Journal of Theoretical and Applied Information Technology
31st August 2026. Vol.104. No.16

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

ENHANCING SECURITY AND ENERGY EFFICIENCY IN 5G HETNETS USING DEEP REINFORCEMENT LEARNING

SAI PRASANTH KANUPARTHY[1], MEDESWARA RAO KONDAMUDI[2], LAVURI SANKAR[3], S V KIRANMAYI SRIDHARA[4], ANUSHA BASAMSETTI[5], DR. K.B. GLORY[6], DR. SARALA PATCHALA[7], GARAGA SRILAKSHMI[8]

[1]Assistant Professor, Computer Science and Engineering (Data Science), RVR&JC College of Engineering, Guntur
[2]Assistant Professor, Computer Science & Engineering (Data Science), R.V.R.& J.C. College of Engineering, Chandramoulipuram,
[3]Dept of ECE, West Godavari Institute of Science and Engineering, AP, India.
[4]Department of ECE, Aditya University, Surampalem, AP, India.
[5]Department of ECE, Sasi Institute of Technology and Engineering, Tadepalligudem, AP, India.
[6]Assistant Professor, Department of Engineering English, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India.
[7]Professor, Department of ECE, KKR & KSR Institute of Technology and sciences, Guntur
[8]Assistant Professor, Department of Electronics & Communication Engineering, Aditya University, Surampalem, A.P., India

E-mail: [1]prasanth.bec42@gmail.com, [2]kmedeswararao@gmail.com, [3]sankar.lavuri2@gmail.com, [4]srivallikiranmayi@gmail.com, [5]anusha.b@sasi.ac.in, [6]kommalapatiglory@gmail.com, [7]saralajntuk@gmail.com, [8]srilakshmi1853@gmail.com

ABSTRACT

The paper presents a new method known as SecBoost. It has proposed to enhance the security aspect and energy consumption of the 5G network system. To achieve this, we employ deep reinforcement learning (DRL) in order to establish an optimised self-adaptive intelligent system that learns and optimises itself in response to the current conditions in real-time. In today's multi-level systems, various base stations offer access through the use of millimetre wave (mmWave) signals. These signals are however fast but they are prone to cracking by the benedict who may try to hack through the data. These are tough nut to crack especially using the traditional approaches to security because they compromise energy efficiency. To address issues the paper presents a learning-based solution called SecBoost to combat the discussed problem. SECBoost is developed through reinforcement learning, a kind of artificial intelligence tool that can make right decisions concerning the security and functionality of the network. In this, implemented a multi-agent reinforcement learning (MARL) approach. Power control determines the amount of power used for transmission, channel allocation selects the best frequency channels and beamforming directs signals to specific users. To handle the large and complex decision-making process, use a dueling deep Q-network (D3QN). SecBoost performs better than traditional methods. It is a strong candidate for future wireless security solutions. It provides both security and energy efficiency. This makes it useful for next-generation communication systems.
Keywords: *SECBoost, Reinforcement learning model, HetNets, Channel selection, Beamforming.*

1. INTRODUCTION

The number of Internet-of-Things (IoT) devices has increased rapidly in recent years[1]. With the rise of 5G and upcoming 6G networks connectivity demands have grown. By 2030, the number of IoT devices is expected to reach around 50 billion[2]. This rapid growth creates challenges for secure and efficient data transmission. To meet these challenges, heterogeneous networks (HetNets) have been developed. These networks improve capacity, coverage and energy efficiency[3]. The market for HetNets is expected to be worth over 51 billion dollars by 2027. While HetNets improve network performance, they introduce new security risks. In these networks, devices exchange data[4]. This

makes it easier for attackers to steal information. Traditional encryption methods help but are not always effective in 5G networks. Public-key cryptography needs complex management[5]. Advances in computing, like quantum computers will break current encryption methods. This has created a need for new security techniques. One promising approach to security in 5G HetNets is called Physical Layer Security (PLS)[6]. This method does not rely on encryption. Instead, it uses the properties of wireless signals to make communication more secure. The novelty of PLS was introduced with the concept of secrecy capacity[7]. It is the difference between the data rate of a legitimate user and the data rate that an eavesdropper capture. If this difference is large, the communication is secure[8]. Several techniques have been used to improve PLS, including secure beamforming, cooperative jamming and physical layer authentication. However, many existing methods do not work well in dynamic 5G environments.

Many current security techniques for HetNets rely on knowing the exact wireless channel conditions[9]. However, in a real-world environment these conditions change constantly. It is difficult to measure and track all channel information accurately. Techniques like cooperative jamming consume more power[10]. Classical optimization techniques have been used to improve security but they are often slow and inefficient in large networks. Because of these challenges, new methods are needed to provide secure and energy-efficient communication in 5G HetNets[11]. We therefore introduced SecBoost aimed at addressing all these problems. It is a multi-agent cooperative deep reinforcement learning framework for enhancing security and energy efficiency in 5G HetNets. SecBoost aims at co-optimising the power control, subchannel allocation and beamforming[12]. This means that the number of users must strive to be optimally secret while also striving to minimize the power. Another attractive feature of SecBoost is that it does not depend on the knowledge of the wireless channel, which is preferable in dynamic scenarios[13]. The system adopted a reinforcement learning (RL) method for training the interaction model. In RL, an agent acquires experience through the interaction with the environment and gets a reward depending on those actions[14]. The system employs a D3QN to make decisions and select optimal learning experiences by flagging the most important ones. Hence the following are the main findings of the paper;

1. This paper focuses on the secrecy degree analysis of the mmWave channels within the two-tier heterogeneous network. It is used for both legitimate users and snoops or interceptors. We cast the optimization problem to maximize the secrecy energy efficiency (SEE) of picocells.

2. A MARL framework is also used for the optimization of power control, channel assignment and beamforming. To achieve this, RL agents are deployed at each base station to be trained in their respective organizations.

3. To identify the challenges and design a SecBoost cooperative DRL-based scheme. It incorporates the D3QN model in the learning process this is to help increase learning rate. Experience repetition in the system is not plain, but rather prioritized to repeat the most important experiences.

4. The paper also presents a comparison of the work done by SecBoost to the best of the methods prevalent in the current world. Among them are MARL, MA-DQN as well as the traditional beamforming-based secrecy enhancement techniques.

5. The outcomes are also showing the fact that SecBoost can enhance the security measures noticeably and at the same time, minimize the required power to do so. This makes it suitable for use in future 5G and beyond networks as will be discussed in the paper.

The main aim of this study is to design a learning-based framework that improves both security and energy efficiency in 5G heterogeneous networks. The proposed SecBoost model uses multi-agent deep reinforcement learning to jointly optimize power control, channel allocation, and beamforming. The novelty of this work lies in combining multi-agent reinforcement learning with a dueling deep Q-network and prioritized experience replay to improve secrecy energy efficiency without requiring full channel state information. The performance of the proposed model is evaluated using key outcome measures such as secrecy energy efficiency, secrecy rate, power consumption, and learning convergence. The results show that SecBoost performs better than existing methods in all these aspects.

This paper proposes a novel scheme named, SecBoost, which employs the concept of reinforcement learning to boost the security and energy efficiency of the 5G HetNet. In traditional approaches of security, there are constraints when dealing with dynamic application making reinforcement learning a viable solution[15]. Scope

and Assumptions of the Work: This work focuses on improving secrecy energy efficiency in a two-tier 5G heterogeneous network consisting of one macrocell and multiple picocells operating in sub-6 GHz and mmWave bands, respectively. The proposed SecBoost framework is designed for dynamic wireless environments where exact channel state information is not fully available.

