Energy-Efficient Secure Transmission Techniques for 5G-Enabled HetNet

A

Thesis submitted

for the award of the degree of

DOCTOR OF PHILOSOPHY

By

Himanshu Sharma
(901803020)

Under the guidance of

Dr. Neeraj Kumar
(Professor, CSED)

Dr. Raj Kumar Tekchandani
(Assistant Professor, CSED)

Computer Science and Engineering Department
Thapar Institute of Engineering and Technology
Patiala - 147004, India

October 2023

CERTIFICATE

This is to certify that the Thesis entitled "**Energy-Efficient Secure Transmission Techniques for 5G-Enabled HetNet**", submitted by **Himanshu Sharma** (901803020), a research scholar in the *Computer Science and Engineering Department, Thapar Institute of Engineering & Technology, Patiala*, for the award of the degree of **Doctor of Philosophy**, is a record of an original research work carried out under the supervision of Dr. Neeraj Kumar, Dr. Raj Kumar Teckchandani and refers work of other researchers which are duly listed in reference section. The Thesis has fulfilled all requirements as per the regulations of the Institute and in our opinion has reached the standard needed for submission.

The results embodied in this Thesis have not been submitted to any other University or Institute for the award of any degree.

Himanshu

Himanshu Sharma
Registration No. 901803020

This is to certify that the above statement made by the candidate is correct and true to the best of my knowledge and belief.

Date:
Place: Patiala

Dr. Neeraj Kumar
Professor,
Computer Science and Engineering Department,
Thapar Institute of Engineering and Technology, Patiala,
India - 147004

Rajkumar

Dr. Raj Kumar Tekchandani
Assistant Professor,
Computer Science and Engineering Department,
Thapar Institute of Engineering and Technology, Patiala,
India - 147004

ACKNOWLEDGEMENTS

First and foremeost, thanks to almighty God for all his blessings without which nothing of my work would have been possible. The successful completion of any task would be incomplete without acknowledging the people who made it possible. I would like to take this opportunity to express my gratitude to all those who made this journey possible. Words are often too less to express one's deepest regards, but lets give it a go.

I would like to express my sincere gratitude to my supervisors, Dr. Neeraj Kumar (Professor, CSED), and Dr. Raj Kumar Tekchandani (Assistant Professor, CSED), who have supported me throughout my Ph.D. work with their patience, motivation, enthusiasm and immense knowledge. Apart from providing me with excellent supervision, strong cooperation and constant encouragement throughout this journey, they also shared their invaluable experiences with me to succeed in life. They have truly been the source of real inspiration for me and I will always remain indebted to them.

I also take the opportunity to thank head of deaprtment Dr. Shalini Batra, the former head Dr. Maninder Singh for providing me the necessary administrative assistance and infrastructure that helped me in the completion of my research work. I would also like to extend my sincere thanks to the doctoral committee members Dr. Anil Kumar Verma (Professor, CESD), Dr. Shreelekha Panday (Assistant Professor, CSED), and Dr. Sudhanshu Tyagi Verma (Assistant Professor, ECED) for their helpful suggestions and ensuring the progress of my research work regularly.

My deepest gratitude to my soul mate Mrs. Gitika Sharma for being a pillar of support and encouragement throughout my research work. Her constructive criticism, involvement with my work, and critical reading of the text, helped me tremendously in improving upon the thesis. I would also like to pay my sincere regards to my parents and in-laws family for their constant motivation and support.

The chain of my gratitude will definitely be incomplete if I forget to thank my father Rajesh Sharma, mother Ansuya Sharma and brother Shubham Sharma, for their unconditional love, support and encouragement in every phase of my life. It was their confidence in me that I started my Ph.D. Since then, the journey of Ph.D. has been a sweet and bitter ride at times which lead to a special mention for my mother who was with me through thick and thin, and gave me courage at the times when I felt really low. Her continuous motivation showed me the silver lining in the dark clouds.

I would also like to thank my friends and colleagues with whom I have traveled this journey of research. A special thanks to my Ph.D fellow Dr. Ishan Budhiraja and Dr. Rajat Chaudhary for their guidance and always being there as elder brothers. I would especially thank my friends Krishan, Niyaz and Deepak for their much needed moral and emotional support. As one cannot mention the names of all well-wishers, friends and beloved ones, I would like to pay my regards to one and all who supported me during the journey of knowledge.

Himanshu Sharma

List of Publications

Journal Publications (SCI/SCIE):

1) **H. Sharma**, N. Kumar, and R. Tekchandani, "Physical Layer Security using Beamforming Techniques for 5G and beyond Networks: A Systematic Review," Physical Communication, Elsevier. https://doi.org/10.1016/j.phycom.2022.101791 (Impact Factor- 2.39).

2) **H. Sharma**, N. Kumar and R. Tekchandani, "Mitigating Jamming Attack in 5G Heterogeneous Networks: A Federated Deep Reinforcement Learning Approach," in IEEE Transactions on Vehicular Technology, 2022, doi: 10.1109/TVT.2022.3212966. (Impact Factor- 6.24).

3) **H. Sharma**, N. Kumar and R. K. Tekchandani, "SecBoost: Secrecy-Aware Deep Reinforcement Learning Based Energy-Efficient Scheme for 5G HetNets," in IEEE Transactions on Mobile Computing, doi: 10.1109/TMC.2023.3235429. (Impact Factor- 6.075)

ABSTRACT

Heterogeneous Networks (HetNets) play an essential role in enhancing the quality-of-service (QoS) for end-users by increasing the spectral efficiency of the network and reducing the power consumption of user equipment (UE). With an exponential increase in the number of Internet of Things devices (IoTD), data traffic flow demands, and the complex network structure of 5G, the HetNets are also growing rapidly to increase the spectral efficiency of the wireless network. The market size of HetNets is expected to reach about 51.1 billion USD by the year 2027 as compared to 18.3 billion USD in the year 2020 with a compound annual growth rate (CAGR) of 15.2%. In contrast to traditional homogeneous networks, HetNets allow small cells to collaborate in macrocell networks, which increases the possibility of spatial resource reuse and improves the quality-of-service (QoS) for user equipment (UE).

However, the dynamic and distributed nature of HetNets makes them susceptible to various types of attacks (e.g., eavesdropping, jamming). Also, new technologies of 5G such as Massive MIMO, mmWave, NOMA brings unique security concerns to 5G HetNets, which were not present in pre-5G HetNets. Implementing traditional security techniques such as access control, encryption, and network security seems to be insufficient for 5G HetNets and their inherent vulnerabilities. Also, HetNet's architecture is more open and varied than traditional single-tier cellular networks, making information sharing more vulnerable to security threats. Thus, designing and implementing effective eavesdropping countermeasures is essential for secure wireless transmissions in 5G HetNets. Although, cryptography-based solutions have been widely used to provide network security at the upper levels. But, these solutions are limited in their ability to meet the security needs of 5G-and-beyond networks due to the following constraints i) It is extremely difficult to use cryptographic approaches using public keys in large, decentralized networks ii) Public-key infrastructure (PKI) has remained unbreakable until now in light of the usage of extremely long key pairs; however, advances in computing power, such as strong quantum systems, can now crack the cryptographic keys.

PLS is based on the fundamentals of information theory and focuses on the security of propagation channel. It can be used in the 5G HetNets to efficiently degrade signal transmission efficiency at unauthorized receivers and applications to prevent them from obtaining sensitive data from the received signal. It ensures safe and efficient communications, even though eavesdroppers (illegitimate smart devices) are fitted with powerful computational devices in these networks. Some of the commonly

used anti-eavesdropping techniques in HetNets include secure beamforming, cooperative jamming, and physical layer authentication (PLA). Beamforming (BF) is one of the promising PLS techniques to solve the issues mentioned above. At the transmitters and receivers, BF matrices can be used to shape the beam patterns of antennas to maximize a specific security parameter, such as signal-to-interference noise ratio (SINR), secrecy rate. Direct-sequence spread spectrum (DSSS) and frequency hopping (FH) techniques have been widely adopted as antijamming strategies in literature to mitigate the aforementioned issues. Particularly, FH is a sophisticated and commonly used technique which allows user equipment (UEs) to change their operating frequency to another frequency spectrum, which in turn avoids malicious jamming assaults. Besides DSSS and FH, power control is also an effective anti-jamming technique. Jamming cognition, decision-making, and joint optimization of beamforming and power allocation are the three fundamental phases of the power control aided anti-jamming communication cycle

Although the existing proposals include efficient usage of various signal-processing PLS techniques, but these techniques suffer from the following constraints: i) Most of the pre-existing PLS techniques require prior and accurate knowledge of channel state information (CSI) values for effective PLS designs. However, it is difficult to obtain instantaneous global CSI of dynamic HetNet, since global CSI often varies frequently ii) Usage of large number of relays and active antennas in PLS techniques increases the power consumption. Also, cooperative jamming and broadcasting artificial noise require increased transmit power to achieve perfect secrecy iii) Most of the above mentioned studies utilize classical optimization techniques to optimize beamforming and power allocation vectors, which are less efficient in dynamic large-scale networks. Also, FH and DSSS based anti-jamming techniques are constrained by their inherent dependency on pre-shared secrets (i.e., spreading codes and hopping sequences) between the communicating parties. Also, by using intelligent radio devices such as software-defined radio (SDR), the jammers can work together to block the wireless channels and disrupt the transmissions of FH-based UEs. Moreover, eavesdropping attacks and spectrum sensing on the control channel of 5G boost the jamming strength of malicious jammers in FH-based HetNets.

As evidenced by widespread use of AI in different application areas of PLS, it is undoubtedly one of the most necessary elements for enhancing the PLS of 5G HetNets. It can be used to learn about normal and aberrant behaviors of HetNets based on how users and base stations communicate with one another. AI techniques can successfully anticipate future new instances by learning from existing instances. AI techniques can also be used to forecast new attacks, which are usually mutations of previous attacks. AI has been used in various PLS applications such as security ori- ented beamforming, cooperative jamming, PLA, secure handover schemes, etc.

Deep learning (DL) methods, in general, rely on training and experience to improve task completion performance. This learning approach, which is a subset of machine learning (ML), analyzes the data for categoriza- tion and decision-making without programming. Also, Reinforcement learning (RL) algorithms can be used to design an optimal policy using the Markov decision process (MDP). Although RL-based techniques are viable solutions in designing PLS schemes, but the use of Q-learning method for the large state and action spaces suffers from stagnant learning speed, which may result in performance degradation of PLS techniques. DL has recently been integrated into RL techniques, allowing them to tackle a wide range of complicated problems. Deep reinforcement learning (DRL) is a set of approaches for estimating value functions or policy functions using deep neural networks. It uses Markov decision models to help choose between several actions based on state transition models. DRL have recently piqued the interest of the research community in designing intelligent PLS techniques for wireless networks.

In this research work, the following schemes have been proposed to rectify the aforementioned issues:

- Firstly, we propose a secrecy-aware energy-efficient scheme for a two-tier heterogeneous network (HetNet), consisting of a sub-6 GHz macrocell and multiple millimeter wave (mmWave) picocells. Each picocell is assumed to have several users and an eavesdropper (Eve) which intercepts the signal of the picocell users. In the proposed scheme, firstly, to maximize the secrecy energy-efficiency (SEE) of picocell users, a joint optimization problem of power control, channel allocation, and beamforming is formulated by considering the minimum secrecy rate and signal-to-interference-plus-noise ratio (SINR) constraints. Due to the non-convex nature of the aforementioned optimization problem in a highly dynamic HetNet environment, we transform it into a reinforcement learning (RL) problem using the Markov decision process (MDP). Then, a multi-agent reinforcement learning (MARL) technique is used to obtain the maximum long-term reward. Moreover, we propose a multi-agent cooperative deep reinforcement learning (DRL) scheme known as *SecBoost* to solve the MDP with large number of action and state spaces. It uses the dueling and double-Q architecture of dueling double deep Q-network (D3QN) to optimize power control, channel allocation, and beamforming vectors to maximize the SEE of picocells. Also, prioritized experience replay is used to increase the sampling efficiency of *SecBoost*. The SEE performance of *SecBoost* is compared with MARL, multi-agent deep Q-network (MA-DQN), state-of-the-art joint beamforming based secrecy energy efficiency maximization (JBF-SEEM) scheme, and one-time pad based encrypted data transmission (O-EDT). Simulation results demonstrated that the proposed *SecBoost*

scheme achieves 14.7%, 8.33%, 30%, and 69% better average SEE in comparison to MARL, MA-DQN, JBF-SEEM, and O-EDT schemes, respectively, which reveals its effectiveness in improving SEE of picocells.

- Further,we propose a federated deep reinforcement learning (DRL) based anti-jamming technique for two-tier 5G HetNets. In the proposal, each femtocell of 5G HetNets is assumed to have multiple single antenna femto users (FUs) and a multi-antenna jammer used to jam the downlink signals from femto base station (FBS) to FUs. Aiming to improve the achievable rate at FUs in the presence of jammers, a joint optimization problem of beamforming and power allocation at FBSs is formulated by considering the quality-of-service (QoS) requirements of FUs. Due to the non-convex nature of the aforementioned optimization problem, we have used the Markov decision process (MDP) to transform the optimization problem into a multi-agent reinforcement learning (MARL) problem. Then, to solve this MDP with large number of states and action spaces, a federated deep reinforcement learning (DRL) scheme is proposed to maximize the achievable rate at FUs. The proposed scheme uses federated learning and dueling architecture of dueling double deep Q network (D3QN) to optimize the beamforming vectors and power allocation jointly at FBSs. The achievable rate performance of the proposed federated DRL scheme is compared with double deep Q network (DDQN) and deep Q network (DQN). Simulation results show that the proposed federated DRL scheme achieves 19.39% and 23.85% better achievable rate in comparison to DDQN and DQN schemes.

