughput for Operation in Bands below 6 GHz, IEEE Std 802.11ac-2013 (Amendment to IEEE Std 802.11-2012), IEEE, 2013. doi: 10.1109/IEEESTD.2013.6687187.

[3] R. Sutton and A. Barto, Reinforcement Learning: An Introduction*, MIT Press, 2018.

[4] IEEE Standard for Information Technology—Telecommunications and Information Exchange between Systems Local and Metropolitan Area Networks—Specific Requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 1: Enhancements for High-Efficiency WLAN, IEEE Std 802.11ax-2021 (Amendment to IEEE Std 802.11-2020), IEEE, 2021, pp. 1–767. doi: 10.1109/IEEESTD.2021.9442429. Available: https://doi.org/10.1109/IEEESTD.2021.9442429.

[5] A. Bardou, T. Begin, and A. Busson, ”Improving the Spatial Reuse in IEEE 802.11ax WLANs: A Multi-Armed Bandit Approach,” in *Proceedings of the 24th International ACM Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, Alicante, Spain, 2021, pp. 135–144. doi: 10.1145/3479239.3485715. Available: https://doi.org/10.1145/3479239.3485715.

[6] A. Bardou, T. Begin, and A. Busson, ”Mitigating starvation in dense WLANs: A multi-armed Bandit solution,” *Ad Hoc Networks*, vol. 138, p. 103015, 2023. doi: 10.1016/j.adhoc.2022.103015. Available: https://www.sciencedirect.com/science/article/pii/S1570870522001871.

[7] A. Bardou and T. Begin, ”INSPIRE: Distributed Bayesian Optimization for ImproviNg SPatIal REuse in Dense WLANs,” in *Proceedings of the 25th International ACM Conference on Modeling Analysis and Simulation of Wireless and Mobile Systems*, Montreal, Quebec, Canada, 2022, pp. 133–142. doi: 10.1145/3551659.3559050. Available: https://doi.org/10.1145/3551659.3559050.

[8] P. E. Iturria-Rivera, M. Chenier, B. Herscovici, B. Kantarci, and M. Erol-Kantarci, ”Cooperate or Not Cooperate: Transfer Learning With Multi-Armed Bandit for Spatial Reuse in Wi-Fi,” *IEEE Transactions on Machine Learning in Communications and Networking*, vol. 2, pp. 351-369, 2024. doi: 10.1109/TMLCN.2024.3371929.

[9] J. Liu, Y. Liu, J. Zhang, X. Ge, A. Xu, and M. Zhao, ”A Bayesian Optimization Algorithm to Improve the Spatial Reuse in the Next-Generation WLANs,” in *2024 International Wireless Communications and Mobile Computing (IWCMC)*, 2024, pp. 1048–1053. doi: 10.1109/IWCMC61514.2024.10592439.

[10] S. Barrachina-Muñoz, A. Chiumento, and B. Bellalta, ”Multi-Armed Bandits for Spectrum Allocation in Multi-Agent Channel Bonding WLANs,” *IEEE Access*, vol. 9, pp. 133472–133490, 2021. doi: 10.1109/ACCESS.2021.3114430.

[11] F. Wilhelmi, J. Hribar, S. F. Yilmaz, E. Ozfatura, K. Ozfatura, O. Yildiz, D. Gündüz, H. Chen, X. Ye, L. You, and others, ”Federated spatial reuse optimization in next-generation Decentralized IEEE 802.11 WLANs,” *arXiv preprint arXiv:2203.10472*, 2022.

[12] W. Wydmański and S. Szott, ”Contention Window Optimization in IEEE 802.11ax Networks with Deep Reinforcement Learning,” in *2021 IEEE Wireless Communications and Networking Conference (WCNC), 2021, pp. 1–6. doi: 10.1109/WCNC49053.2021.9417575.

[13] W. Lee and J.-B. Seo, ”Deep Learning-Aided Channel Allocation Scheme for WLAN,” *IEEE Wireless Communications Letters*, vol. 12, no. 6, pp. 1007–1011, 2023. doi: 10.1109/LWC.2023.3257128.

[14] V. Ramesh, M. Maneesha, and S. Surendran, ”A reinforcement learning approach for improving internet connectivity in maritime network,” J. Adv. Res. Dyn. Control Syst., pp. 598-604, 2019.

[15] S. Simi and M. V. Ramesh, ”Intelligence in wireless network routing through reinforcement learning,” Int. J. Commun. Netw. Distrib. Syst., vol. 23, no. 2, pp. 231–251, 2019. Available: https://doi.org/10.1504/ijcnds.2019.101224