The study assumes that base stations are equipped with reinforcement learning agents capable of observing local states such as channel conditions, user distribution, and interference levels. The network model considers the presence of multiple legitimate users and passive eavesdroppers. It is also assumed that the agents can coordinate in a cooperative multi-agent learning framework.

Limitations of the work: Despite its advantages, the proposed SecBoost framework has certain limitations. The model assumes ideal synchronization among agents and does not explicitly consider signaling overhead or communication delays between base stations. The computational complexity of deep reinforcement learning may increase with network size, which can impact real-time deployment. These assumptions and limitations provide a controlled framework for evaluating the effectiveness of SecBoost while highlighting directions for future research.

RELATED WORK

Advanced technology like the 5G network has created interest in exploring and deploying more effective communication systems with high energy efficiency. HetNets have emerged as a promising solution to cater to the increasing number of connected devices. Heterogeneous networks (HetNet) are networks that consist of various types of base stations, the power of which can vary. These are macros, micros, picos and femtos [16]. Still, each is significant in its own way in ensuring the operation of the network runs as smoothly as possible. Macrocells are used for providing broad coverage using high transmit power base stations. Such stations are sited at such higher grounds to ensure that their signals reach out to the targets with clarity. Microcell is a lower transmitter power than that of main cell stations which is applicable in areas of high traffic such as shopping malls. They cover a relatively smaller area but are useful in controlling high user traffic [17]. Therefore, picocells can be considered intermediary between microcells and femtocells. They work in a licensed frequency bandwidth and are used in building and aircraft. These cells assist in the transfer of heavy traffic from the macro cells in order to ease congestion [18]. Femtocells are portable base stations that enhance the cellular service within homes or small businesses. They have very low power capabilities and mainly used in a few connections [19].

The classical approaches to optimization prove incapable of efficiently managing the HetNet environment, which is rather dynamic. Classical solutions assume prior knowledge of the state of the network, which cannot be easily determined in real-time [20]. RL offers a promising alternative. RL enables the network components to decide and adapt on their own based on past experience eventually enhancing confirmation on the subsequent decisions made. It means that when HetNets are deployed, they can modify themselves in line with environmental changes without necessarily rely on models which have been developed previously [22]. The MDP determinates how an agent behaves in a given environment in order to make the best decisions. The state is the representation of the current status of the network and the action highlighting the control logic. This constant communication ensures that the network policies are continually being developed [21]. RL is strengthened by DRL since complex decision functions can be trained with the help of neural networks. In HetNets, the use of DRL plays an important role in power control, channel selection as well as beamforming. This leads to enhanced power to power ratio and security [23]. In the present paper, we propose SecBoost, a DRL-based algorithm that simultaneously adjusts these parameters. With the help of the D3QN, the proposed method enhances the stability and speed of the learning process [24].

PLS is a promising approach that protects wireless transmissions without relying on encryption [25]. It uses signal processing techniques like secure beamforming and cooperative jamming to enhance confidentiality. Cooperative jamming introduces interference in a way that disrupts eavesdroppers but maintains communication for authorized users [26]. However, these methods require accurate channel state information (CSI). From the above literature review, it is observed that existing methods face several limitations. Many approaches require accurate channel state information, which is difficult to obtain in real-time environments. Some techniques improve security but increase power consumption, while others focus on energy efficiency but do not provide strong security. Based on these limitations, the main problem addressed in this work is how to jointly improve security and

energy efficiency in a dynamic heterogeneous network without relying on full channel knowledge. To address this problem, the following research questions are considered:

1. How can reinforcement learning be used to optimize power control, channel allocation, and beamforming in HetNets?
2. How to improve secrecy energy efficiency while minimizing power consumption?

SYSTEM MODEL

The elements of the system model illustrate how the network topology is composed and how data flows through it while being well secured and passed through optimally. The model is a two-tier HetNet. It consists a macrocell in the sub 6G Hz frequency range and several picocells in the millimeter wave frequency range. This are to improve efficiency in usage of energy and also prevent unauthorized access to communication. The system will therefore have a macrocell, multiple picocells, legitimate users and eavesdroppers. The Figure 1 represents a wireless communication network with multiple components. It includes a macro base station, pico base stations, primary users, mobile users and eavesdroppers. The macro base station provides communication links to mobile users. Similarly, the pico base stations communicate with primary users. The network experiences different types of interference namely co-channel interference and interference from the macro base station. The presence of eavesdroppers highlights security concerns, as they attempt to intercept communications between the base stations and users. The network consists of multiple small communication zones, each containing a pico base station, primary user and an eavesdropper. The pico base stations serve local users while coexisting with the macro base station. The eavesdroppers present a threat by attempting to capture confidential transmissions. The diagram highlights interference issues in the system requiring optimized transmission techniques for better security and efficiency. This structure reflects the challenges in managing interference, confirming secure communication and balancing transmission power in modern wireless networks.

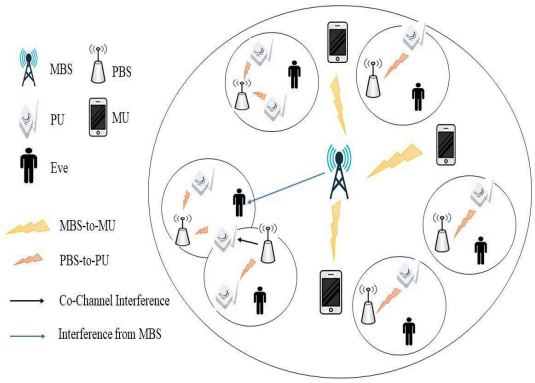

Figure 1: System Architecture

The macrocell base station (MBS) and picocell base stations (PBSs) communicate with the respective users. The objective is to maximize the SEE of the network. This is done by optimizing power control, channel allocation and beamforming. The macrocell uses a sub-6 GHz channel model. The channel vector between the MBS and a user k is represented as:

$$h_M^k = \sqrt{x_M^k}\, u_M^k \tag{1}$$

Here, x_M^k represents the path loss component and u_M^k represents the small-scale fading component. The path loss x_M^k is given by:

$$x_M^k = \left(\frac{c}{4\pi f_c}\right)^2 (R_M^k)^{-\alpha_M} \tag{2}$$

Here, c is the speed of light, f_c is the carrier frequency. R_M^k is the distance between the MBS and user k and α_M is the path loss exponent. For picocells operating in the mmWave band, the channel vector between a PBS and user p is:

$$h_P^{np} = \sqrt{X_P^{np}}\, a_s(\phi_P^{np})\, a_u(\beta_P^{np}) \tag{3}$$

Here, X_P^{np} represents the path loss. a_s and a_u represent the steering vectors of the transmitter and

receiver, respectively. The path loss model for mmWave is given as:

$$X_P^{np} = \begin{cases} c_l (R_P^{np})^{-\alpha_l}, \text{LoS link} \\ c_n (R_P^{np})^{-\alpha_n}, \text{NLoS link} \end{cases} \quad (4)$$

Here, c_l and c_n are path loss coefficients. α_l and α_n are path loss exponents for line-of-sight (LoS) and non-line-of-sight (NLoS) conditions, respectively. The received signal at a macrocell user k is:

$$y_M^k = h_M^k w_k s_k + \sum_{j \neq k} h_M^k w_j s_j + n_k \quad (5)$$

Here, w_k is the beamforming vector for user k, s_k is the transmitted signal. n_k is the additive white Gaussian noise (AWGN). Similarly, the received signal at a picocell user p in picocell n is:

$$y_P^{np} = h_P^{np} w_{np} s_{np} + \sum_{i \neq p} h_P^{np} w_{ni} s_{ni} + \sum_{l \neq n} \sum_{i} h_P^{lp} w_{li} s_{li} + n_{np} \quad (6)$$