Contents

List of Figures

List of Tables

List of Major Acronyms

AI	Artificial Intelligence
HetNets	Heterogeneous Networks
5G	Fifth Generation Network
ML	Machine Learning
PLS	Physical Layer Security
Eve	Eavesdropper
BF	Beamforming
CSI	Channel State Information
CJ	Cooperative Jamming
CC	Channel Coding
CNN	Convolution Neural Network
RNN	Recurrent Neural Network
LSTM	Long Short Term Memory
GAN	Generative Adversarial Network
DBN	Deep Belief Network
RL	Reinforcement Learning
DL	Deep Learning
DRL	Deep Reinforcement Learning
SNR	Signal to Noise Ratio
mmWave	milli-meter wave
SINR	Signal to Interference Noise Ratio
CR	Cognitive Radio
IoV	Internet of Vehicles
UAV	Unmanned Aerial Vehicles
NOMA	Non Orthogonal Multiple Access
SWOT	Strength-Weakness-Opportunities-Threats
MIMO	Multiple Input Multiple Output
IoT	Internet of Things
CAGR	Compound Annual Growth Rate
SOP	Secrecy Outage Probability
CC	Channel Coding

LDPC	Low Density Parity Check
AN	Artificial Noise
MBS	Macro Base Station
NR	New Radio
BER	Bit Error Rate
MSE	Mean Squared Error
SE	Spectral Efficiency
UEs	User Equipments
PLA	Physical Layer Authentication
IoV	Internet of Vehicles
D2D	Device-to-device
QoS	Quality-of-service
QoE	Quality-of-experience
MITM	Man-in-the-middle
QNN	Quantized Neural Network
RQNN	Reccurent Quantized Neural Network
OFDMA	Orthogonal Frequency-Division Multiple Access
MISO	Multiple Input Single Output
SWIPT	Simaltaneous Wireless Information and Power Transfer
LTE	Long-Term Evolution
MIMO	Multiple Input Multiple Output

Chapter 1

Introduction

A rapid increase in the usage of intelligent devices, base stations (BS), and the exuberant multimedia content have contributed to an exponential increase in data traffic in the recent years. According to the report of International Telecommunication Union Radiocommunication Sector (ITUR) [1], the Global mobile traffic data will be around 4.394k- Exabytes (M2M traffic not included) and around 5.016k- Exabytes (M2M traffic included) by the year 2030. Modern wireless applications such as e-healthcare, augmented and virtual reality (AR/VR), tactile internet, connected robotics and autonomous systems, wireless computer brain interfaces (WCBI), and the internet of vehicles (IoV) have become a reality, leading to an explosion in wireless data traffic, end-to-end devices, and massive internet of things (IoT). Thus, due to the shift in users' preference for wireless access, present communication infrastructure is experiencing significant capacity constraints. For example, meeting the ever-increasing traffic flow demands from the aforementioned applications will be a difficult task for the existing fourth. generation (4G) and recently deployed fifth-generation (5G) networks. Therefore, the 5G and sixth generation (6G) wireless networks will see a variety of technological revolutions in the upcoming years. Apart from the high data traffic demand, these networks will also have to deliver low latency, ultrareliable and secure communication to maintain quality-of-service (QoS) and quality-of-experience (QoE) for the endusers. Thus, to support high traffic flow demands from these applications, a variety of technological revolutions are expected in the fifth generation (5G) and beyond wireless networks across the globe in the years to come [2], which also increases the demand for smooth data transmission, wireless coverage, and connectivity. In this direction, the heterogeneous network (HetNet) [3] is among the most optimistic technologies to improve network capacity, coverage, and energy-efficiency in 5G networks. The market size of HetNet is expected to be around $51.1 billion by the end of 2027 [4] as shown in Fig. 1.1.

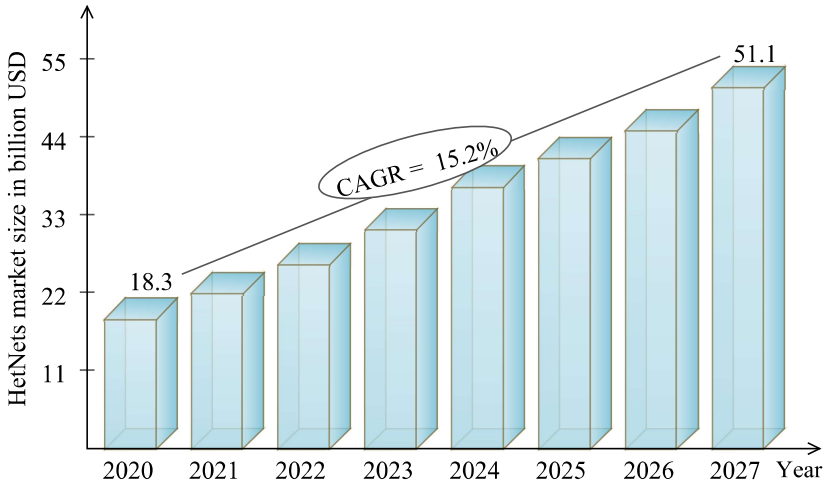

Figure 1.1: HetNets Market Size

1.1 Heterogeneous Networks

HetNets allow dense deployment of the small cells to collaborate in the macrocell, which significantly increases spectrum efficiency (SE), network coverage, and system throughput of wireless networks. HetNet consists of various cellular base stations with varying transmit power levels, such as macrocells, picocells, and femtocells [5]. Macrocells use a high power base station, i.e., macro base station (MBS) to enable all macrocell users (MUs) with wide area coverage up to a few kilometers. MBSs are always installed at a high elevation to allow for a clear view of the surrounding obstacles and buildings. A low-power BS ($\leq$ 20 W) serves a microcell installed in heavily populated areas, such as shopping centers [6]. The micro base station has a coverage area (200m-1km), which is smaller than the MBS. Picocells are a fusion between high capacity-long-range microcells and short-range femtocells. Picocell technology is expected to improve spectrum efficiency, capacity, and coverage footprints [3]. These cells have low transmit powered ($\leq$2 W), operator-managed pico base station (PBS) to work in a licensed spectrum and are generally used for a small areas, aircraft, and buildings. Picocells offload data traffic from macrocells and increase the network capacity by reusing the same spectrum in small dimensions. Femtocells are equipped with low-powered ($\leq$0.2 W) small femto base stations (FBS) to enhance the communication quality in a residential or small business firm. The major difference between FBSs and PBSs is that FBSs serve a less number of users than PBSs. Fig. 1.2 shows the architecture of 5G-enabled HetNet.

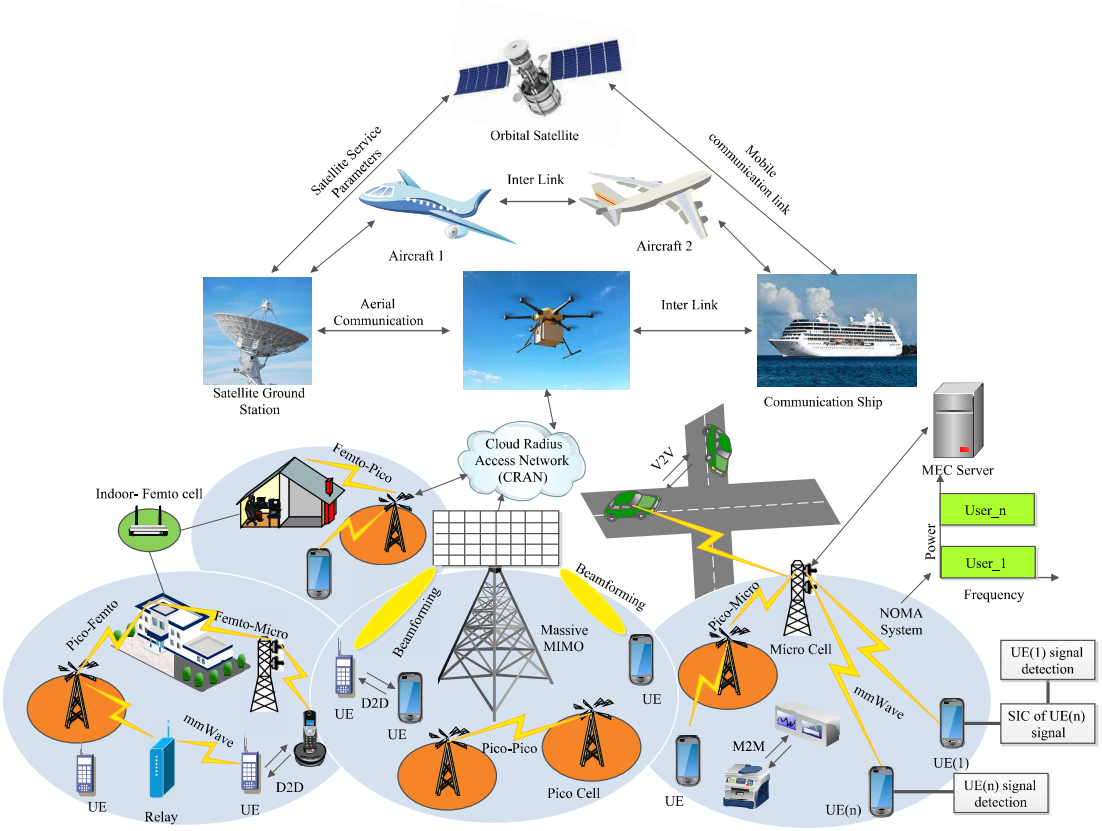

Figure 1.2: 5G HetNet Architecture

1.2 Need of Security in 5G HetNets

The broadcast nature of 5G wireless networks makes them more vulnerable to a broad set of adversarial attacks [7–9] that could jeopardize the security and privacy of the end-users. Traditional security techniques used in previous generations of telecommunication networks focus on encrypting communication data to ensure proper billing system functionality and radio interface security. For example, two-way authentication was used in 3G to prevent the establishment of connections with forged BS. While 4G uses advanced cryptographic techniques for the authentication of users. However, some privacy concerns were addressed to some extent in pre-5G networks as the user data was retained in databases owned by mobile operators. Also, due to the increase in the number of user equipment (UE), services, heterogeneity of connected UEs, high privacy concerns, and new requirements to support various IoT technologies, 5G and beyond networks will face a new set of security concerns. Also, most 5G applications are decentralized, which means that UEs can join or leave the network at any time. Distribution and management of cryptographic keys become extremely difficult in such cases. Also, the use of cryptographic techniques in wireless communications has certain limitations, such

as the need for individual secret keys for every wireless communication link, and eavesdroppers with sufficient computational capacities can solve a complex mathematical problem. Thus, the existing cryptographic techniques can not guarantee secure data transmission in 5G and beyond networks [10]. Also, HetNet is open and diversified in comparison to standard single-tier cellular networks, making the information exchange between different devices more sensitive to eavesdropping [11]. Thus, designing and implementing effective eavesdropping countermeasures is essential for secure wireless transmissions in 5G HetNets.

Bit-level cryptography-based approaches are traditionally used at the top levels to ensure network security. However, those techniques are constrained in their ability to meet the needs of 5G-and-beyond applications due to the following factors: i) In large-scale and decentralized networks, deploying cryptographic approaches relying on public keys is extremely difficult. ii) By employing very long key pairs, public-key encryption has remained impenetrable until now; but, advances in computing powers, such as powerful quantum computers, can crack the key pairs. iii) Further, methods based on cryptography are vulnerable to replay attacks. In replay attacks, an attacker retrieves the physical-layer bit-stream instead of cracking the cryptography-based method and then transmits the recovered signal to the legal receiver. Since the upper-layer signaling is not altered, the replayed signal can easily impersonate the actual receiver. iv) Because of the sophisticated upper-layer processes, such as encryption and decryption, cryptography-based methods have a significant complexity and communication overhead. But since low-cost terminals, such as the internet of vehicles (IoV), unmanned aerial vehicle (UAV), and massive internet of things (IoT), are implicitly power-limited and delay-sensitive, high complexity and communication overhead inevitably raise UE cost.

1.3 Physical Layer Security

Physical layer security (PLS) has gained a lot of attention as it can be an alternative to cryptography-based techniques [12–14]. PLS uses the physical properties of wireless channels to improve the reliability and secrecy of wireless communications. It enables secure data transmission between base stations and the user equipment (UE) without using the secret keys. The notion of PLS was first designed by defining the secrecy capacity of a wiretap channel as explored in [15] and [16]. It is described as the maximum secrecy rate at which private information can be transferred from a transmitter to the receiver in the presence of eavesdroppers. The secrecy rate is referred to as the the difference between the legitimate communication channel's achievable data rate and the eavesdropper's maximum rate (eavesdropping rate).

Some of the commonly used anti-eavesdropping techniques in HetNets include secure beamforming [17], cooperative jamming [18], and physical layer authentication (PLA) [19]. For instance, secure beamforming can be used to reduce the received signal strength at the eavesdroppers. In contrast, physical layer authentication can be used to quickly verify valid UEs, eliminating unnecessary signal processing for undesired transmissions. However, in view of the dynamic nature of 5G HetNets, existing PLS techniques may not guarantee the security in 5G HetNet with fast data transmissions. Moreover, new wireless applications, such as the wireless computer-brain interface (WCBI), massive IoT, and augmented and virtual reality (AR/VR), brings new risks to the physical layer security in 5G HetNets. Hence, there is a strong requirement to design new intelligent PLS techniques for 5G HetNets.

1.4 AI enabled PLS Design

As evidenced by widespread use of AI in different application areas of PLS [20–23], it is undoubtedly one of the most necessary elements for enhancing the PLS of 5G HetNets. It can be used to learn about normal and aberrant behaviors of HetNets based on how users and base stations communicate with one another. AI techniques can successfully anticipate future new instances by learning from existing instances. AI techniques can also be used to forecast new attacks, which are usually mutations of previous attacks. AI has been used in various PLS applications such as security oriented beamforming [20], cooperative jamming [21], PLA [22], secure handover schemes [23], etc. The Ml based classification approach can identify the relay, jammer, and eavesdropper for PLS, while ML-based regression can assist pick the optimal antenna selection from a group. Hoang *et al.* [24] proposed support vector machine (SVM) based eavesdropper detection technique for wireless networks in the absence of channel state information (CSI). Wang *et al.* [25] proposed a decision tree based relay node selection technique by converting it into multi-class classification problem.

DL methods, in general, rely on training and experience to improve task completion performance. This learning approach, which is a subset of ML, analyzes the data for categorization and decision-making without programming. A LSTM-based impersonation attack detection technique was proposed in [26]. The data sequences are preprocessed as a sliding window, and the sliding windows are then trained with an LSTM network. In MIMO-OFDMA systems, Liao *et al.* [27] developed a CNN-based multi-user authentication method to prevent spoofing threats. The authors used CSIs from multiple transmitters as input and matching tags as output to train the CNN. To construct finite alphabet iterative decoders (FAIDs) for LDPC codes, Xiao *et al.* [28] presented a recurrent quantization neural network (RQNN) based

method. In a recurrent structure, the trainable parameters are shared across all iterations, leading to fewer biases and weights and the capacity to extend to more iterations than in typical multi-layer DNNs. Thomas *et al.* [29] created an adversarial network-based system to learn secure coding techniques over a noisy wiretap channel. The authors utilized an adversary approach to solve a minimax game in which a genuine autoencoder network competes against an adversarial network and then validated their approach for secure picture transmission using the CIFAR-10 dataset. Xiaopeng *et al.* [30] proposed a DBN-based intrusion detection method for UAVs that was improved using particle swarm optimisation. Further, to improve the PLS of wireless communications, O'Shea *et al.* [31] developed an end-to-end autoencoder.

DL has recently been integrated into RL techniques, allowing them to tackle a wide range of complicated problems. DRL is a set of approaches for estimating value functions or policy functions using deep neural networks. It uses Markov decision models to help choose between several actions based on state transition models. It has gained a lot of success in a variety of fields, including video games (e.g., Atari) to other realistic applications such as autonomous surgery, robotics, and autonomous vehicles. PLS and wireless networks are two topics that have recently piqued the interest of the DRL research community. For instance, to improve the PLS of MIMO wiretap channels, Youbing *et al.* [32] developed a DQN-based optimum antenna selection method. Also, to enhance the PLS of large MIMO systems, Zhang *et al.* [33] developed an asynchronous advantage actor-critic (A3C) based secure beamforming approach. Since it does not rely on channel estimation, the proposed technique also overcomes the problem of poor CSI in large MIMO systems.

Deep-Q Networks (DQNs) were initially suggested by DeepMind in 2015 as an attempt to apply the benefits of DL techniques in reinforcement learning (RL). DQN uses two unique techniques: experienced replay and iterative updating to solve the DL instability problems [34]. In this, the agent interacts with the environment by performing a series of activities in order to maximize the cumulative reward. Youbing et al. [35] proposed DQN based optimal antenna selection technique to enhance the PLS of the MIMO wiretap channel. DL has significantly outperformed key-based encryption techniques, as well as ML based security techniques. DL is a useful technique for dealing with intricate problems such as detection and mitigation of jamming attacks, power allocation, and physical layer authentication since it summaries hybrid and large physical layer parameters' intrinsic patterns. Furthermore, channel modulation and secure channel coding techniques such as Polar codes and LDPC codes need large computations in current networks. The DL approach has the potential to dramatically increase the performance of these techniques. For instance, as mentioned in the above literature, jamming and eavesdropping attacks

can be mitigated by using DL and DRL based beamforming and power allocation PLS techniques. Spoofing attacks can be detected using efficient PLA schemes, further DL can be used to enhance the authentication accuracy and speed up the process of PLA. DL based channel estimation and PLA can be utilized for efficient CSI and RSSI estimation, which plays an important role in designing intelligent PLS techniques and detecting MITM attacks. Moreover, DRL techniques are efficient in handling dynamic environment, so these techniques can be used in effective design of PLS techniques for highly dynamic 5G and beyond environment.

1.5 Primer on RL and MDP

Reinforcement learning (RL) is a machine learning (ML) based approach which uses one or more agents to constantly interact with the environment to select the optimal policy, which can be viewed as a possible solution to classical optimization problems. The environment responds to the agent's actions by providing rewards and transitioning to a new state. The RL has the following elements:

• *Environment*: It is a representation of the optimization problem that needs to be addressed. It can be a simulated or real-world environment with which the RL agent interacts.

• *State*: State $s_t \in \mathcal{S}$ defines the agents' place in the environment at time-step t.

• *Action*: Action $a_t \in \mathcal{A}$ is the agent's method to interact with the environment and changes its state from $s_t \in \mathcal{S}$ to $s_{t+1} \in \mathcal{S}$. Agents in RL chose their actions by using a policy π.

• *Policy*: It is the mapping between the environments' action and state spaces, i.e., $\pi : s_t \times a_t \longrightarrow [0, 1]$.

• *Reward Function*: The intrinsic desirability of each observed state of the environment is specified by a reward function, which maps each perception of state to a single number. It enables RL-based algorithms to make judgments.

An environment in RL vane be described mathematically using the Markov Decision Process (MDP). It is described as a tuple of $(\mathcal{S}, \mathcal{A}, r, \mathcal{T}_{ss'}, \lambda)$, where $\mathcal{S}$ represents the set of states, $\mathcal{A}$ denotes the set of actions, r is the reward function, $\mathcal{T}_{ss'}$ shows the transition probability from state s to s', and $\varpi \in (0, 1)$ is the discount factor. The aim of MDP is to discover a policy π, that produces the best long-term reward.

1.6 Thesis Organization

The thesis will be organized as follows with the brief description of each chapter.

Chapter 1: Introduction

This chapter introduces the basics of HetNets, PLS, and AI in the 5G. Moreover,

the challenges associated with conventional PLS techniques are also listed.

Chapter 2: Literature Review

In this chapter, we present an in-depth analysis of various PLS techniques for 5G HetNets. These applications include security oriented beamforming, cooperative jamming, resource allocation and power control, game theoretic approaches, channel secrecy codes, secure handover schemes, and physical layer authentication. We provide challenges associated with each PLS technique and how the use of AI can solve these challenges in optimizing the PLS-aware secure data transmission techniques for 5G HetNets.

Chapter 3: SecBoost: Secrecy-Aware Deep Reinforcement Learning based Energy-Efficient Scheme for 5G HetNets

The secrecy level of the mmWave channel model for multiple picocells consisting of legitimate users and eavesdroppers is investigated. Moreover, we formulate a joint power control, channel allocation, and beamforming optimization problem with an aim of maximizing the secrecy energy-efficiency of picocells. A MARL based framework is presented to obtain an optimal policy for joint power control, channel allocation, and beamforming, such that an RL agent (controller) is placed at each pico base station, which try to cooperatively optimize the policy by using Markov decision process (MDP). A multi-agent cooperative DRL based scheme, SecBoost is proposed to boost the secrecy energy efficiency of picocells. The proposed SecBoost scheme exploits the channel allocation, beamforming, and power control domain, along with the dueling structure of dueling double deep Q-network (D3QN). Also, a prioritized experience replay is used with multi-agent D3QN architecture to increase its efficiency. Finally, the SEE performance of SecBoost is compared with MARL, MA-DQN, and JBF-SEEM schemes.

Chapter 4: Mitigating Jamming Attack in 5G HetNets

In this chapter, the achievable rate for multiple mmWave femtocells having FUs and jammers is investigated in this chapter. Moreover, we have formulated a joint optimization problem of beamforming and power allocation at FBSs to maximize the achievable rate at FUs. A multi-agent reinforcement learning (MARL) problem is formulated using the Markov decision process (MDP) to obtain an optimal strategy for joint beamforming and power allocation optimization. A federated DRL scheme is proposed to maximize the achievable rate of mmWaveenabled FUs in a two-tier downlink 5G HetNet. Using extensive simulations, we have compared the achievable rate of proposed scheme with double deep Q network (DDQN) and deep Q network (DQN) schemes.

Chapter 5

Conclusion and Future Scope This chapter concludes the work by highlighting the main contributions made by the proposed techniques. Moreover, this chapter also

provides the future directions in the research areas of AI enabled PLS for 5G HetNets.

Chapter 2

Literature Review

The number of Internet connected wireless devices is growing rapidly and is anticipated to reach around 50 billion by 2030 as per the report by Statistica [36], that is almost 2.3 times from 22 billion in 2018. Modern wireless applications such as e-healthcare, augmented and virtual reality (AR/VR), tactile internet, connected robotics and autonomous systems, wireless computer brain interfaces (WCBI), and the internet of vehicles (IoV) have become a reality, leading to an explosion in wireless data traffic, end-to-end devices, and massive internet of things (IoT).

Thus, due to the shift in users' preference for wireless access, present communication infrastructure is experiencing significant capacity constraints. For example, meeting the ever-increasing traffic flow demands from the aforementioned applications will be a difficult task for the existing fourth generation (4G) and re- cently deployed fifth-generation (5G) networks. Therefore, the 5G and sixth generation (6G) wireless networks will see a variety of technological revolutions in the upcoming years. Apart from the high data traffic demand, these networks will also have to deliver low latency, ultra-reliable and secure communication to maintain quality-of-service (QoS) and quality-of-experience (QoE) for the end-users [37].

Physical layer security (PLS) has been recognized as a possible approach for achieving confidentiality at the physical layer by utilizing the inherent randomness of wireless communications [38]. It can be used to provide secure wireless communications without the use of a key to encrypt them. The advantage of using PLS over cryptographic techniques is that PLS does not depend on computational complexity. Thus, even if the eavesdroppers have high computational capacities, the communication secrecy and reliability can be achieved. Also, physical layer authentication (PLA) can instantly authenticate legitimate UEs before demodu- lating and decoding signals, avoiding wasteful signal processing for unwanted transmissions. PLS methods can also be utilized as an extra layer of security, working in conjunction with existing security measures to provide an adequate safeguard for 5G and beyond networks. Different types of PLS techniques can be used to achieve secrecy. Some

of them are as follows: artificial noise (AN) injection to degrade the signal quality at the eavesdroppers, secure beamforming to eliminate the signal reception at eavesdroppers, and automatic modulation classification for intrusion detection.

2.1 Overview

2.1.1 Scope of the Survey

There exist many review articles in which PLS and other security techniques have been discussed to provide practical guidance and future research opportunities for security concerns in 5G and beyond networks. For instance, Kakkar [39] presented a survey on secure communication techniques for 5G HetNets, mainly focusing on cryptographic algorithms. The authors in [12] discussed the fundamentals and applications of PLS techniques for confidentiality in wireless networks. Giougus *et al.* [40] provide a detailed analysis of secure handoff optimization schemes for wireless HetNets. In [41], the authors presented an exhaustive survey on physical layer authentication schemes in wireless networks, covering their limitations, challenges and future recommendations. Haider *et al.* [42] presented a short survey on the use of AI and ML for 5G network security, mainly focusing on anomaly detection. In [43], the fundamentals, key technologies, and future research directions on PLS in wireless networks are summarized. Yongjun *et al.* [44] provide a detailed overview of resource allocation schemes in 5G HetNets. In [45], the authors presented a comprehensive survey on the physical layer security of 5G networks. However, these previous efforts do not concentrate on the role of AI in optimizing the PLS techniques for secure data transmission in 5G heterogeneous networks. Based on the research questions formulated and listed in Table 2.1, we have compared this survey to other existing most similar surveys in Table 2.2.

2.1.2 Our Contributions

The main contribution of the survey is to present the role of AI in the design and optimization of PLS-aware secure data transmission techniques for 5G HetNets. Our key contributions include:

• First, we present a comprehensive study of different types of physical layer attacks and significant security concerns in 5G HetNets. • Further, we present a comprehensive review on various types of PLS techniques available in the literature for 5G and beyond networks.

• Then, we present an in-depth analysis of various PLS techniques for 5G HetNets. These applications include security oriented beamforming, cooperative jamming, resource allocation and power control, game theoretic approaches, channel secrecy

Table 2.1: Common research questions and their objectives

S.No.	Research Question	Objective
RQ1	What is the role of physical layer security (PLS) in designing secure data transmission techniques for 5G and beyond HetNets?	It aims to explore the need of PLS in designing secure data transmission techniques in cellular HetNets, and the advantage of using PLS over cryptographic techniques.
RQ2	What are the different types of physical layer threats and security concerns in 5G and beyond HetNets?	New 5G technologies will bring new threats and security concerns to HetNets. It serves the purpose to discuss different physical layer threats and security concerns in the design of secure data transmission techniques for 5G and beyond HetNets.
RQ3	What are the existing PLS techniques available for secure data transmission in HetNets?	The purpose is to provide a detailed overview of existing secure data transmission techniques for HetNets.
RQ4	What are the challenges in these PLS techniques concerning 5G and beyond HetNets?	It aims to explore the challenges of using existing PLS techniques for 5G and beyond HetNets.
RQ5	How can AI be used with existing PLS techniques to mitigate the threats and address the security concerns?	The purpose is to investigate how the use of AI can effectively address the security challenges in the design of secure data transmission techniques.
RQ6	What are the applications of AI in designing secure data transmission techniques for 5G and beyond HetNets?	It aims to investigate the use of AI in various applications of PLS-aware secure data transmission techniques.
RQ7	What are the different performance evaluation metrics used to evaluate the performance of AI-enabled secure data transmission techniques?	The objective is to investigate different performance evaluation metrics used in literature to evaluate the performance of AI-enabled secure data transmission techniques.
RQ8	What are these secure data transmission techniques' strengths, weaknesses, opportunities, and threats (SWOT)?	The purpose is to provide a SWOT analysis of the secure data transmission techniques to the readers.
RQ9	What are the challenges and future research trends of using AI-based secure data transmission techniques for 5G and beyond HetNets?	It aims to provide future research trends and challenges of using AI-enabled secure data transmission techniques for 5G and beyond HetNets.

Table 2.2: Comparision of existing survey articles on AI enabled PLS

Survey Papers	RQ1	RQ2	RQ3	RQ4	RQ5	RQ6	RQ7	RQ8	RQ9
Kakkar *et al.* [39]	×	✓	×	×	×	×	×	✓	×
Jehad *et al.* [12]	✓	✓	✓	✓	×	×	×	×	✓
Giorgus *et al.* [40]	×	✓	✓	×	×	×	×	×	×
Lin *et al.* [41]	✓	✓	×	✓	×	×	×	×	×
Haider *et al.* [42]	×	✓	×	×	✓	✓	✓	×	✓
Mukherjee *et al.* [43]	✓	✓	✓	✓	×	×	×	×	✓
Yongjun *et al.* [44]	×	×	×	×	×	×	×	×	×
David *et al.* [45]	✓	✓	×	✓	×	×	×	×	×
This Survey	✓	✓	✓	✓	✓	✓	✓	✓	✓

codes, secure handover schemes, and physical layer authentication. We provide challenges associated with each PLS technique and how the use of AI can solve these challenges in optimizing the PLS-aware secure data transmission techniques for 5G HetNets.

- Also, we have performed SWOT analysis on the PLS aware secure data transmission techniques to assist readers in determining the strengths, weaknesses, opportunities, and threats of these techniques.
- Finally, we discuss future research directions and challenges in the design of AI-enabled secure data transmission techniques for 5G and beyond HetNets.

2.2 Physical layer threats and security concerns in 5G HetNets

This section discusses different types of physical layer threats and significant security concerns in 5G HetNets.

2.2.1 Eavesdropping

The unauthorized and imperceptible observation of a live, private conversation is known as eavesdropping. Eavesdropping occurs when attackers listen in on the data traffic moving across machines, networks, UEs, and IoT devices. Passive eavesdroppers intercept messages while remaining silent, so their CSIs are inaccessible to transmitters. While active eavesdroppers acting as communication parties accidentally send some messages to transmitters, whose CSIs may be acquired using CSI estimation. Due to the open and more diversified structure of 5G HetNets as compared to traditional single-tier cellular networks, 5G HetNets are more vulnerable to eavesdropping attacks [18, 46].

2.2.2 Impersonation

Identity impersonators demolish the identity-centric credibility. These attackers may create a large number of false identities or steal the identities of other genuine nodes [47]. These types of attacks are also known as spoofing attacks. Impersonation attacks can be more sophisticated for 5G and beyond HetNets. For example, if the attacker possesses a full-duplex radio, it can broadcast false identifying data while concurrently monitoring the victims. Sybil attack and identity spoofing are the most common types of impersonation attacks.

2.2.3 Jamming

Jammers are malicious wireless devices installed by an adversary to create malicious interference in wireless networks [48]. It significantly reduces the secrecy capacities of wireless channels. Many studies [49,50] have shown that 5G HetNets are susceptible to jamming attacks. Many jamming techniques for SVD-based MIMO systems, including a practical and robust channel traffic threat, are demonstrated in [51]. Some of the most commonly studied jamming attacks are reactive jamming and pilot jamming.

2.2.4 Contaminating

In contaminating attacks, attackers seek to contaminate the phase of channel estimation to gain unfair advantages in the subsequent communication phase [52]. This type of attack can be classified into feedback, pilot, and contamination based on various channel estimate phases. The beamforming design at the base stations in massive MIMO and NOMA communications is dependent on the CSI, which is determined using pilot sequences. If attackers contaminate the pilot signals, they will have a better chance of eavesdropping on legal communication or severely degrading the communication performance of UEs.

2.2.5 Man in the Middle (MITM) Attack

A man in the middle (MITM) attack occurs when a perpetrator inserts oneself into a dialogue between a user and an application, either to listen in or mimic one of the parties, making it look as though a regular information exchange is taking place. Some new MITM attacks possible on 5G and beyond networks are device bidding down and battery drain [53].