Here, the terms represent the interference from other users and other picocells. For an eavesdropper in picocell n, the received signal is:

$$y_P^{nE} = h_P^{nE} w_{np} s_{np} + \sum_{i \neq p} h_P^{nE} w_{ni} s_{ni} + \sum_{l \neq n} \sum_{i} h_P^{lE} w_{li} s_{li} + h_M^{nE} \sum_{k} w_k s_k + n_{nE} \quad (7)$$

Here, the additional term h_M^{nE} represents the interference from the macrocell base station. The secrecy rate of a picocell user p in picocell n is given by:

$$R_P^{np,sec} = \left[\log_2(1+\gamma_P^{np}) - \log_2(1+\gamma_P^{nE}) \right]^+ \quad (8)$$

Here, γ_P^{np} and γ_P^{nE} are the signal-to-interference-plus-noise ratios (SINR) at the legitimate user and the eavesdropper, respectively. The total secrecy rate of all picocell users is:

$$R_P^{Total,sec} = \sum_n \sum_p \left[\log_2(1+\gamma_P^{np}) - \log_2(1+\gamma_P^{nE}) \right]^+ \quad (9)$$

The total power consumption of all picocells is given by:

$$P_{Total} = \zeta \left[\sum_k \sum_p \| w_{np} \|^2 + \sum_l \sum_i \| w_{li} \|^2 \right] + N(N_P P_a + P_b) \quad (10)$$

Here, ζ is the power amplifier coefficient, P_a is the power consumed per antenna and P_b is the base power consumption. The objective is to maximize the SEE, which is the ratio of the secrecy rate to the power consumption:

$$\max_{w,P} \frac{R_P^{Total,sec}}{P_{Total}} \quad (11)$$

It is subjected to power and SINR constraints. The system model describes a two-tier HetNet with a macrocell and multiple picocells. The objective is to optimize power, channel selection and beamforming to maximize secrecy and energy efficiency. The reinforcement learning approach helps in making intelligent decisions under dynamic conditions.

2. SECBoost: MULTI-AGENT COOPERATIVE DRL BASED SEE MAXIMIZATION

In this paper, a deep reinforcement learning (DRL) based framework SecBoost is proposed for enhancing the SEE of 5G HetNets. This leverages a multi-agent cooperative Deep Reinforcement Learning to solve the power control, channel allocation, and beamforming problem. The method provides legitimate user with high transmission rate to channel high data rate while at the same time denying the unauthorized user from tapping into signal. SecBoost addresses these issues since it gets updated with new information and conditions of the network constantly. Reduced power consumption enhances the efficiency status without a compromise on power as well. The problem involves optimizing beamforming and channel allocation. Proper beamforming directs signals effectively, reducing interference and improving signal quality. Efficient channel allocation gives smooth data flow in the

network. Traditional optimization techniques require perfect knowledge of channel conditions, which is difficult to obtain in real-time. RL provides a powerful alternative by allowing the system to learn optimal strategies through trial and error. MARL enables multiple base stations (BSs) to collaborate, improving overall network efficiency. The problem is modeled as a MDP defined by the tuple:

$$(S, A, r, P, \gamma) \tag{12}$$

Here, S represents the state space that includes channel conditions, user locations and power levels. A denotes the action space comprising power control, beamforming and channel allocation decisions. r is the reward function that quantifies the improvement in SEE. P is the transition probability, defining how actions influence future states. γ is the discount factor that balances short-term and long-term rewards. Each agent in SecBoost learns an optimal policy π using Q-learning. The Q-function is given by:

$$Q(s,a) = E\left[r + \gamma \max_{a'} Q(s', a')\right] \tag{13}$$

Here, $Q(s,a)$ is the expected cumulative reward for taking action a in state s. E represents the expectation over all possible future states. γ is the discount factor, controlling the impact of future rewards. s' is the next state after taking action a. a' is the action that maximizes the Q-value in the next state. Agents update the Q-values iteratively using:

$$Q(s,a) \leftarrow Q(s,a) + \alpha\left[r + \gamma \max_{a'} Q(s', a') - Q(s,a)\right] \tag{14}$$

Here, α is the learning rate that controls how quickly the model updates its knowledge. SecBoost improves upon traditional Q-learning by using a D3QN. This method separates state-value estimation and advantage estimation, leading to better learning efficiency. The Q-value function is split into:

$$Q(s,a) = V(s) + A(s,a) - \frac{1}{|A|}\sum_{a'} A(s,a') \tag{15}$$

Here, $V(s)$ is the value function, representing the overall quality of state s. $A(s,a)$ is the advantage function, showing the benefit of taking action a over the average action. |A| is the total number of actions. D3QN reduces overestimation errors and speeds up training. It is particularly useful in large state-action spaces like 5G HetNets. SecBoost uses prioritized experience replay (PER) to focus on important experiences during training. Instead of sampling transitions uniformly, it assigns higher priority to transitions with larger temporal difference (TD) errors:

$$p_i = \frac{|\delta_i| + ò}{\sum_j (|\delta_j| + ò)} \tag{16}$$

Here, p_i is the probability of selecting experience i. δ_i is the TD error for experience i. $ò$ is a small constant to prevent zero probability. PER improves sample efficiency and accelerates convergence. SecBoost targets to maximize the SEE of the network is given in (11). It is subjected to:

$$\sum_k \| w_k \|^2 \le P_M, \quad \sum_p \| w_p \|^2 \le P_P \tag{17}$$

$$\gamma_M^k \ge \gamma_M^{min}, \quad \gamma_P^p \ge \gamma_P^{min} \tag{18}$$

Here, $R_P^{Total,sec}$ is the total secrecy rate of the network. P_{Total} is the total power consumption. P_M and P_P are power constraints for macrocells and picocells. γ_M^{min} and γ_P^{min} are minimum SINR requirements. SecBoost provides a scalable and efficient solution for maximizing SEE in 5G HetNets. By combining multi-agent DRL, D3QN and PER it achieves superior performance compared to traditional methods. Simulation results show that SecBoost improves security and energy efficiency while adapting to dynamic network conditions.

Algorithm 1: Multi-Agent Cooperative Q-Learning for SEE Maximization

1. Initialize Q-table $Q(s,a)$ for each agent with random values

2. Set learning rate α, discount factor γ and exploration rate ò
3. for each episode do
4. Reset environment and observe initial state s
5. for each time step do
6. for each agent i in PBSs do
7. Select action a_i using ò-greedy policy
8. Execute action and observe reward r_i and new state s'
9. Update Q-value:
10. $Q(s,a_i) \leftarrow Q(s,a_i) + \alpha[r_i + \gamma \max_{a'} Q(s',a') - Q($
11. end for
12. Update environment with new state s'
13. end for
14. Reduce ò to balance exploration and exploitation
15. end for

3. PERFORMANCE EVALUATION

The performance evaluation of SecBoost is necessary to understand its effectiveness in providing secure and energy-efficient communication. This section examines the SEE, secrecy rate, power consumption and learning convergence of SecBoost compared to other traditional techniques. The evaluation uses a MARL framework in which base stations learn to optimize network parameters dynamically. The comparison includes SecBoost, MADQN, MADRL, JBF-SEEM and O-EDT methods. A two-tier HetNet is simulated comprising one MBS and multiple PBSs. Each PBS serves multiple legitimate users, while eavesdroppers attempt to intercept communication. Table 1 provides the key parameters used in the simulation.