2.2.6 Semantic Information Attacks

5G networks have a much smaller coverage area than 4G networks, and unlike 4G, their signals cannot penetrate barriers. As a result, indoor and outdoor 5G networks will require many smaller antennas and base stations. Identifying which cell tower or antenna a UE communicates with can reveal crucial UE location data. In semantic attacks, attackers can use false information to weaken the trustworthiness of target resources by exploiting the location data of UEs.

Summary and Insights: In HetNets, there are countermeasures available that provide security against many sorts of physical layer attacks. However, as wireless networks become more advanced, new risks and security problems emerge in designing secure transmission techniques for 5G HetNets. Thus, the existing countermeasures should be strengthened to address new security concerns and potential threats on HetNets.

2.3 PLS in 5G and beyond networks

This section discusses various conventinal PLS techniques used in the key technologies of 5G and beyond as follows: NOMA, Full Duplex network, Massive MIMO, Cognitive Radio, Relay Network, SWIPT, HetNet.

2.3.1 Non-Orthogonal Multiple Access (NOMA)

Non-orthogonal multiple access (NOMA) systems have received a lot of interest in recent years for 5G cellular networks [54]. The capacity of NOMA to serve a large number of users by sharing the same time and frequency resources is the central motivation behind its adoption in 5G. It delivers high device performance, high efficiency, increased coverage, low latency, and huge networking [55]. In contrast to point-to-point transmission techniques, there are two types of eavesdroppers in NOMA: external eavesdropper (passive) whose CSI cannot be identified by the transmitter, and internal user (active) whose CSI can be recognized at the transmitter. As a physical layer technology, the security of NOMA networks is a significant issue that needs more investigation. However, certain issues must be addressed during the design phase of PLS for NOMA, such as the disparity in transmit power and the diverse security demands of users.

NOMA is recently used within the 3GPP long-term evolution advanced (LTE-A) model because of its spectral performance gain, demonstrating the role of NOMA in 5G wireless networks. As a result, one of the top concerns in the design and operation of 5G wireless networks is to provide a peerless level of security for NOMA [56]. To effectively integrate physical layer protection with NOMA, a considerable amount

of work is needed. Zhang *et al.* [57] analyzed the stable NOMA transmission for a single transmitter, multiple receivers having perfect CSI, and one eavesdropper device with no CSI knowledge. Furthermore, in [58], a joint beamforming and power allocation method for NOMA based satellite terrestrial integrated network is presented. A summary of the latest beamforming-based PLS techniques for NOMA is given in Table 2.3.

2.3.2 Full Duplex Network

Full-duplex PLS is an exciting area that recently received a lot of attention. This type of transmission includes sending and receiving data on the same frequency range at the same time. It will theoretically double spectral efficiency as compared to half-duplex ones. Thus, physical layer security for full duplex systems is a potential research subject that has recently achieved a lot of attention. The following three categories can be used to categorize the research on full-duplex physical layer security: i) Full-duplex base station ii) Full Duplex transmitter and receiver iii) Full duplex active eavesdropper.

Li *et al.* [59] studied a single transmitter, a full-duplex receiver, and a single-eavesdropper wireless communication model in which the full-duplex legitimate receiver uses one antenna to obtain the signal and another antenna to transmit AN to the eavesdropper. Then, a joint transmitting and receiving beamforming architecture for a single-antenna receiver, multiple-antenna transmitter, and a single eavesdropper with incomplete CSI is studied by Zheng *et al.* [60]. Meanwhile, Akgun *et al.* [61] suggested a similar MISO multiple-antenna eavesdropper wiretap channel, in which the transmitter uses the zero-force beamforming to prevent interference caused by multiple users. A summary of the latest beamforming-based PLS techniques for the full-duplex network is provided in Table 2.4.

Table 2.3: PLS techniques for NOMA

Authors	BF	CJ	AN	CC	Security Parameter	Provides security from attack	Description	Advantages	Limitations
Cao *et al.* [62]	✓				Secrecy Rate, Transmission Rate	Eavesdropping, Signal Interference	Beamforming is used to preserve the privacy of data transmission in MISO NOMA systems	Improved secrecy rate of Private users	Proposed scheme should be tested for multiple eavesdropper scenerio
Hao *et al.* [63]	✓				Secrecy Sum Rate	Eavesdropping	Beamfrming is used along with SWIPT to improve the secrecy sum rate of the NOMA Network	Improved secrecy sum rate	High computational complexity
Zhao *et al.* [64]	✓	✓			Secrecy Rate	Signal Interferecnce, Eavesdropping, Jamming	Beamforming and artificial noise are used to secure the transmission channel in NOMA Network	Improved secrecy rate, reduced eavesdropping rate	Proposed Technique should be tested on more security parameters
Deng *et al.* [65]	✓				Average Secrecy Sum Rate	Eavesdropping, Signal Interferecnce	Hybrid Artificial jamming and beamforming is used to improve the security of NOMA network in the presence of an eavesdropper	Works well for both perfect and imperfect CSI	Tested only for single eavesdropper case

Table 2.3: PLS techniques for NOMA

Authors	BF	CJ	AN	CC	Security Parameter	Provides security from attack	Description	Advantages	Limitations
Li *et al.* [66]	✓				Sum Achievable Secrecy Rate	Eavesdropping	Beamforming is used to secure the downnlink MISO NOMA network	Improved Sum Achievable Secrecy Rate	Proposed Technique should be tested on more security parameters
Jiang *et al.* [67]	✓				Secrecy Rate	Eavesdropping, Signal Interference	Beamforming is used to secure the downlink MIMO NOMA network	Better secrecy rate as compared to zero-force beamforming technique	High computational complexity
Feng *et al.* [68]	✓		✓		Secrecy Sum Rate	Eavesdropping	Beamforming along with jamming is used to secure the physical layer of NOMA network	Useful in power allocation problem	Proposed Technique should be tested on more security parameters
Yin *et al.* [69]	✓				SINR, Secrecy Rate	Eavesdropping	Beamforming is used to improve the physical layer security of the UAV enabled NOMA network	Improved Secrecy unicasting rate	Proposed scheme should be tested for multiple eavesdropper scenerio

Table 2.4: PLS techniques for full duplex networks

Year	BF	CJ	AN	CC	Security Parameter	Provides security from attack	Description	Advantages	Disadvantages
Fengchao *et al.* [70]	✓	✓			Average Secrecy Rate	Eavesdropping	Joint beamfroming and cooperative jamming is used to improve PLS at both the transmitter and receiver of the full duplex base station	Improved average secrecy rate	High computational complexity
Chalise *et al.* [71]	✓				Secrecy Rate	Eavesdropping	Beamforming based secure full duplex network having perfect and partial values of CSI	Improved secrecy rate	Proposed Technique should be tested on more security parameters
Juhwan *et al.* [72]	✓	✓			Average Secrecy Rate	Eavesdropping, Signal Interference	Joint cooperative jamming and beamforming is used to secure data transmission in Full duplex systems	Improved average secrecy rate	Tested only for single eavesdropper case
Wei *et al.* [73]	✓		✓		Secrecy Rate	Eavesdropping, Signal Interference	Joint artificial noise and beamforming is used to secure data transmission in Full duplex systems having self energy recycling phenomenon	Low power consumption	Proposed Technique should be tested on more security parameters

Table 2.4: PLS techniques for full duplex networks

Year	BF	CJ	AN	CC	Security Parameter	Provides security from attack	Description	Advantages	Disadvantages
Zhengmin *et al.* [74]	✓	✓		✓	Secrecy Rate	Eavesdropping	Secure beamfroming based on artificial noise and cooperative jamming is used to improve PLS	Improved worst case secrecy rate	High computational complexity
Qiang *et al.* [75]	✓			✓	Sum Secrecy Rate	Eavesdropping	Joint artificial noise and two rank beamforming for maximizing the sum secrecy rate of the full duplex network	Improved sum secrecy rate	High computational complexity
Cepheli *et al.* [76]	✓			✓	Data rate	Eavesdropping	Joint artificial noise and beamforming for maximizing the secrecy secrecy rate while maintaining the data rate and throughput of the full duplex network	Improved Throughput	Proposed Technique should be tested on more security parameters

2.3.3 Massive MIMO

Massive MIMO is one of the "big three" 5G innovations [77], and it is an exciting solution for the effective transfer of massive data. One of the advantages of massive MIMO is that it significantly increases the capability of physical layer security against passive eavesdropping assaults. However, the eavesdropper can take countermeasures. For instance, it might place itself near to the legitimate user, so that the routes to the legitimate user and the eavesdropper gets highly connected. Therefore, physical layer security for massive MIMO system has to be advanced. A recent study presents PLS for large MIMO networks with passive eavesdroppers. In this, Zhu *et al.* [78] investigate stable massive MIMO communications for multicell-user networks over Rayleigh fading wireless channel, in which an eavesdropper tries to decipher the data transmitted to one of the users.

Regularized inversion of the wireless channel along with AN injection is presented in [79] for the massive MIMO to increase the secrecy rate further. Wang *et al.* [80] analyze AN-assisted stable massive MIMO transmission over a fading channel with Rician fading. Secure communication for massive MIMO systems with restricted radio frequency and hardware abnormalities [81, 82] secure tactics in the existence of a massive MIMO eavesdropper [83,84], confidentiality outage likelihood study for massive MIMO systems [85] are the recent developments in the design of PLS of massive MIMO systems. A summary of the latest beamforming-based PLS techniques for massive MIMO is presented in Table 2.5.

Table 2.5: PLS Techniques for Massive MIMO

Year	BF	CJ	AN	CC	Security Parameter	Provides security from attack	Description	Advantages	Disadvantages
Qingjiang *et al.* [86]	✓		✓		Secrecy Rate	Eavesdropping	Inexact block coordinate descent (IBCD) algorithm based on beamforming and artificial noise is proposed to maximize the secrecy rate	Maximized secrecy rate	Does not solve the interference issue between the users.
Mukherjee *et al.* [87]	✓				Bit Error Rate, SNR	Eavesdropping	Zero force beamfrming and optimal power beafmorming techniques are analyzed for their performance in multicase MIMO network	Reduced bit error rate, low power consumption	Tested only for single eavesdropper case
Jayasinghe *et al.* [88]	✓				Bit Error Rate, SNR	Eavesdropping	Secure beamforming technque is proposed to improve the physical layer security of MIMO system	Reduced computatinal complexity	Eavesdropper can intercept the transmitted data

Table 2.5: PLS Techniques for Massive MIMO

Year	BF	CJ	AN	CC	Security Parameter	Provides security from attack	Description	Advantages	Disadvantages
Kaushik *et al.* [89]	✓				Data Rate	Eavesdropping	Beamfroming is used to improve the datta rate of the MIMO systems	Energy efficient	Proposed Technique should be tested on more security parameters
Tian *et al.* [90]	✓				Secrecy Rate, Spectral Efficiency	Eavesdropping, Interference bewteen users	Secure beamforming technque is proposed to improve the secrecy rate of MIMO system	Improved Secrecy Rate	Tested only for single eavesdropper case
Zhao *et al.* [91]	✓				Secrecy Sum Rate,	Eavesdropping, Interference bewteen users	Secure beamforming is used to maximize the secrecy sum rate of the full-duplex MIMO systems	Improved Secrecy Sum Rate	Proposed Technique should be tested on more security parameters
Nabil *et al.* [92]	✓		✓		Secrecy Rate, SNR	Eavesdropping	Hybrid artificial noise and beamforming is used to improve the security of the physical layer of MIMO-OFDM systems	Low power consumption	High computational complexity

2.3.4 Cognitive Radio Network

Cognitive Radio (CR) network is a technology to address the issue of under-utilized wireless bandwidth by enabling secondary users to access approved networks without interfering with primary users' transmissions [93]. It adjusts its transmission parameters in response to its interaction with other devices [94]. Although cognitive radio is an effective strategy for alleviating the pressures of wireless spectrum depletion, its characteristics have also introduced entirely new forms of PLS risks and challenges [95]. In [96], authors studied robust secure beamforming for wirelessly powered cognitive staellite-terrestrial networks. Further, Lin *et al.* [97] presented a secure beamforming technique for Rate-splitting multiple access (RSMA) based cognitive networks having multiple eavesdroppers. Furthermore, in [98], a cooperative NOMA (co-NOMA) framework is presented to evaluate the security-reliability tradeoff for primary user networks. Simulation findings reveal the co-NOMA's effectiveness compared to non-cooperative NOMA (nco-NOMA) in terms of total connection outage probability and secrecy outage probability.

As a result, one of the most critical criteria for deploying CR networks is its security architecture. A summary of the latest beamforming-based PLS techniques for Cognitive Radio is provided in Table 2.6.

Table 2.6: PLS Techniques for Cognitive Radio Newtorks

Year	BF	CJ	AN	CC	Security rameter	Pa-	Provides security from attack	Description	Advantages	Disadvantages
Fengchao *et al.* [99]	✓				SINR		Eavesdropping	Cooperative beamforming is used to improve the security of both the primary and secondry users	Low power consumpion by secondry users	Requires prior infromation of eavesdropper channel
Jiang *et al.* [100]	✓				Secrecy Rate		Eavesdropping	Beamfrming is used to maximize the data rate of the secondry system which leads to improved secrecy rate of the primary system	Better secrecy rate compared to the energy harvesting system	Secrecy Rate decreases with the increase in the number of eavesdroppers
Tang *et al.* [101]	✓				Secrecy Rate		Eavesdropping	A cooperative spectrum sharing techhnique is designed to improve the secrecy rate of the primary users	Better secrecy rate of the primary users compared to the zero - forcing beamforming	Throughput of the secondry users decrease with the increase in the number of eavesdroppers

Table 2.6: PLS Techniques for Cognitive Radio Newtorks

Year	BF	CJ	AN	CC	Security Parameter	Provides security from attack	Description	Advantages	Disadvantages
Antony *et al.* [102]	✓				Bit Error Rate, Secrecy Outage Probability and Secrecy Rate	Interferecnce bewteen users	Secure beamforming is used to secure the communication between the primary and the secondry cells	Provdes better secrecy rate even in the presence of large number of anteenas	Limited to Cognitive Radio Networks only
Wu *et al.* [103]	✓	✓			Secrecy Rate	Eavesdropping, Interferecnce bewteen users	Artificial Noise with beamforming is proposed to enhance the quality of the signal and enhancing the security of sharing of the spectrum	Improved Secrecy rate	Low secrecy rate if the number of anteenas are more then six
Pei *et al.* [104]	✓				Secrecy Rate, Secrecy Capacity	Eavesdropping	Secure beamforming is used to secure the communication channel of MISO based cognitive radio	Improved Secrecy capacity	Proposed scheme needs to be campared with different existing schemes to validate the performance

2.3.5 Relay Network

Relay Networks are classified into two types: full-duplex and half-duplex [105, 106]. The term full-duplex relay, also known as a two-way relay, refers to the ability of a relay to send and receive wireless signals on the same channel at the same time. In comparison, for the half-duplex relay (oneway relay) to send and receive wireless signals, two different channels are required. Recently, relaying techniques have also gained substantial interest in the field of PLS over 5G wireless networks, benefiting from information-theoretical techniques used in cooperative communications [107]. For example, analogous to cooperative communications, relay nodes maybe used as trustworthy nodes to resend an improved version of the wireless signal obtained from the transmitter with an appropriate power amplification coefficient, namely amplify-and-forward (AF) relay network.

The advantage of using beamforming in multiple antenna devices for two-way AF relaying was explored in [108]. In [109], authors investigated confidential message transmission over multihop connectivity using a chain of linked untrusted relays. The SOP in a three-node AF relay network has an untrusted relay node was investigated in [110], where authors demonstrates that confidentiality can be maintained if the transmitter and receiver are kept apart. A summary of the latest beamforming based PLS techniques for relay network is illustrated in Table 2.7.

2.3.6 Simaltaneous Wireless Information and Power Transfer (SWIPT)

SWIPT technology is used for simultaneous energy and information exchange within a variety of modern wireless networks in the context of 5G connectivity. By allowing simultaneous transmission of both information and power, SWIPT provides significant gains in terms of energy consumption, spectral performance, interruption control, and transmission delay [111].

However, it has been demonstrated that in certain situations, SWIPT receivers might have a more powerful medium to obtain information directed at receivers, thus putting the security of the data transmitted in danger. PLS has been used on SWIPT in a variety of studies [112–115] in order to resolve these issues. Authors of [66] proposed a safe beamforming architecture for SWIPT in a heterogeneous network consisting of a single macrocell with multiple users and a single femtocell with multiple eavesdroppers and a single receiver scenario. A summary of the latest beamforming based PLS techniques for SWIPT is provided in Table 2.8

Table 2.7: PLS Techniques for Relay Newtorks

Year	BF	CJ	AN	CC	Security Parameter	Provides security from attack	Description	Advantages	Disadvantages
Zhang *et al.* [116]	✓			✓	Secrecy Sum Rate, Secrecy Rate	Eavesdropping	Secure beamforming is used along with network coding to improve the security of two-way relay network	Improved secrecy rate	Proposed scheme is tested only for single anteena
Junwei *et al.* [117]					Secrecy Rate, SNR	Eavesdropping	Callobrative beamfrming is used w.r.t. secrecy conditions to improve the secrecy rate of the relay network	Reduced comutational complexity	Works only for perfect CSI
Yang *et al.* [118]	✓				Secrecy Rate	Eavesdropping	Two cooperative secure beamforming technques are proposed to improve the secrecy rate with power constraints	Improved data rate for multiple eavesdroppers	Proposed Technique should be tested on more security parameters

Table 2.7: PLS Techniques for Relay Newtorks

Year	BF	CJ	AN	CC	Security Parameter	Provides security from attack	Description	Advantages	Disadvantages
Wang *et al.* [119]	✓		✓	✓	Secrecy Sum Rate	Eavesdropping	Improved physical layer security of bidirectional transmission for multiiple relays in the presence of an eavesdropper	Works well for imperfect CSI scenerio	Proposed Technique should be tested on more security parameters
Li *et al.* [120]	✓		✓		Average Secrecy Sum Rate	Eavesdropping, Interference between users	Artificial Noise with beamforming is proposed to enhance the quality of the signal and enhance the security of two-way relay network	Improved Average Secrecy Sum Rate	Proposed Technique should be tested on more security parameters
Chengmin *et al.* [121]	✓	✓			Average Secrecy Sum Rate, Average secrecy capacity	Eavesdropping	Cooperative beamforming along with jamming is used in DF relay networks to secure physical layer	Low computational complexity	Tested only for single eavesdropper case

Table 2.8: PLS Techniques for SWIPT

Year	BF	CJ	AN	CC	Security Parameter	Provides security from attack	Description	Advantages	Disadvantages
Bin *et al.* [122]	✓				SINR	Eavesdropping	Under SINR constraint, a beamforming based technique is designed to secure the data transmission in SWIPT	Reduced Power consumption	Proposed Technique should be tested on more security parameters
Schober *et al.* [123]	✓		✓		SINR, CSI, Average Secrecy Capacity	Eavesdropping	Beamforming and Artificial Noise based secure communication for SWIPT in presence of passive eavesdroppers	Improved average secrecy capacity	Tested only for passive eavesdroppers
Dong *et al.* [124]	✓		✓		Secrecy Rate	Eavesdropping, Signal Interference	Joint artificial noise and beamforing is used to secure data transmission in Full duplex SWIPT systems	Improved secrecy rate	High computational complexity
Zheng *et al.* [125]	✓	✓	✓		Secrecy Rate	Eavesdropping, Interference bewteen users	Secure beamforming for MISO SWIPT using artificial noise and cooperative jamming	Improved Secrecy Rate	High computational complexity

Table 2.8: PLS Techniques for SWIPT

Year	BF	CJ	AN	CC	Security Parameter	Provides security from attack	Description	Advantages	Disadvantages
Jiang *et al.* [67]	✓				Secrecy Rate	Eavesdropping	Secure beamforming is used to improve the secrecy rate of the SWIPT in D2D networks	Improved Secrecy Rate	Proposed Technique should be tested on more security parameters
Quanzhong *et al.* [126]	✓				Average Secrecy Rate	Eavesdropping	Relay based secure beamforming is used to secure the data transmission in SWIPT Relay networks	Low power consumption	High computational complexity
Zhengyu *et al.* [127]	✓	✓			Transmit Power, Secrecy Rate	Eavesdropping	Joint artificial noise and beamforming is used to improve the security of the physical layer at SWIPT systems	Improved Secrecy Rate	Tested only for single eavesdropper case

2.3.7 Heterogeneous Network (HetNet)

A heterogeneous network, in general, is made up of several layers of networks that use the same bandwidth. Recent research on PLS of heterogeneous network focuses on how to build transmission schemes that protect multi-tier communications. The use of heterogeneous networks in 5G can intelligently and efficiently merge various nodes into a multi-tier proposed network, which includes macro-cells with high-power nodes for broad radio access networks, small cells with limited power nodes for small radio access networks, and so on [128]. Compared to the traditional single-tier topology, this multi-tier architecture presents new problems in the analysis of PLS. The positions of high/low power nodes would directly affect the PLS architecture and must be carefully modelled and evaluated. Further, heterogeneous networks can cause substantial cross-tier interference. This should be considered when developing reliable and secure data transfer techniques. Furthermore, in HetNets, users can access an arbitrary tier, such as open access. As a result, specific user association policies that coordinate both QoS and secrecy are required.

Lv *et al.* [129] investigated the PLS in a two-tier heterogeneous downlink network having multiple users and an eavesdropper in each cell. By comparing the average received RSSI at the user with a given threshold, Wu *et al.* [130] proposed a user association strategy. Then, for a downlink K-tier HetNet with multiple nodes, a closed-form expression of secrecy outage probability is derived by modeling the nodes' positions as separate Poisson Point Processes, assuming that both intra-cell and inter-cell interference would not occur. Xu *et al.* [131] suggested a complex organized multipoint transmission strategy to expand the stable connectivity reach in a downlink K-tier HetNet for intra-cell interference. A summary of the latest beamforming based PLS techniques for HetNet is given in Table 2.9.

Table 2.9: PLS Techniques for HetNets

Year	BF	CJ	AN	CC	Security Parameter	Provides security from attack	Description	Advantages	Disadvantages
Lv *et al.* [129]	✓				Secrecy Rate, SINR	Eavesdropping	Secure beamfroming is used to to improve the PLS of two-tier downlink hetnet	Low power consumption	High computational complexity
Jahandideh [132]	✓		✓		Secrecy Rate	Eavesdropping, Signal Interference	Joint beamforming and artificial noise is used to secure the data transmission in two relay hetnet	Improved secrecy rate	Proposed Technique should be tested on more security parameters
Bin *et al.* [133]	✓		✓		Secrecy Rate	Eavesdropping	Joint beamforming and artificial noise is used to secure the data transmission SWIPT two tier hetnet	Low power consumption	Tested only for single eavesdropper case
Chang *et al.* [134]	✓	✓	✓		Secrecy capacity, SNR	Eavesdropping, Signal Interference	Beamforming is used with artificial noise and cooperative jamming to improve the PLS of heterogeneous massive MIMO system	Improved secrecy capacity	High computational complexity

2.3.8 UAV Communication Networks and Space Information Networks

Unmanned Aerial Vehicles (UAVs) have gained much popularity in both commercial and military applications for their ability to leverage evolving wireless networks and deliver reliable communication due to their flexible deployment, high mobility, and inherent line-of-sight (LOS) air-to-ground (A2G) channels. Despite the remarkable advancements made by UAVs, the open nature of air-to-ground wireless channels makes secure information transfer difficult [135]. Therefore, the security of UAV based communication network is critical. Li *et al.* [136] conducted a comprehensive survey on physical layer security in space information networks. In the survey, the authors have briefly introduced the satellite Internet of Things (IoT), related research issues, security performance metrics, and the state-of-art PLS approaches related to space information networks. In [137], the authors have introduced PLS to UAV communication networks to address the problem of data leakage induced by potential eavesdropping. The authors have also discussed PLS schemes for two application scenarios of UAVs, i.e., UAV as a base station and UAV as an aerial node. Further, in [138], secure transmission for cognitive satellite-terrestrial networks is investigated to enhance the security of satellite links. In particular, the authors examined the challenge of downlink beamforming for secure cognitive satellite-terrestrial networks by evaluating the realistic scenario of incomplete CSI for the primary user's, secondary user's, and Eve's links.

2.3.9 Other 5G and beyond Scenarios

Physical layer security techniques have a wide range of uses in 5G and beyond. Traditional security key systems rely heavily on the sharing of secret keys. The key generation in the physical layer was first studied in [139, 140], where coupled measurements of noisy patterns can be used to produce secret keys over a dedicated network. Key distribution is one of the few physical layer security measures that can be used in today's wireless applications. Many prototypes involving physical layer secret-key distribution have been published in [141–144]. Further, RIS is considered as one of the key technologies for B5G wireless networks. It comprises of a large number of low-cost passive reflectors, which can reflect the signal independently by adjusting its phase or amplitude to achieve signal enhancement or passive beamforming. RIS empowered wireless networks have the capacity to regulate the electromagnetic wave propagation environment [145]. Lin *et al.* [146] proposed joint beamforming design and optimization for RIS based hybrid satellite-terrestrial relay networks. Simulation results show the effectiveness of the scheme in enhancing the QoS for satellite communications.

Internet of Things (IoT) is a network of devices equipped with radio-frequency identification (RFID), actuators, sensors, and networking to enable them to communicate with operators, suppliers, and other connected devices to achieve shared goals. 5G would be a crucial component for IoT by allowing many machine-type communication (MTC) devices to link to the network. In the literature on PLS, these dimensions have received comparatively little consideration. For example, there is still a need for a technically sound and holistic approach to define uncertainty and energy constraints in PLS designs specifically.

Summary and Insights: We have found that the study of beamforming based PLS techniques have provided a lot of literature with themes spanning from security-related theoretical studies to practical criteria designs. Integration of beamforming based PLS in different promising technologies of 5G is a potential research path in security paradigms. For example, NOMA with Beamforming (NOMA-BF) can use both the power and spatial domains to improve spectral efficiency and secrecy by increasing SINR at UEs. In the case of full-duplex wireless communications, beamforming-based approach can be effectively used to mitigate the signal interference. Furthermore, using beamforming cancellation, several standard RF, antenna, and baseband cancellation processes can be eliminated, which reduces the system complexity. In massive MIMO systems, beamforming can be used to improve the system security, spectral efficiency, and energy efficiency. In massive MIMO systems, the likelihood of an eavesdropper receiving the beamformed signal is lower as compared to systems with conventional antennas. Further, beamforming can be effectively used with artificial noise injection to enhance the secrecy of primary and secondary networks in cognitive radio networks. Also, beamforming can be used for confidential data transmission in relay networks by using relay nodes as trustworthy nodes and optimizing the beamforming vectors for cognitive receivers' signal. Beamforming is further used in the literature for designing secure transmission scheme while satisfying the energy harvesting constraints for SWIPT enabled wireless networks. In terms of HetNets, beamforming can be used at macro base station, micro base station, femto base station and pico base station to enhance the security of multi-tier wireless networks. The physical layer characteristics of the wireless channel, in particular, may be used to develop novel security and privacy-preserving techniques in 5G and beyond. In most of the literature, the eavesdropper is frequently believed to have the same or worse channel characteristics as the legitimate receiver. However, this may not always be the case in practice, like Eve can have more antennas than base stations and user devices, leading to security issues.

2.4 Key applications of AI in designing PLS aware secure transmission techniques for 5G HetNets

AI plays an essential role in the design of secure data transmission schemes. One of the benefits of using AI-based PLS over traditional approaches is that they can anticipate and detect future attacks by learning from prior incidents. In this section, we discuss some of the key applications of AI in designing PLS-aware data transmission techniques for 5G HetNets. We also discuss challenges associated with traditional PLS techniques and how the use of AI can enhance the performance of PLS for 5G HetNets. Finally, we also provide a SWOT analysis of the secure data transmission techniques.

2.4.1 Security Oriented Beamforming

Beamforming is a signal processing technique used to guide data transmission. It ensures the best possible signal quality difference between the genuine receiver and the eavesdropper. It is considered as one of the promising techniques in the design of PLS [17], [99], [147]. Consider a simple PLS system model having one transmitter (T), one receiver (R), and one eavesdropper (E), all of which have multiple antennas in the quantity of N_T , N_R and N_E respectively. The legitimate receiver (R) receives the signal vector $r_{R(j)}$ as

$$r_{R(j)} = H_{TR}(j)w(j)s(j) + n_R(j) \tag{2.1}$$

where $H_{TR}(j)$ is the wireless channel parameters matrix between transmitter and receiver of size $N_T \times N_R$, $w(j)$ is the weight coefficient matrix of the channel, $s(j)$ is the transmitted signal and $n_R(j)$ is the AWGN vector of size N_R. Similarly, the eavesdropper (E) receives the signal vector $r_E(j)$ as

$$r_E(j) = H_{TE}(j)w(j)s(j) + n_E(j) \tag{2.2}$$

where $H_TE(j)$ is the wireless channel parameters matrix between transmitter and eavesdropper of size $NT \times NE$, $w(j)$ is the weight coefficient matrix of the channel, $s(j)$ is the transmitted signal and $n_E(j)$ is the AWGN vector of size N_E. The SINR at the receiver and Eavesdropper can be given as

$$\psi_R = \sum_{k=1}^{N_R} P_S\Big(\frac{w^H R_{TR,k} w}{{\sigma_R}^2}\Big) \tag{2.3}$$

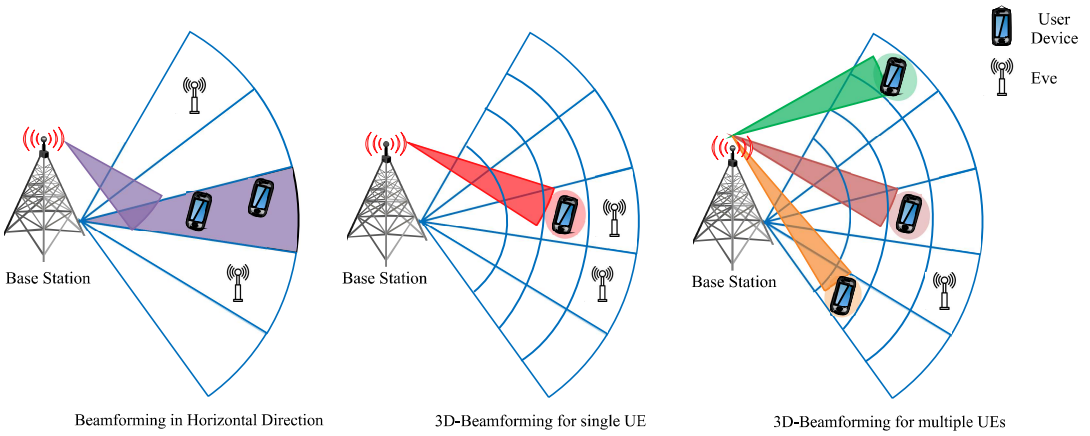

Figure 2.1: Differet types of Beamforming

$$\psi_E = \sum_{k=1}^{N_E} P_S\left(\frac{w^H R_{TE,k} w}{{\sigma_E}^2}\right) \tag{2.4}$$

where $R_{TR,k}$ and $R_{TE,k}$ are the Channel State Information (CSI) of the Receiver and Eavesdropper which are available to Transmitter, P_S is the power of transmitted signal. The main objective of applying beamforming is to maximize the signal power and minimize the other signal interference in the main wireless channel such that the legitimate receiver can receive good quality signal. It can be formulated as SINR at receiver must be greater then a threshold value R_{th}, i.e., $\psi_R \geq R_{th}$. The optimal beamforming design to maximize the secrecy capacity (C_s) can be mathematically formulated as

$$\begin{aligned} \max_{w} \quad & C_s(w) \\ \text{s.t.} \quad & \| w \|^2 \leq P^{max} \end{aligned} \tag{2.5}$$

where w is the beamforming matrix with the maximum transmit power constraint P^{max}. Different types of beamforming are illustrated in Fig. 2.1. 3D beamforming can be used if the beam has to be transmitted in multiple directions.