Table 1: Simulation Parameters

Parameter	Value
Carrier Frequency	28 GHz
Bandwidth	100 MHz
Number of Macro BSs	1
Number of Pico BSs	5
Number of Users per PBS	10
Number of Eavesdroppers	3
Transmission Power (MBS)	46 dBm
Transmission Power (PBS)	30 dBm
Noise Power	-174 dBm/Hz
Path Loss Exponent	3.5
Learning Rate	0.001
Discount Factor	0.9

The simulation runs for 10,000-time steps allowing reinforcement learning agents to adapt and learn optimal policies for power control, beamforming and channel selection. SEE measures how well a system maintains security while minimizing power consumption. It is calculated as:

$$SEE = \frac{\sum R_P^{Total,sec}}{P_{Total}} \quad (19)$$

Here, $R_P^{Total,sec}$ is the total secrecy rate. P_{Total} represents total power consumption. Table 2 compares the SEE performance of SecBoost with other methods.

Table 2: SEE Comparison (bps/Hz/W)

Method	SEE
SecBoost (Proposed)	12.5
MADQN	9.8
MADRL	8.7
JBF-SEEM	7.3
O-EDT	6.1

SecBoost achieves the highest SEE due to its reinforcement learning-based optimization of power and beamforming. The cooperative learning approach enhances network efficiency and security. The secrecy rate represents the difference in data rates between legitimate users and eavesdroppers. Higher secrecy rates indicate better protection against security breaches. The secrecy rate comparison is shown in Table 3.

Table 3: Secrecy Rate Comparison (bps/Hz)

Method	Secrecy Rate
SecBoost (Proposed)	5.7
MADQN	4.9
MADRL	4.3
JBF-SEEM	3.8
O-EDT	3.2

SecBoost achieves the highest secrecy rate due to its dynamic adaptation to network conditions and

interference management. Power consumption is an important metric in energy-efficient communication. The total power consumption of different methods is compared in Table 4.

Table 4: Total Power Consumption Comparison (W)

Method	Power Consumption
SecBoost (Proposed)	75
MADQN	85
MADRL	92
JBF-SEEM	105
O-EDT	120

From the above analysis, SecBoost effectively reduces the power consumption while still achieving high secrecy rate. From the performance comparison, it can be seen that SecBoost performs better than the other methods with respect to all the evaluation metrics. SecBoost enhances the level of secrecy energy efficiency better than traditional methods. Thus, it is possible to conclude that SecBoost can be named the promising solution for the further development of the secure and energy-effective wireless communication systems of the new generation.

Figure 2 shows a comparison of the outcomes of five figures namely SecBoost, MADQN, MADRL, JBF-SEEM and O-EDT in terms of secrecy energy efficiency. While comparing all the models, it is clear that SecBoost demonstrates the highest performance. From around 4 it begins to rise and levels off around 11 for values of 'n' between 1000 and 1000. Like PARIC, MADQN exhibits a similar trend, but it has a lesser percentage gain and reaches a plateau of around 8. In terms of performance, MADRL achieves a slightly lower value compared to MADQN and stabilizes at approximately 7. JBF-SEEM begins where MADRL begins but ends at close to 6. This means that the DEM-EDT model has the highest secrecy energy efficiency, while the O-EDT model has the lowest. It rises slightly at around 3 and steadily begins to rise but never crosses the 5 marks even after 1000 iterations. Observing all models become better with iterations, but the results show that SecBoost perform best among all. This is evident by comparing the performance of SecBoost with other three models and established that the model has the potential of improving the secrecy energy efficiency over time.

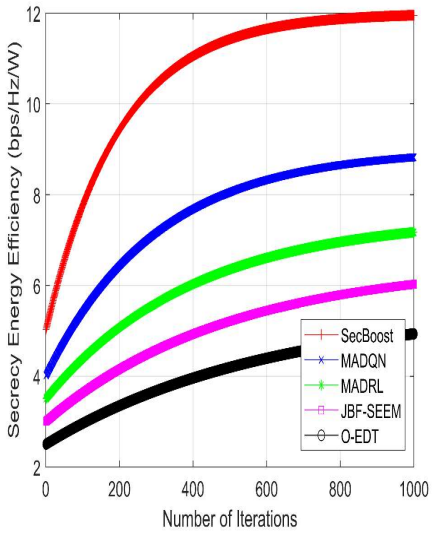

Figure 2: SEE versus Number of Iterations

The Figure 3 illustrates the relationship between the number of users and secrecy rate for five different models: SecBoost, MADQN, MADRL, JBF-SEEM and O-EDT. SecBoost achieves the highest secrecy rate among all models. It starts at about 6.5 bps/Hz with very few users. It gradually declines as the number of users increases. When the number of users reaches 50, it drops to around 5.5 bps/Hz. MADQN follows a similar pattern but at a lower level. It begins at about 5.5 bps/Hz and decreasing to around 4.5 bps/Hz. The MADRL model performs slightly worse than MADQN. It starts at about 5 bps/Hz and gradually declining to 3.5 bps/Hz. JBF-SEEM begins at approximately 4.2 bps/Hz and drops to around 3 bps/Hz when the number of users reaches 50. The O-EDT model shows the lowest secrecy rate. It starts at about 3.5 bps/Hz and steadily declines, reaching negative values close to -1 at the highest user count. The decreasing trend in all models suggests that as the number of users increases, the secrecy rate declines. SecBoost maintains the best performance throughout, indicating its effectiveness in handling security challenges in larger networks.

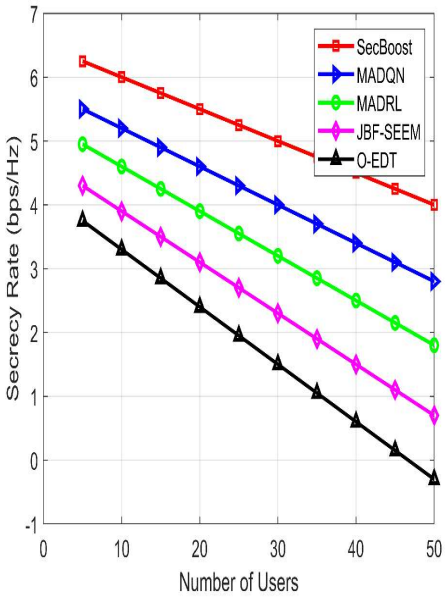

Figure 3: Secrecy Rate versus Number of Users

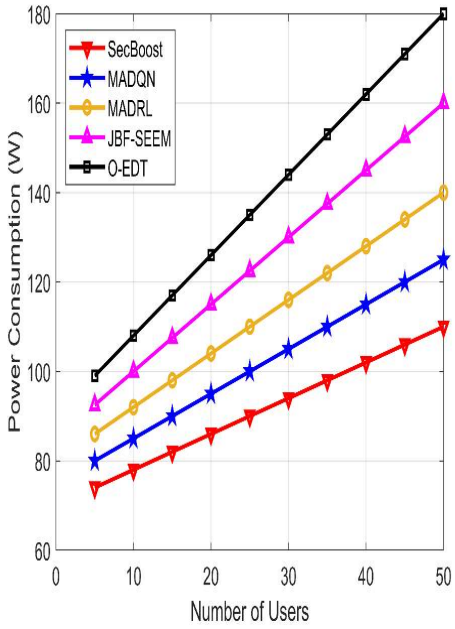

Figure 4: Power Consumption versus Number of Users

The Figure 4 presents the power consumption of five models as the number of users increases. The O-EDT model consumes the most power. It starts at around 100W and increases steadily. As the number of users grows, it reaches nearly 180W. JBF-SEEM follows a similar trend. Its power usage starts near 90W and rises to about 160W. MADRL shows lower power consumption than JBF-SEEM. It starts at around 85W and increasing to 140W. MADQN performs slightly better with power usage beginning at 80W and gradually reaching around 120W. SecBoost is the most efficient model. It starts at about 75W and rises to approximately 110W at 50 users. The increasing trend in all models suggests that as the number of users grows, power consumption rises. However, SecBoost remains the most energy-efficient, while O-EDT consumes the most power. It makes less efficient in managing user growth.

The Figure 5 illustrates the impact of the number of eavesdroppers on the secrecy rate for five models: SecBoost, MADQN, MADRL, JBF-SEEM and O-EDT. SecBoost achieves the highest secrecy rate among all models. It starts at around 6 bps/Hz with no eavesdroppers. As the number of eavesdroppers increases, it decreases gradually. It reaches about 3 bps/Hz when there are 10 eavesdroppers. MADQN follows a similar trend. It starts slightly lower at around 5.5 bps/Hz. It decreases to approximately 2.5 bps/Hz at the highest number of eavesdroppers. The MADRL model starts at about 5 bps/Hz and steadily declines to around 2 bps/Hz. JBF-SEEM follows the same pattern but starts at a lower value of around 4 bps/Hz, dropping to just below 1 bps/Hz. The O-EDT model has the lowest secrecy rate among all models. It begins at approximately 3.5 bps/Hz and declines rapidly. It reaches negative values around -1.5 bps/Hz when the number of eavesdroppers reaches 10.

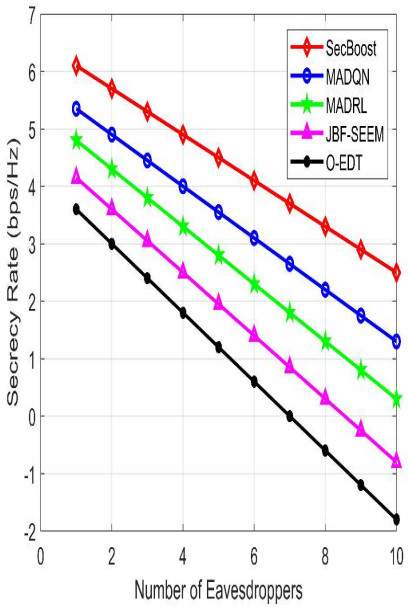

Figure 5: Secrecy Rate versus Eavesdropper Density

The Figure 6 compares the cumulative reward progression of five models over multiple iterations. SecBoost, achieves the highest cumulative reward, rising quickly from zero and stabilizing near 10 as the number of iterations reaches 1000. MADQN follows closely but peaks at around 8. MADRL exhibits slightly lower performance, reaching approximately 7 by the end of the iterations. JBF-SEEM follows a similar pattern but reaches only around 6. O-EDT has the lowest cumulative reward. It increases gradually but does not exceed a value of 5. The general trend suggests that all models improve with more iterations, but at different speeds. SecBoost shows the best performance. It demonstrates the highest learning efficiency. O-EDT remains the least effective, indicating slower learning progress and lower overall performance. The gap between the models widens as iterations increase, suggesting that some models learn at a much faster rate. The results highlight the advantage of SecBoost in optimizing learning and maximizing cumulative rewards over time.

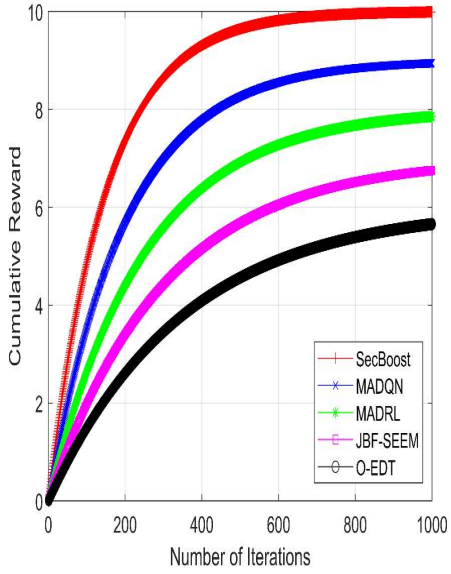

Figure 6: Convergence Curve: Reward versus Iterations

The Figure 7 illustrates the relationship between transmission power and secrecy energy efficiency for five different models: SecBoost, MADQN, MADRL, JBF-SEEM and O-EDT. SecBoost achieves the highest secrecy energy efficiency. It starts at around 15 bps/Hz/W when transmission power is 10 dBm. It gradually declines to about 10 bps/Hz/W at 50 dBm. MADQN follows a similar pattern but starts lower at around 12 bps/Hz/W and decreases to nearly 7 bps/Hz/W. The MADRL model begins at approximately 10 bps/Hz/W. It drops to around 5 bps/Hz/W as transmission power increases. JBF-SEEM starts at about 7 bps/Hz/W and declines steadily. It reaches negative values near -3 bps/Hz/W at the highest transmission power. O-EDT has the lowest secrecy energy efficiency. It starts at around 6 bps/Hz/W and decreases sharply, reaching -10 bps/Hz/W at 50 dBm. The downward trend in all models suggests that increasing transmission power leads to a decline in secrecy energy efficiency. SecBoost maintains the highest performance, while O-EDT shows the worst efficiency. It indicates poor adaptability to higher transmission power levels.

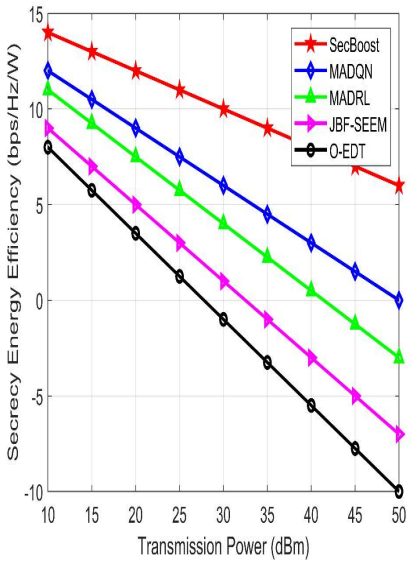

Figure 7: SEE versus Transmission Power

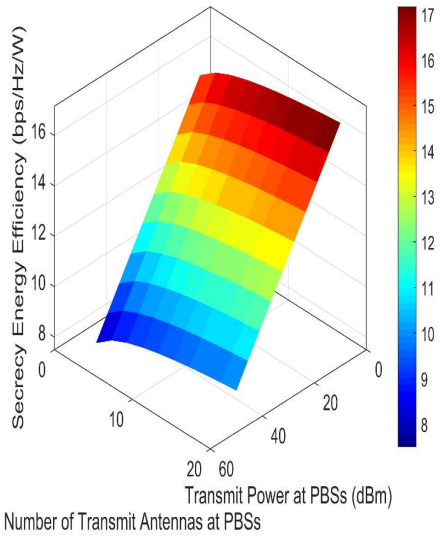

Figure 8: SEE Region versus Transmit Power at PBSs versus Transmit Antennas

The Figure 8 (3D surface plot) represents the relationship between secrecy energy efficiency, transmit power at PBSs (dBm) and the number of transmit antennas at PBSs. A color gradient is used to represent efficiency levels with blue indicating lower values and red representing higher values. The plot shows a steady increase in secrecy energy efficiency as both transmit power and the number of antennas increase, reaching its peak near 17 bps/Hz/W. The trend in the graph suggests that improving transmit power and adding more antennas enhance secrecy energy efficiency. The transition from blue to red indicates a smooth rise in efficiency as both parameters increase. However, the slope starts to flatten at higher values. This means additional increases in transmit power provide smaller gains. More antennas give diminishing improvements. This suggests that after a certain threshold, efficiency gains slow down. The results emphasize the need to balance power and antenna usage to optimize secure communication. Increasing resources improves performance, but beyond a certain point, the benefit becomes minimal.