It has been studied as a potential PLS technique for 5G HetNets [18, 148, 149]. Yan *et al.* [18] proposed a robust beamforming algorithm to enhance the secrecy rate of macrocell users in a two-tier 5G HetNet. The simulation results show the effectiveness of the proposed scheme against collusive eavesdoppers in a massive MIMO and NOMA enabled HetNet. Zheng *et al.* [148] studied distributed beamforming for secure content delivery in a two-tier 5G HetNet. It was observed that the distributed beamforming and direct transmission scheme maximizes the overall secrecy throughput of a cache-enabled mmWave HetNet. Malcolm *et al.* [149] implemented fast converging beamforming technique for macrocell users in massive MIMO HetNet scenario. In this, the optimization problem of SINR balancing is investigated for imperfect CSI. In [150], a joint beamforming and relay selection

optimization problem was investigated to mitigate negative impacts of the jamming signals in relay-based systems, and was solved using the semi-determined relaxation scheme. A beam extraction method based on spatial data was proposed in [151] to eliminate the eavesdropping and suppress the jamming signals in massive MIMO systems. iejun *et al.* [129] presented a PLS technique based on secure beamforming for a two-tier HetNet. The simulation results revealed the technique's efficacy in improving HetNet's secrecy rate performance.

Challenges in security oriented beamforming: However, because of the simultaneous optimization of multiple variables and the constant modulus constraint, the beamforming design is widely acknowledged to be a challenging non-convex problem, and it is rare to find a closed-form optimal solution [152]. Also, the above-mentioned beamforming techniques are often practically challenging to deal with highly dynamic 5G and B5G HetNet environments. Intelligent beamforming techniques are the necessity to process excellent CSI by the base station, optimization of beamforming vectors, and reduce the computational complexity. *AI assisted secure beamforming*: AI has been acknowledged as a successful approach for solving intractable difficulties, and thus it has been used in numerous research studies to solve the issue of BF optimization. Zhihan *et al.* [20] proposed a DNN based technique to improve the PLS of 5G HetNets. In this, joint beamforming and modulation information recognition is studied with unsupervised deep learning. The scheme reduces the computational complexity of beamforming design under different number of transmit antennas and also solve the modulation recognition problem in 5G HetNets. However, the quantity and quality of the dataset can affect the security performance of the DL based schemes in 5G HetNets. In some studies, [153,154], CNN has been used to improve beamforming vectors in cellular networks. By accepting CSI as an input feature, CNNs are trained to generate effective beamforming for a specified power constraint at the BSs. Specifically, in [154], the authors proposed a CNN based fast beamforming design technique that converts the downlink beamforming to uplink beamforming and power allocation form. It was observed from simulation analysis that the CNN based fast beamforming technique reduces the computational complexity and iterations as compared to the weighted minimization mean square error (WMMSE) method. Wenchao *et al.* [153] proposed a CNN based DL framework to optimize downlink beamforming in multiuser MISO systems. The DL framework significantly reduces the computational complexity of beamforming design to achieve the near-optimal solution to the SINR balancing problem. However, the use of CNN in designing beamforming optimization techniques is confusing about the input format and phase/magnitude transformation.

The authors of [155] studied a DL based approach to optimize the beamforming vector at the MISO base station, which can be used for any power constraint. The

authors used the transmit power constraint of base stations as side information to learn the impact of power constraint on universal MISO beamforming. The scheme reduces the computational complexity of beamforming optimization while enhancing the average sum rate of the system. However, again the performance of DNN in beamforming optimization is dependent upon the quantity of available dataset. Qisheng *et al.* [156] proposed PrecoderNet, a deep deterministic policy gradient (DDPG) [157] based DRL approach, to maximise the average rate upper bound. The digital beamformer and analog combiner from the previous learning iteration are utilized as state, while the matrices from the current learning iteration are used as an action. Numerical results revealed the effectiveness of the PrecoderNet in reducing bit error rate and time consumption. Although, PrecoderNet is robust to CSI imperfection, but it is based on policy learning method and suffers from the issue of slow learning rate. Xiao *et al.* [158] suggested a decaying DQN framework for integrating reconfigurable intelligent surfaces (RIS) in wireless networks with unmanned aerial vehicles (UAVs), where the RIS is used to improve the UAV's QoS requirements. Results revealed the effectiveness of this DRL based technique in reducing the average energy consumption of the system. DQN works well for discrete state and action spaces but is not well suited for continuous state and action spaces. A relative comparison of the AI enabled beamforming techniques is presented in Table 2.10.

Lessons Learned: Intelligent beamforming design, which can optmize the variables of beamforming vectors and variable radio resources is essential for all emerging 5G and B5G HetNets. Conventional beamforming techniques cannot handle the ultimate complexity of giant antenna arrays with multiple incorporated beamforming active control components. The use of AI techniques can efficiently reduce the computational complexity of beamforming design while maintaining the QoE and QoS for the end-users. According to the literature, AI has been applied in the design of intelligent beamforming techniques by embedding the AI based algorithms on the various types of base stations.

Table 2.10: A relative comparision of the AI enabled beamforming techniques for HetNets

Author(s)	Technique Used	Performance Evaluation Metric	Advantage	Downsides
Zhihan *et al.* [20]	DNN enabled beamforming	SNR and accuracy of pattern recognition	Reduced Computational Complexity, Good performance at low values of SNR	Quality and quantity of the available dataset can affect the performance of DL based technique
Huang *et al.* [154]	CNN based fast beamforming design	Sum Rate	Reduced computational complexity and number of iterations	Uncertainity about input format to CNN and phase/magnitude transformation
Wenchao *et al.* [153]	CNN based DL framework for beamforming optimization	SINR	Reduced computational complexity and near optimal solution to SINR balancing problem	Uncertainity about input format to CNN and phase/magnitude transformation
Junbeom *et al.* [155]	DNN enabled MISO beamforming	Average Sum Rate	Reduced Power Consumption	Quality of dataset affect the performance
Qisheng *et al.* [156]	DDPG based hybrid beamforming	Bit Error Rate, Spectral Efficiency	Robustness to imperfect CSI	Slow learning rate
Xiao *et al.* [158]	DQN based passive beamforming	Energy Consumption, Convergence Rate	Reduced average energy concumption and improved QoS for end-users	Not suitable for continous state-action pairs

2.4.2 Cooperative Jamming/Injection of Artificial Noise

Injection of Artificial Noise (AN) is another technique used in PLS design. In this technique, the transmitter injects AN in the original signal to degrade the quality of the signal at the wiretap channel used by the eavesdropper. AN injection was first used in [159] to improve the physical layer security in wireless communication systems. The generation of AN is carried out based on the status of CSI of the eavesdropper. If the CSI of the eavesdropper is not known to the transmitter, then identical AN is generated. The AN injection is performed in such a manner that it does not affect the legitimate receiver's channel since the AN is generated in its nullspace. Artificial Noise can be used with beamforming to improve the PLS in wireless communication systems further. A PLS model using AN and beamforming is as shown in Fig. 2.2. The transmitted signal after AN injection can be mathematically represented as follows

$$x = ms_a + ns_j \tag{2.6}$$

where x is the final transmitted signal after addition of AN, m and n are the beamforming vectors for data and jamming signal respectively, s_a and s_j are the data jamming signals. Also, the signals received by the legitimate receiver and eavesdropper can be given as follows

$$y_R = h_R ms_a + g_R ns_j + n_R \tag{2.7}$$

$$y_E = h_E ms_a + g_E ns_j + n_E \tag{2.8}$$

where h_R and h_E are the wireless channel responses of the main and wiretap channel, n_R and n_E represent the Gaussian Noise for Receiver and Eavesdropper, respectively.

Cooperative jamming is seen to be a potential physical layer-based technique to secure wireless signals in the vicinity of eavesdroppers. It is a method in which a helping interferer introduces artificial noise to mislead the eavesdropper. A Cooperative Jamming assisted wireless network consists of a transmitter that transmits the signal to the legitimate receiver and a jammer that transmits the jamming signal towards the wiretap channel used by the eavesdropper to degrade it and improve the PLS.

Cooperative relaying can be used along with cooperative jamming, where a relay node is used to transmit data from the transmitter and legitimate receiver. After applying cooperative jamming, the received signal at the receiver y_R and eavesdropper y_E can be given as follows

$$y_R = \sqrt{P_S} h_{TR} s + w h_{JR} z + n_R \tag{2.9}$$

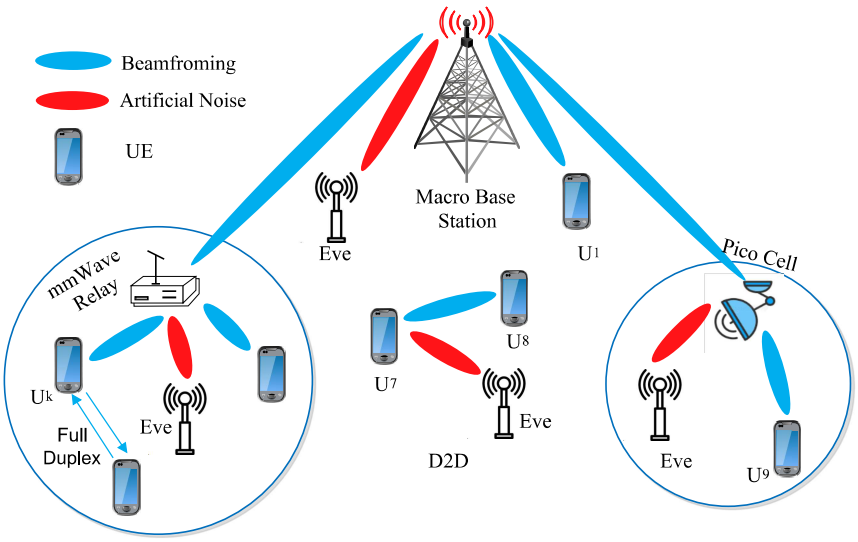

Figure 2.2: Beamforming aided Artificial Noise

$$y_E = \sqrt{P_S} h_{TE} s + w h_{JE} z + n_E \tag{2.10}$$

where P_S is the transmit power, z denotes the jamming signal with w as weight vector. h_{TR}, h_{JR}, h_{TE}, h_{JE} are the channel parameters of Transmitter Receiver channel, Jammer Receiver channel, Transmitter Eavesdropper channel and Jammer Eavesdropper channel, respectively.

Yan *et al.* [18] explored cooperative jamming for a two-tier 5G HetNet having large scale antenna arrays at macro cell base stations. Results revealed the effectiveness of the technique in improving the secrecy performance of the network. Jingjing *et al.* [160] studied cooperative jamming for secure data transmission in D2D enabled HetNets. Simulation results depict the enhanced security performance of uplink cellular HetNet. Zhang *et al.* [47] proposed an artificial noise-aided secure transmission technique for SWIPT in HetNet. Potential eavesdroppers are jammed, and the energy harvesting capability of the system is enhanced by constructing the downlink information beamforming and artificial noise matrix of femtocell base stations and macrocell base stations together. Shiwei *et al.* [161] designed an artificial noise-aided scheme to enhance the PLS of D2D enabled cellular HetNet. More specifically, authors [161] aim to improve D2D users' secrecy capacity in a single cell when D2D users share downlink resources with cellular users. Gong *et al.* [162] proposed an artificial noise (AN) aided beamforming technique to maximize the secrecy rate of wirelessly powered HetNet. The technique maximizes the AN power subject to the secrecy rate requirements of femtocell users. Yan *et al.* [163] investigated cooperative jamming in a two-tier HetNet with massive antenna arrays at the MBS. The results demonstrate the effectiveness of the scheme in improving HetNet's secrecy performance. Also, the authors of [164] presented a secure data transmission system

for D2D-enabled HetNets based on cooperative jamming. According to the simulation results, uplink cellular HetNet has improved security performance.

Challenges in cooperative jamming/artificial noise injection: Almost all known cooperative jamming techniques are centered on the premise that all UEs in a wireless system have perfect or statistical channel state information (CSI). However, it is impossible to precisely assess an eavesdropper's CSI, especially in passive eavesdropping mode. Also, the secure transmission methods described above based on cooperative jamming are at the expense of jamming power usage, which is inappropriate for a low-power communication system such as IoT. *AI assisted cooperative jamming/artificial noise injection*: Recently, AI has got a lot of attention in the design of cooperative jamming based techniques and to solve the aforementioned issues. Yan *et al.* [21] proposed a back-propagation neural network (BPNN) aided cooperative jamming technique for secure data transmission in IoT environment. Simulation results reveal that the proposed technique has better security performance than continuous jamming under energy constraints. However, the problem with BPNN is that any new learning induces catastrophic forgetting once a network has learned one set of weights. Also, it is highly sensitive to noisy/complex data, which makes it less suitable for highly dynamic 5G and beyond HetNet environments. Sangsoek *et al.* [165] proposed a DL-based precoding scheme for MISO wiretap channels. It exploits a DNN to optimize the precoder for artificial noise and information signals. It was observed that the proposed DL based AN scheme performs better than the conventional AN scheme. Although the scheme enhances the secrecy rate of the system but the training of DNN requires high computational complexity, and due to this it is very difficult to achieve *ideal* training in delay-constrained systems. Zhang *et al.* [166] proposed a multi-agent DRL technique for secure communications in UAV enabled wireless network. It aims to solve the trajectory design and cooperative jamming issues in the multi-UAV communication scenario. Specifically, it uses continuous action attention multi-agent deep deterministic policy gradient (CAA-MADDPG) to maximize the secrecy capacity of the system and improve the convergence rate of DDPG based scheme. However, the DDPG based schemes are highly sensitive to hyper-parameters and suffers form issue of instability due to its sensitivity.