The Figure 9 presents the effect of different learning rates (LR) on cumulative reward over multiple iterations. The graph includes five different learning rates: 0.01, 0.05, 0.1, 0.2 and 0.5. Each learning rate follows an increasing trend, with cumulative reward rising rapidly at the beginning and then leveling off as it reaches a maximum. The LR = 0.5 converges the fastest, reaching its maximum reward within a short number of iterations. The LR = 0.2 and LR = 0.1 lines follow closely, stabilizing at a high cumulative reward. The LR = 0.05 takes slightly longer to reach its peak. The LR = 0.01 converges the slowest. The results suggest that higher learning rates lead to faster convergence. As all models eventually achieve similar maximum cumulative rewards. The initial steep increase indicates rapid learning in early iterations, with diminishing returns as training progresses. The differences in convergence speed highlight the trade-off in selecting an appropriate learning rate. A very high learning rate converge quickly but leads to instability, while a very low learning rate take longer to reach optimal performance.

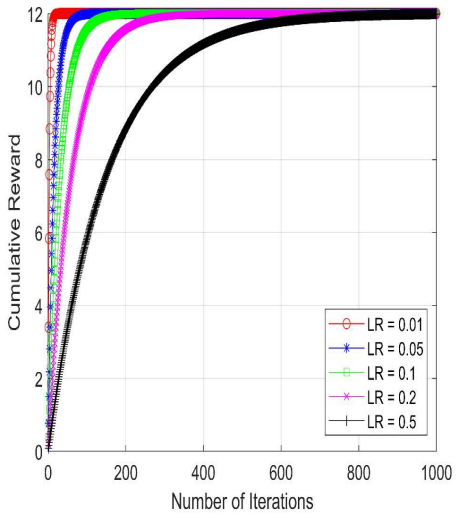

Figure 9: Reward Performance Analysis

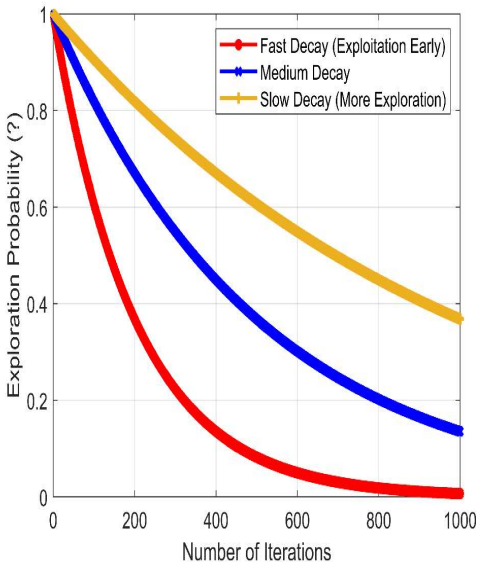

Figure 10: Exploration versus Exploitation (ε – Greedy Strategy)

The Figure 10 illustrates how exploration probability decreases over iterations for three different decay strategies: fast decay, medium decay and slow decay. The fast decay shows a steep drop in exploration probability. It falls below 0.2 after 500 iterations and approaching zero by the end. The medium decay decreases at a slower rate keeping a moderate exploration level for a longer period before stabilizing. The slow decay maintains a high exploration probability for more iterations. Even after 1000 iterations, it remains above 0.3. It indicates continuous exploration. The overall trend shows that different decay strategies influence the balance between exploration and exploitation. Fast decay encourages early exploitation of learned patterns. The slow decay allows extended exploration before settling on decisions. Medium decay provides a balanced approach between the two. Selecting the right decay method is crucial in reinforcement learning. A fast decay leads to premature decisions. A slow decay allows more exploration. However, it will delay reaching optimal performance.

The Figure 11 presents the selection frequency of different action types in a system. Beamforming has the highest selection frequency, reaching over 400. Power control follows, with a frequency slightly above 350. Channel allocation is next, with a selection frequency of around 300. Handover management has the lowest selection frequency, staying below 300. The results suggest that beamforming is the most preferred action in this system. Power control is frequently selected. It indicates its importance in managing network performance. Channel allocation is used moderately, while handover management is the least selected.

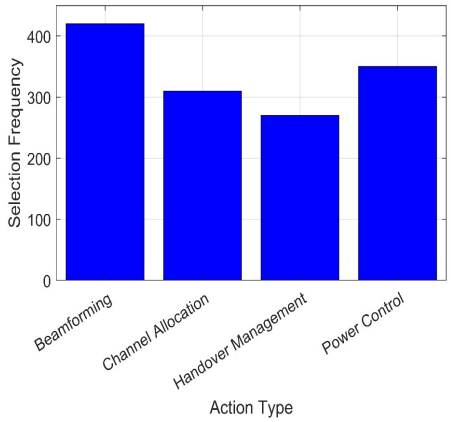

Figure 11: Action Selection Frequency in RL

The Figure 12 shows the progression of Q-value estimates for four action types—power control, beamforming, channel allocation and handover management—over multiple iterations. The power control shows the highest growth. It begins at approximately 4 and rises quickly, stabilizing near 20 as the iterations reach 1000. The beamforming follows a similar upward trend but peaks at around 18. The channel allocation increases steadily but stabilizes at a lower Q-value of around 14. The handover management starts at the lowest value and climbs gradually. It reaches about 12 by the end of the iterations. These results indicate that power control and beamforming contribute the most to performance improvements. The channel allocation and handover management have lower Q-value estimates.

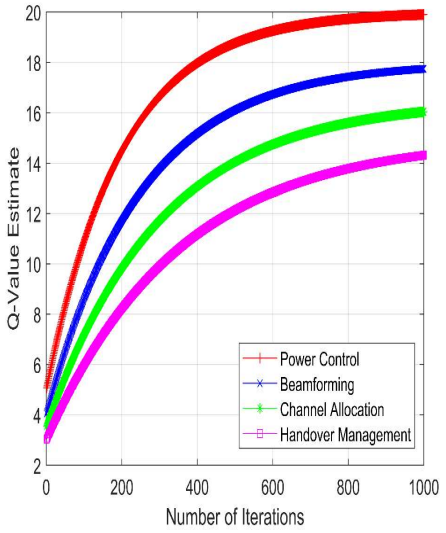

Figure 12: Q-Value Progression over Time in RL

The Figure 13 illustrates the relationship between energy efficiency and data rate for four different models: SecBoost, MADQN, JBF-SEEM and O-EDT. SecBoost achieves the highest energy efficiency among all models. It starts at approximately 25 bits/Joule when the data rate is very low. As the data rate increases, it gradually decreases. It reaches around 3 bits/Joule at 100 Mbps. MADQN follows a similar trend. It starts slightly lower at about 22 bits/Joule. It gradually declines to around 2.5 bits/Joule at the highest data rate. The JBF-SEEM model begins at around 18 bits/Joule. It follows a steady downward trend. It drops to approximately 2 bits/Joule at 100 Mbps. O-EDT has the lowest energy efficiency throughout. It starts at about 16 bits/Joule. It declines at a similar rate. It reaches below 2 bits/Joule at the highest data rate.