Yuxin *et al.* [167] designed a neural network based encoder to optimize the design of cooperative communications in wireless networks. This method employs a two-stage technique to train and reduce the intended losses of a conventional three-node cooperative system using a single autoencoder model. It was observed that the proposed scheme achieves a similar block error rate (BLER) as compared to decode-and-forward (DF) based scheme and outperforms amplify-and-forward (AF) scheme. One advantage of the autoencoder-based scheme is that we don't have to choose a

precise training SNR value based on heuristics. Another is reduced sensitivity to the accuracy of channel estimation. In another study, Yuxin *et al.* [168] proposed a neural network-based autoencoder technique for relay-enabled cooperative communications. Simulation results depict the effectiveness of DL-based technique's over the conventional amplify-and-forward (AF) and decode-and-forward (AF) relay schemes in terms of BLER performance. The advantage of this autoencoder based technique is that it eliminates the need of CSI and noise variance and is adaptive to any input block size. However, autoencoders need a large amount of dataset to generate meaningful results and require high training time. Yuhan *et al.* [169] presented a DRL based technique for cooperative communications and relay selection. In this, a DQN is trained using mutual information and outage probability, and then the best relay is chosen from a pool of relay nodes without using previous data. Simulation results revealed that the proposed scheme reduces the outage probability and energy consumption of the system. However, the authors of [169] have considered the case of stationary sensor nodes, and the effectiveness of the proposed scheme on the mobility of sensor nodes have to be tested, and also the use of DQN for continuous action-state pairs is considerably limited. Table 2.11 depicts a relative comparison of the existing cooperative jamming techniques for HetNets.

Lessons Learned: The deployment of AI based techniques such as ML, DL, and DRL have been adopted as a promising tool to facilitate the cooperative jamming/AN injection in the design of PLS for 5G HetNets. For instance, AI based cooperative jamming methods can deliever comparable or better security performance without compromising the expense of jamming power usage. Also, it has been observed from the literature that the AI based techniques do not need perfect CSI knowledge to secure data transmission from passive eavesdropping attacks. Notably, it is shown that integration of neural network and reinforcement learning with online training can facilitate hyper-parameter tuning and optimize the convergence rate, improving the cooperative jamming performance without having perfect CSI.

Table 2.11: A relative comparision of the AI enabled cooperative jamming techniques for HetNets

Author(s)	Technique Used	Performance Evaluation Metric	Advantage	Downsides
Yan *et al.* [21]	Back propogation neural network aided cooperative jamming	Jamming efficieny	Low energy consumption	Highly sensitive to noisy/complex datasets and suffers from catastrophic forgetting issue
Sangsoek *et al.* [165]	DNN aided artificial noise for MISO wiretap channels	Secrecy Rate	Improved secrecy rate	High computational complexity
Zhang *et al.* [166]	CAA-DDPG based cooperative jamming in UAV enabled wireless networks	Secure rate, transmission power	Low power consumption, works for multple eavesdroppers scenerio	Highly sensitive to hyperparameters and can lead to unstability
Yuxin *et al.* [167]	Autoencoder based cooperative communication	Block error rate	Reduced block error rate at low values of SNR	Requires large amount of dataset and have high training time
Yuxin *et al.* [168]	Autoencoder based learning for relay assisted cooperative communications	Block error rate, SNR	Eliminates the need of CSI and noise variance	Requires large amount of dataset and have high training time
Yuhan *et al.* [169]	DQN aided cooperative communication	Outage probability, system capacity	Enhances system capacity and energy efficiency	Not tested for mobile nodes and DQN is not suitable for continous state-action pairs

2.4.3 Resource Allocation and Power Control

The enormous performance gain of HetNets comes at the cost of complex resource allocation such as subcarrier, power, etc. Resource allocation is critical for controlling signal interference and designing secure data transmission techniques in 5G, and B5G HetNets [44, 170]. In the scope of PLS, resource allocation algorithms are mainly used to optimize the secrecy capacity of wireless channels [171]. Also, power control in a communication system is the adaptive selection of transmitters' output power to ensure optimal security performance. Irrum *et al.* [170] proposed secrecy-ensured optimization of resource allocation in 5G and beyond HetNets. It was observed that the algorithm enhances the secrecy rate, energy-efficiency, and throughput for D2D enabled HetNet. Humayun *et al.* [172] studied secure resource management for secrecy rate maximization in mmWave enabled N-tier B5G HetNet. Simulation results revealed the effectiveness of downlink uplink decoupled user association technique in terms of secrecy rate and secure user association performance. Wang *et al.* [173] proposed a power control based secure transmission scheme for vehicular 5G HetNets. Specifically, the authors used an on board unit (OBU) equipped in vehicles and proposed a multi-level security architecture for OBU to enhance the secrecy outage probability of the network. Marabissi *et al.* [174] proposed a user cell association and cell-activation selection scheme for ultra-dense HetNet. The simulation results show the effectiveness of the proposed scheme in minimizing the energy consumption and to maximize the security of HetNet. Zheng *et al.* [175] explored a jammed wireless network in which the operator adjusts the broadcast power to enhance the system rate and ensuring the QoS requirements of receivers. Also, in [176], a jamming-resistant receiver was proposed to enhance the resilience of the wireless communications against the jamming, and an optimal power allocation scheme was designed to increase the achievable data rate. To deal with jamming attacks more effectively, Luliang *et al.* [177] explored an anti-jamming defense scheme, which incorporates both the power and the spectrum domain of HetNet. Zhang *et al.* [57] formulated a resource allocation optimization problem to maximize the secure capacity of device-to-device (D2D) user equipment in D2D underlaying HetNet. The proposed resource allocation solution significantly enhances the secrecy capacity with a rapid convergence speed. Wang *et al.* [178] proposed a PLS scheme for two-tier HetNet based on average received signal power (ARSP). By incorporating pilot attacks from active eavesdroppers, the authors utilized the stochastic geometry to analyze the secrecy performance of two-tier HetNet. Irrum *et al.* [179] explored a resource allocation optimization problem for D2D enabled 5G HetNets. Simulation results show that the proposed technique increased the network's throughput, secrecy rate, and energy efficiency. In [180], Xin *et al.* explored an alternative di-

rection multiplier method (ADMM) to enhance the secrecy energy efficiency (SEE) of wirelessly powered HetNet.

Challenges in resource allocation and power control: The resource allocation and power control challenges vary depending on the application context. For instance, in heterogeneous vehicular networks, the goal is to modify the user's transmit power appropriately and consider cache optimization and compute offloading in the communication system. In this situation, focus must be on a more pragmatic and complex application domain. However, the conventional resource allocation schemes lack these features. Also, most conventional resource allocation techniques do not consider the mobility of users, which changes the interference scenario and requires more intelligent techniques to manage interference. Furthermore, from the standpoint of the solution process, a more intelligent and self-optimizing algorithm should be developed and created for future HetNets.

AI assisted resource allocation and power control: AI-based techniques can replicate the surrounding radio settings adaptively. To adapt to the needs of secure wireless networks, the training system can dynamically modify its optimization parameters (e.g., transmit power, subcarrier assignment). Thus, the RA issues in HetNets can be solved using intelligent AI-based methods. Yong *et al.* [181] designed a DRL framework (HetDQN) for optimization of resource allocation in HetNets. Specifically, the authors have used DQN to jointly optimize resource allocation, power allocation, and user association under load coupling conditions. Simulation results revealed the effectiveness of HetDQN in terms of fast convergence rate. Although the DQN works well for a discrete set of transmit power in the power allocation problem but is not suitable for continuous values of power. Ismail *et al.* [182] proposed a ML based technique for energy-efficient resource allocation for 5G enabled heterogeneous cloud radio access networks (EE-HCRAN). They used centralized Q learning for joint optimization of resource blocks and power allocation for remote radio heads. According to simulation results, the suggested resource allocation approach can considerably reduce interference, enhance energy and spectral efficiency, and meet users' QoS expectations. Although the Q-learning method works well for a limited set of state-action pairs, when the values of Q-table increase, it becomes very difficult to store a large amount of state-action pairs in Q-learning. Tang *et al.* [183] investigated DRL for dynamic uplink/downlink resource allocation in 5G HetNets. In this, a deep neural network is used to extract the features of the complex network data. Also, a dynamic Q-value iteration-based RL with experience replay memory method is used to adaptively modify the time-division duplex (TDD) uplink/downlink ratio by assessed incentives. The scheme achieves significant improvements in packet loss rate and network throughput compared to conventional TDD-based resource allocation schemes. However, the proposed DRL based scheme

suffers from high computation overhead.

Ding *et al.* [184] studied DRL for joint optimization of power control and user association in HetNets. In this, the authors used a multi-agent DQN framework to maximize the energy efficiency of UEs in OFDMA-based HetNets. Compared to traditional RL, simulation results show that the multi-agent DQN approach offers better convergence and energy efficiency. However, the problem with multi-agent DQN is that it requires considerable time to train all agents. The received agent data becomes less directed and relevant to UEs. Jing *et al.* [185] proposed DRL-based mobility–aware robust proactive resource allocation (MRPRA) technique for HetNets. They explored multi-actor DDPG to exploit and handle prediction for proactive resource allocation optimization efficiently. Simulation results show that the proposed method adapts well to the users' mobility intensities and rate requirements. Although the proposed scheme achieves efficiency and robustness, but it suffers from the issue of slow learning rate. Zhang *et al.* [186] proposed DRL based multi-agent power control technique for HetNets. To be more specific, the authors proposed a multiple-actor-shared-critic (MASC) technique to train the local DNNs individually in a trial-and-error method by treating each access point as an agent with a local DNN. Simulation results revealed that the proposed technique outperforms conventional power control schemes with respect to computational complexity and sum rate. However, it requires a large amount of time to train a DNN and calculate the transmit power. A relative comparison of the existing resource allocation and power control techniques for HetNets is shown in Table 2.12.

Lessons Learned: In the literature, we've seen how AI techniques can be used to solve problems like power control and resource allocation in 5G HetNets. When learning capabilities are built into devices, the processing load on base stations is distributed and the decision-making process becomes decentralized. Due to this, 5G HetNets' base stations will be empowered to conduct additional network services such as cell association, mobility management, power control, and other responsibilities as a result of this decentralized decision. Moreover, the security performance can also be increased using AI enabled resource allocation and power control techniques.

Table 2.12: A relative comparision of the AI enabled resource allocation and power control techniques for HetNets

Author(s)	Technique Used	Performance Evaluation Metric	Advantage	Downsides
Yong *et al.* [181]	DQL based resource and power allocation	Transmit power, spectral efficiency	Good system capacity gain	DQN works well for discrete set of transmit power but not well suited for continous values
Ismail *et al.* [182]	Q learning based resource allocation	Data rate, energy efficiency	Increased convergence speed	It is very difficult to store huge state action-pairs in Q-table and also authors have only considered the case of downlink
Tang *et al.* [183]	Priortized experienced replay-DQN based resource allocation	Throughput, packet loss rate	High network throughput, both uplink and downlink case are considered	High computation overhead
Ding *et al.* [184]	DQN based power control and user association	Energy efficiency	Improved energy efficiency and better convergence performance compared to Q learning	Requires considerable amount of time to train all agents and the agent data becomes less relevant to UEs
Jing *et al.* [185]	Multi-critic based DDPG for proactive resource allocation	Average data rate, average service delay	Good adaptiveness to data rate requirements	Slow learning rate
Zhang *et al.* [186]	Multiple-actor shared-critic based multi agent power control	Sum rate	Improved average sum rate and computational complexity performance	It requires large amount of time to train DNN

2.4.4 Game theoretic approaches

Game theory can be defined as the mathematical analysis of individual or/and cooperative behaviors of players who pick a certain strategy/action to meet their self-interest [187]. In recent years, there has been a surge in interest in using game-theoretic approaches to address various PLS concerns of HetNets. Ahmed *et al.* [188] studied coalitional game for secrecy ensured resource allocation in D2D enabled 5G HetNets. More specifically, the authors investigated a coalitional game to enhance the information-theoretic security of D2D and cellular users in 5G HetNets. Lalropuia *et al.* [189] designed a Bayesian game model to mitigate denial of service (DoS) attacks for small cells in 5G HetNets. It uses a game-theoretical model to analyze the interactions between a network defender and an attacker during a spoofing attack. Yulan *et al.* [190] designed a stochastic game for analyzing the interactions between a user and jammer, where the transmitter and jammer competes with one another to obtain the best defense policies. In [191], a Stackelberg game based beam domain anti-jamming transmission scheme was proposed to minimize the cost of transmission in the presence of a jammer for a downlink massive multiple-input multiple-output (MIMO) system.

Challenges in game theoretic approches: It's nearly hard to define reward functions for both attackers and defenders. Assumption-based strategies cannot be applied in real-time. The most common application of non-cooperative game theory is mechanism design. Although the system can attain Nash equilibria (NE) through mechanism design, it typically only achieves a sub-optimal solution. Also, the cooperative game theory techniques also take only the ideal computing model into account. This abstraction ignores the associated QoS limitations and the variety of operations and dynamic and heterogeneous settings.

Integration of game theory and AI: Weerasinghe *et al.* [192] suggested a DL based game theoretical technique to reduce the risk of jamming attacks. In this, using a non-cooperative security game, the authors describe the interaction between a transmitter that employs chaotic pseudo-random patterns for channel hopping and a smart jammer that uses LSTM model to anticipate the transmitter's frequency hopping patterns. Simulation results show that the LSTM based technique improves the security performance of the system against jamming attacks. However, the problem with LSTM models is that these models take longer time to train, require more memory and are highly sensitive to random weight initialization. Gupta *et al.* [193] studied game theory based privacy-preserving DL technique for IoT devices. Specifically, the authors used a game theory approach in collaborative DL to assess the rationality of mobile edge devices. The proposed scheme addressed the issue of unfair cooperation in collaborative DL among rational IoTD. The proposed scheme

has to be tested on more datasets to validate its accuracy.

Keywhan *et al.* [194] proposed a game theory based Q-learning approach to secure communication systems from adversarial actors. The authors compensated for the lack of data by allowing the ideal policy to be learned, which simulates circumstances where attackers probe system weaknesses and defenders train and renew security policies and devices based on previous data. Simulation results revealed the effectiveness of the proposed Q-learning based technique against irrational attackers. However, the proposed scheme is not capable of detecting new attacks and also there is a lack of coverage of undefined states or actions. Joseph *et al.* [195] suggested a hybrid reinforcement learning and game theory technique to frame cyber-physical systems security. Here, the strategic level is depicted as imperfect information, extended form game, in which the human administrator and malware creator choose protection and attack methods, respectively. Although the scheme enhances the security performance of cyber-physical systems, but it is not suitable in the scenario with imperfect CSI. Feng *et al.* [196] proposed a RL based smart mode selection technique for secure VR broadcasting in D2D enabled 5G HetNets. The authors designed a Q-function and reward function to obtain the optimal transmission rate for VR and optimize VR quality in mmWave HetNet. It has been observed that the proposed scheme enhances the system throughput with a moderate resource cost as compared to conventional broadcasting schemes in 5G HetNets. However, the scheme is not suitable for the imperfect CSI scenario. A relative comparison of the existing game theoretic techniques for HetNets is discussed in Table 2.13.