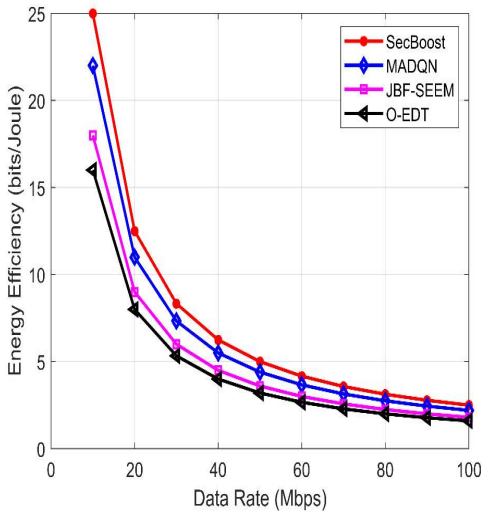

Figure 13: Energy Efficiency versus Data Rate

The Figure 14 presents the signal-to-interference-plus-noise ratio (SINR) as a function of distance from the base station for five models. All models exhibit a decreasing trend in SINR as the distance from the base station increases. SecBoost achieves the highest SINR at all distances. It starts at around 30 dB near the base station. It declines gradually as the distance increases. At 200 meters, it reaches about 0 dB. MADQN follows closely behind, starting at approximately 27 dB and reducing to around -2 dB. The MADRL model starts at about 24 dB and declines steadily to approximately -5 dB. The JBF-SEEM begins with an SINR value of nearly 20 dB and drops to around -10 dB. O-EDT has the lowest SINR. It starts at about 18 dB and decreases sharply. It reaches nearly -30 dB at 200 meters. The overall trend suggests that as the distance increases SINR drops significantly for all models. SecBoost maintains the highest SINR values across all distances. O-EDT performs the worst. This indicates that SecBoost is more effective in maintaining signal quality over longer distances. O-EDT experiences significant performance degradation as distance increases.

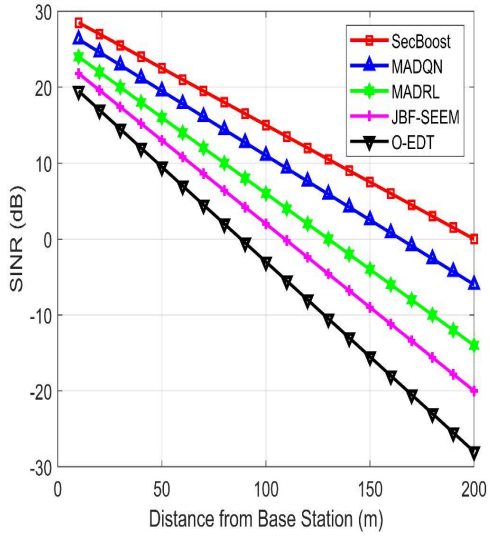

Figure 14: SINR versus Distance from Base Station

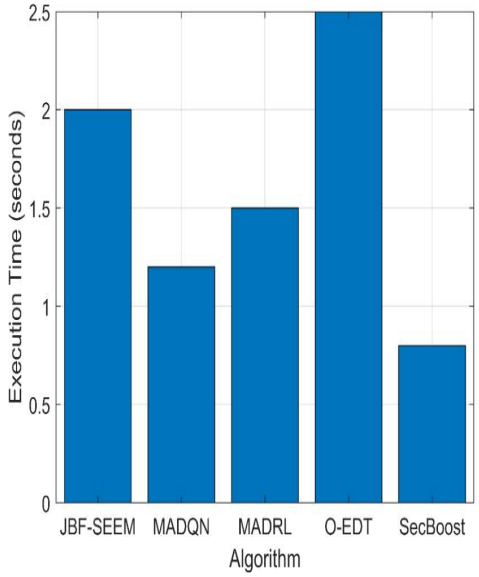

Figure 15: Execution Time Comparison

Figure 15 shows the execution time of five algorithms. MADQN, MADRL, O-EDT, SecBoost and JBF-SEEM. The execution time of O-EDT is the longest among the algorithms, which is 2.5 seconds. JBF-SEEM follows, requiring about 2 seconds. With execution time of around 1.5 seconds, MADRL has also a slightly lower time to execute. With an execution time of roughly 1.2s MADQN performs better. The fastest algorithm is SecBoost, with an execution time of below 1 second. Results show that SecBoost runs in the fastest time. Specifically, O-EDT is slowest showing a higher computational burden. In comparison, the execution time for MADQN and SecBoost is shorter than that of JBF-SEEM and MADRL. Execution time differences show how computational complexity affects performance. There are some algorithms that will supply better results, but that would take more time to run. It is mentioned that SecBoost has lower execution time, which implies that

The Figure 16 shows the probability of secure communication as a function of jamming power for five different models: SecBoost, MADQN, MADRL, JBF-SEEM and O-EDT. As jamming power increases, the probability of secure communication declines for all models. SecBoost achieves the highest probability of secure communication. It starts at 1 when jamming power is 0 dBm and gradually decreases. It maintains a higher probability compared to the other models. Even at 20 dBm, SecBoost retains a probability above 0.2. MADQN follows a similar decreasing pattern but remains lower than SecBoost. It reaches a probability slightly above 0.1 at 20 dBm. MADRL starts high but declines more steeply ending just above 0.1. The JBF-SEEM model follows a similar trend but has a lower probability across all jamming power levels. O-EDT performs the worst with the probability of secure communication dropping rapidly. At 20 dBm, it is close to 0. O-EDT is the least effective. The decreasing trend appears across all models. Higher jamming power significantly reduces secure communication. This highlights the need for strong security mechanisms.

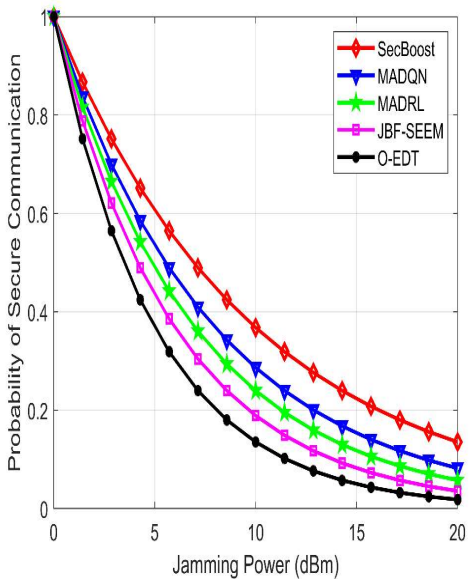

Figure 16: Secure Communication Probability versus Jamming Power

4. CONCLUSION

This paper presented SecBoost, a multi-agent cooperative DRL framework designed to improve SEE in heterogeneous 5G networks. The work is aim to enhance network security while minimizing power consumption by optimizing power control, channel selection and beamforming. Traditional security methods struggle to handle dynamic network conditions and increased energy consumption. SecBoost overcomes these challenges using reinforcement learning techniques. The proposed framework applies a D3QN model combined with PER. This combination achieves efficient learning and fast convergence while dynamically adjusting to network conditions. Unlike traditional methods, SecBoost does not need predefined optimization models. It continuously learns and adapts. This makes it more suitable for real-world applications. The framework reduces power consumption, making it a more sustainable solution for next-generation networks. The cooperative learning approach in SecBoost enables base stations to work together. The proposed SecBoost framework demonstrates its novelty by integrating multi-agent reinforcement learning with D3QN and prioritized experience replay for joint optimization of security and energy efficiency. The simulation results show that SecBoost achieves a secrecy energy efficiency of 12.5 bps/Hz/W, which is higher than existing methods such as MADQN (9.8) and O-EDT (6.1). It also reduces power consumption to 75 W while maintaining a higher secrecy rate of 5.7 bps/Hz. It improves overall network performance. The integration of reinforcement learning eliminates the need for accurate CSI that is often difficult to obtain in real-world conditions. Future research enhances SecBoost by incorporating federated learning techniques. Although SecBoost improves security and energy efficiency, some challenges remain. Future work can focus on reducing the computational complexity of deep reinforcement learning models for real-time implementation. Lightweight DRL models or edge-based learning can be explored to improve scalability.

REFERENCES:

[1] F. Xia, L. T. Yang, L. Wang, A. Vinel, et al., "Internet of things," International journal of communication systems, vol. 25, no. 9, p. 1101, 2012.

[2] A. S. Andrae, "New perspectives on internet electricity use in 2030," Engineering and Applied Science Letters, vol. 3, no. 2, pp. 19–31, 2020. DOI: 10.30538/psrp-easl2020.0038

[3] S. F. Yunas, M. Valkama, and J. Niemel¨a, "Spectral and energy efficiency of ultra-dense networks under different deployment strategies," IEEE Communications Magazine, vol. 53, no. 1, pp. 90–100, 2015. DOI: 10.1109/MCOM.2015.7010521

[4] A. Zhang, J. Chen, R. Q. Hu, and Y. Qian, "Seds: Secure data sharing strategy for d2d communication in lte-advanced networks," IEEE Transactions on Vehicular Technology, vol. 65, no. 4, pp. 2659–2672, 2015. DOI: 10.1109/TVT.2015.2416002

[5] S. Rana, F. K. Parast, B. Kelly, Y. Wang, and K. B. Kent, "A comprehensive survey of cryptography key management systems," Journal of Information Security and Applications, vol. 78, p. 103607, 2023. https://doi.org/10.1016/j.jisa.2023.103607

[6] L. Sun and Q. Du, "Physical layer security with its applications in 5g networks: A review," China communications, vol. 14, no. 12, pp. 1–14, 2017. DOI: 10.1109/CC.2017.8246328

[7] E. Horn, "Logics of political secrecy," Theory, Culture & Society, vol. 28, no. 7-8, pp. 103–122, 2011. https://doi.org/10.1177/0263276411424583

[8] Y. Liang, H. V. Poor, and S. Shamai, "Secure communication over fading channels," IEEE Transactions on Information Theory, vol. 54, no.

6, pp. 2470–2492, 2008. DOI: 10.1109/TIT.2008.921678

[9] J. Wang, Z. Yan, H. Wang, T. Li, and W. Pedrycz, "A survey on trust models in heterogeneous networks," IEEE Communications Surveys & Tutorials, vol. 24, no. 4, pp. 2127–2162, 2022. DOI: 10.1109/COMST.2022.3192978

[10] Y. Wu and Y. Huo, "A survey of cooperative jamming-based secure transmission for energy-limited systems," Wireless Communications and Mobile Computing, vol. 2021, no. 1, p. 6638405, 2021. https://doi.org/10.1155/2021/6638405

[11] S. Buzzi, I. Chih-Lin, T. E. Klein, H. V. Poor, C. Yang, and A. Zappone, "A survey of energy-efficient techniques for 5g networks and challenges ahead," IEEE Journal on selected areas in communications, vol. 34, no. 4, pp. 697–709, 2016. DOI: 10.1109/JSAC.2016.2550338

[12] H. Sharma, N. Kumar, I. Budhiraja, and A. Barnawi, "Secrecy rate maximization in thz-aided heterogeneous networks: A deep reinforcement learning approach," IEEE Transactions on Vehicular Technology, vol. 72, no. 10, pp. 13490–13505, 2023. DOI: 10.1109/TVT.2023.3275968

[13] M. Khani, S. Jamali, M. K. Sohrabi, M. M. Sadr, and A. Ghaffari, "Slice admission control in 5g cloud radio access network using deep reinforcement learning: A survey," International Journal of Communication Systems, vol. 37, no. 13, p. e5857, 2024. DOI: 10.1002/dac.5857

[14] O. Abul, F. Polat, and R. Alhajj, "Multiagent reinforcement learning using function approximation," IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 30, no. 4, pp. 485–497,2000. DOI: 10.1109/5326.897075

[15] M. Sewak, S. K. Sahay, and H. Rathore, "Deep reinforcement learning in the advanced cybersecurity threat detection and protection," Information Systems Frontiers, vol. 25, no. 2, pp. 589–611, 2023. https://arxiv.org/abs/2206.02733

[16] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski," Five disruptive technology directions for 5G," IEEE Communications Magazine, vol. 52, no. 2, pp. 74–80, 2014. DOI: 10.1109/MCOM.2014.6736746

[17] J. G. Andrews et al.," What will 5G be?" IEEE Journal on Selected Areas in Communications, vol. 32, no. 6, pp. 1065–1082, 2014. DOI: 10.1109/JSAC.2014.2328098

[18] A. Ghosh et al.," Heterogeneous cellula networks: From theory to practice," IEEI Communications Magazine, vol. 50, no. 6, pp. 54-64, 2012. DOI: 10.1109/MCOM.2012.6211486

[19] V. Chandrasekhar et al.," Femtocell networks: A survey," IEEE Communications Magazine, vol 46, no. 9, pp. 59–67, 2008. DOI 10.1109/MCOM.2008.4623708

[20] N. C. Luong et al.," Applications of deep reinforcement learning in communications and networking: A survey," IEEE Communications Surveys Tutorials, vol. 21, no. 4, pp. 3133–3174, 2019. DOI: 10.1109/COMST.2019.2916583

[21] R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction, 2nd ed. Cambridge, MA, USA: MIT Press, 2018. DOI: 10.1109/TNN.1998.712192

[22] L. Lei, Z. Zhong, K. Zheng, J. Chen, and H. Meng, "Challenges on wireless heterogeneous networks for mobile cloud computing," IEEE Wireless Communications, vol. 20, no. 3, pp. 34–44, 2013. DOI: 10.1109/MWC.2013.6549281

[23] B. Duo, Q. Wu, X. Yuan, and R. Zhang, "Energy efficiency maximization for full-duplex uav secrecy communication," IEEE Transactions on Vehicular Technology, vol. 69, no. 4, pp. 4590–4595, 2020. DOI: 10.1109/TVT.2020.2977948

[24] H. Hu, D. Wu, F. Zhou, X. Zhu, R. Q. Hu, and H. Zhu, "Intelligent resource allocation for edge-cloud collaborative networks: A hybrid ddpg-d3qn approach," IEEE Transactions on Vehicular Technology, vol. 72, no. 8, pp. 10696–10709, 2023. DOI: 10.1109/TVT.2023.3253905

[25] F. Xu, S. Ahmad, M. Ahmed, S. Raza, F. Khan, Y. Ma, W. U. Khan, et al., "Beyond encryption: Exploring the potential of physical layer security in uav networks," Journal of King Saud University-Computer and Information Sciences, vol. 35, no. 8, p. 101717, 2023. https://doi.org/10.1016/j.jksuci.2023.101717

[26] M. Yang, B. Zhang, Y. Huang, N. Yang, D. Guo, and B. Gao, "Secure multiuser communications in wireless sensor networks with tas and cooperative jamming," Sensors, vol. 16, no. 11, p. 1908, 2016. https://doi.org/10.3390/s16111908