Lessons Learned: We have seen how AI techniques, specifically DL and reinforcement learning techniques can be used with game theory to improve the wireless networks' security performance. In the reinforcement learning environment, agents are modeled as players in a standard form game who are trying to improve their long-term strategies. Deep learning techniques are also integrated with game theory to improve the security against jamming attacks and for privacy preservation.

Table 2.13: A relative comparision of the AI enabled game theoretic techniques for HetNets

Author(s)	Technique Used	Performance Evaluation Metric	Advantage	Downsides
Weerasinghe *et al.* [192]	DL aided non-cooperative security game	Prediction Error	Improved security against jamming attacks	Requires more memory and takes long time to train
Gupta *et al.* [193]	Collaborative DL and game theory	One Dimensional Loss Value	Solve the issue of unfair cooperation in collaborative DL among rational IoT devices	Should be tested on more datasets and the accuracy of the model must be calculated
Keywhan *et al.* [194]	Q-learning based stochastic game	Accumulation rewards	Perform well against irrational attackers	Cannot detect new attacks and there is no clearly defined reward model for lack of agreement in metrics
Joseph *et al.* [195]	Multi agent reinforcement learning and game theory	Win factor	Enhance security for cyber physical systems	Not suitable for imperfect CSI condition
Feng *et al.* [196]	Reinforcement learning and game theory	System throughput	Enhanced system throughput with moderate resource cost	Not suitable for imperfect CSI condition

2.4.5 Channel Secrecy Codes

Error control codes (ECC) play an important role in designing reliable, secure communication systems. The capacity to establish information-theoretic secrecy depends on the investigation of different coding methods such as privacy amplification and channel resolvability, in which ECC are used to affect the dispersion of stochastic processes [197]. The development of realistic codes for PLS has recently got more recognition. The main objective of the coding is to improve the security against eavesdropping and jamming attacks. In this subsection, we analyze the recent developments in the coding-based PLS. Practical physical layer authentication codes that will be used in 5G and beyond networks are as follows.

2.4.5.1 Low Density Parity Check (LDPC) Codes

On the wireless channel, 5G uses LDPC codes for channel coding. The LDPC codes correct the channel errors by keeping parity bits for a subset of the data bits. LDPC codes are used efficiently by researchers to design PLS techniques. Thangaraj *et al.* [198] shows that absolute confidentiality can be accomplished with any wiretap channel by using LDPC codes. This conclusion lays out a framework for developing secure coding schemes for use over the wiretap channel used by the eavesdroppers.

Later, Rathi *et al.* [199] designed the two-edge style LDPC codes to generalise this coding scheme to binary erasure channel (BEC) of both the legitimate receiver and the eavesdropper. Subramanian *et al.* [200] used the Ramauja graph for a noiseless channel of the receiver and the BEC of the eavesdropper to create LDPC codes of broad girth block length, which achieves a high secrecy rate.

2.4.5.2 Polar Codes

Polar codes are another error correction codes used to design the PLS techniques for 5G and beyond. Mahdavifar *et al.* [201] created a polar coding scheme to obtain the confidentiality for the symmetric wiretap channel under the constraint that the eavesdropper's channel is degraded to the intended user's main channel for the poor secrecy criteria. Furthermore, in [202] and [203], authors demonstrated that this coding scheme can achieve the complete rate-equivocation area (defined in [204]). Other polar coding techniques designed for eavesdropper's channel include concatenating two polar codes [205], concatenating polar and LDPC codes to reduce the security gap [206], and so on.

2.4.5.3 Lattice Codes

Belfiore *et al.* [207, 208], describe a notation of confidentiality advantage for wiretap lattice codes, which represents the eavesdropper's accurate decoding probability.

The confidentiality advantage scales exponentially with the lattice dimension, according to asymptotic theory. In addition, authors in [209] proposes a framework to examine the confidentiality benefits for arbitrary unimodular lattices. In [210], authors propose lattice codes for Rayleigh fading wiretap networks that are suitable depending on the confidentiality gain criteria. In addition, Choo *et al.* [211] suggested a superposition lattice code for the Gaussian Binary Channel with secret message and confidentiality. Nested lattices code designs for mutual jamming, interference channels, and relay networks [212–214], the protection of the continuous mod-lattice system with feedback [215] have been used to secure the physical layer of the 5G wireless network.

The first feasible channel secrecy code design, based on Coset coding, was proposed in [216]. Jeong *et al.* [217] studied energy adaptive LDPC and polar codes [218] for error correction in HetNets. Qinghe *et al.* [219] studied fountain codes for multicast security enhancement in IoT. The simulation findings demonstrate that the proposed method can effectively improve information security while increasing transmission efficiency with low accredited complexity. Kiran *et al.* [220] explored LDPC codes along with Tomlinson-Harashima pre-coding scheme for interference mitigation in HetNets. The suggested approach improves mean, median, and edge user rates significantly.

Challenges in channel secrecy codes: The encoding and decoding complexity of the error-correction techniques must be optimized. Further, modern channel channel secrecy codes offer very low error rates at longer block lengths, however, long blocks are not always acceptable for low-latency applications such as D2D assisted 5G HetNets. While short block codes have great error-correction performance when decoded optimally, creating practical, low-complexity decoding algorithms that can produce close-to-optimal outcomes for short codes is a major challenge. Also, these techniques are vulnerable to timing attacks in post-quantum scenarios.

AI assisted channel secrecy codes: Use of AI techniques can significantly optimize the encoding and decoding complexity of error correction techniques and also can be explored to mitigate timing attacks. Xiaoling *et al.* [221] proposed a unified CNN classifier based decoder for coexistent HetNets. In the training process, the unified decoder is trained to learn the individual properties of encoded codewords that are encoded according to different specifications. Then, the classifier extracts discrete structural features and detect the coding pattern in the deployment phase. The simulation results show that the proposed CNN-based decoder outperforms benchmark methods in terms of reliability over both Rayleigh fading channels additive white gaussian noise (AWGN) channels. Further, in [222], CNN was used to correctly anticipate noise before concatenating a conventional LDPC belief propagation (BP) decoder to enhance BER performance. Specifically, a typical BP decoder is em-

ployed to estimate the coded bits, followed by a CNN to eliminate the BP decoder's estimation mistakes and produce a more accurate assessment of the channel noise. The advantage of proposed CNN based decoder is that its complexity is linear in the length of the filter and is even suitable for channels having large memory. Although the CNN-based decoders reduce the hardware cost, they only take the local information of input data into account. The problems of exploding gradient and overfitting are very common during the training process of CNN.

Zhang *et al.* [223] investigated DRL for construction of LDPC codes. The authors integrate the Monte Carlo tree search (MCTS) and a DNN to guide code development state by state with a long-term perspective. The proposed scheme improved the BER performance of the system and is also flexible for the parameters of arbitary codes construction. However, the proposed DRL based scheme suffers from the issue of slow learning rate. Liao *et al.* [224] proposed reinforcement learning technique SARSA(λ) for construction of polar codes. In particular, this study presents a method for mapping polar-code construction to a game in which the agent is trained to navigate a maze. Experimental results show that the proposed technique can approximate existing standard constructions for SC decoding and surpass the standard construction for SCL decoding with moderate training. However, the authors did not evaluate the effect of channel mismatching in the training process and the proposed scheme needs to be tested for longer codes construction. Tadashi *et al.* [225] suggested DL based trainable projected decoding technique for LDPC codes. It is based on projected gradient descent technique. Numerical analysis shows that the proposed method beats belief propagation decoding in terms of bit error rate and number of iterations. One of the disadvantages of this technique is that it suffers from the issue of vanishing gradient and redundant computation. Table 2.14 shows a relative comparison of the existing channel secrecy coding techniques for HetNets.

Lessons Learned: The use of AI in designing error control codes has increased a lot in recent years. The main strength of AI techniques is that they possess the self-adaptive ability to variety of conditions, unlike the codes generated by conventional methods. For instance, the AI based decoder can be trained to learn the properties of codewords, which are encoded using different specifications. Also AI techniques can be used to anticipate the channel noise correctly to enhance the BER performance of existing decoders.

Table 2.14: A relative comparision of the AI enabled channel secrecy coding techniques for HetNets

Author(s)	Technique Used	Performance Evaluation Metric	Advantage	Downsides
Xiaoling *et al.* [221]	CNN based channel decoder	Bit Error Rate, SNR, Classification Accuracy	Better bit error rate performance for both AWGN and rayleigh fading channels	Problem of overfitting and exploding gradient during the training of CNN
Liang *et al.* [222]	CNN based belief propogation decoding	Bit Error Rate, SNR	Reduced bit error rate at low values of SNR	Takes longer time to train and faces issues such as overfitting and exploding gradient
Zhang *et al.* [223]	DRL based LDPC codes	Error Rate, SNR	Improved frame error rate performance	Slow learning rate
Liao *et al.* [224]	Reinforcement learning based construction of polar codes	Frame Error Rate, SNR	Improved frame error rate performance	Proposed Technique should be tested for longer codes construction
Tadashi *et al.* [225]	DL based LDPC decoder	Mean Square Error, Bit Error Rate	Reduced bit error rate at low values of SNR	Vanishing gradient and redundant computation
Liao *et al.* [226]	Q learning based construction of polar codes	Frame error rate	Low computational complexity compared to SCL decoding	Authors didnot explore effect of channel mismatch during the training process
Sebastian *et al.* [227]	Scaling DL based construction of polar codes	Bit error rate	Reduced latency as compared to belief propogation	Longer training time

2.4.6 Secure Handover Schemes

In HetNets, handover occurs when a mobile station transits from one wireless cell to another, terminating its connection with the first base station and connecting to the second. Zhang *et al.* [228] proposed a robust and universal seamless handover authentication scheme (RUSH) for 5G HetNets. According to the comparative simulation analysis, RUSH surpasses other computing and communication efficiency schemes. Kumar *et al.* [229] suggested a universal subscriber identity module (USIM) and elliptic curve cryptosystem-based secure handover authentication technique for 5G HetNets. Compared to other existing handover schemes, the suggested approach is more efficient in reducing time, storage, and communication cost. Kumar *et al.* [230] suggested secure handover authentication technique for D2D communications in 5G HetNets. The security and performance analysis show that the proposed scheme is capable of resisting a variety of attacks.

Challenges in secure handover schemes : With the emergence of various enabling technologies for fifth-generation (5G) networks, such as mm-wave, massive MIMO, network densification, Internet of things (IoT), and so on, secure handover management is expected to become more difficult, as the number of base stations (BSs) per unit area and the number of connections has been steadily increasing. Further, mmWave and higher frequencies are subject to significant attenuation, implying that their transmission reach will be limited. Furthermore, because service interruptions occur during HOs, customer happiness is significantly impacted, undermining the tremendous promises of 5G networks. Also, establishing a trade-off between secure data transmission techniques, dynamic resource allocation, and handover management is important in designing secure data transmission techniques for 5G and beyond HetNets.

AI assisted secure handover schemes: Alotaibi *et al.* [23] proposed a neural network-based technique to manage vertical handover in HetNets. The authors used artificial neural network (ANN) to improve the chances of getting higher success to find another wireless network during the process of handoff across HetNets. Simulation results show that the proposed handover technique can improve the QoS of both data and voice services while also meeting the user's preferences to a large extent. However, the proposed scheme faces the issue of high computational complexity. Zijun *et al.* [231] suggested a DQN based handover management scheme for dense wireless local area networks (WLANs). The proposed approach allows the network to adapt from actual user actions and network state by adjusting its learning mechanism. Simulation results revealed that the proposed scheme enhances the data rate during the handoff in WLANs as compared to conventional handoff techniques. The limitation of the proposed scheme is that the authors only considered

one agent for simulations and also the simulated WLAN scenario can differ from the actual scenarios.

Rihani *et al.* [232] proposed a neural network-based handover scheme for multiple radio access technologies enabled HetNet. The scheme is built on reinforcement learning to make it smart enough to decide whether or not to execute vertical handover across various wireless networks. The chosen System on Chip (SoC) has a partial reconfiguration (PR) functionality on the field-programmable gate arrays (FPGA), which gives flexibility to the adopted high-performance devices. Although the proposed scheme provides high rate of successful decisions but suffers form the issue of high computation overhead. Mahira *et al.* [233] suggested a feedforward ANN based approach for handover decision in wireless HetNets. Depending on the service cost, data rate, received signal strength indicator (RSSI), and speed of mobile devices speed, the neural network assists in taking the handover and selecting the best option. Compared to other existing methods, the experimental findings demonstrate an improvement in successfully lowering the number of handovers. However, the problem with using feed-forward neural networks is that they are not translation invariance and require a large number of parameters to be optimized. A relative comparison of the existing secure handover techniques for HetNets is presented in Table 2.15.

Lessons Learned: AI enabled intelligent handover optimization schemes would aid in determining the best BSs to connect as well as pre-allocating the resources required at the BSs. For instance, supervised learning algorithms can assist in providing user mobility information by predicting future position, trajectory, cell, and so on, which is required for proactive HO optimization in 5G HetNetss to improve user QoS security performance. While, the unsupervised learning techniques can enable decentralized and scalable handover optimization for 5G and B5G HetNets.

Table 2.15: A relative comparision of the AI enabled secure handover techniques for HetNets

Author(s)	Technique Used	Performance Evaluation Metric	Advantage	Downsides
Alotaibi *et al.* [23]	ANN based handover management	Throughput, packet delay	Enhanced system throughput	High computational complexity
Zijun *et al.* [231]	DQN aided handover management	Data rate	Impoved data rate performance	The authors only considered one agent for simulations and the simulated WLAN may differ from actual scenerio
Rihani *et al.* [232]	ML based handover technique	Successful decisions and Quality of Service	Improved handover performance	High computational overhead
Mahira *et al.* [233]	Multilayer feedforward ANN technique for handover decision	Frame Error Rate, SNR	Reduced number of handovers	Requires large number of parameters to be optimized
Koda *et al.* [234]	Q learning based predictive handover in mmWave enabled wireless networks	Throughput, service disruption time	Enhanced system throughput	Stringent learning rate

2.4.7 Physical Layer Authentication

Physical-layer authentication (PLA) is used to verify the authenticity of a wireless signal and its emitter by examining its physical layer characteristics. To maintain security, a PLA system should ensure that the attacker is unable to launch successful attacks. Ting *et al.* [19] suggested a fast and efficient PLA technique for software defined networking (SDN) enabled 5G HetNets. Simulation results proved that the suggested approach gives dependable security performance for supplementary authentication. Feng *et al.* [235] suggested a mutual PLA scheme to secure 5G HetNets. Simulation results revealed the effectiveness of the proposed scheme in enhancing security through PLA. Pinchang [236] studied PLA for heterogeneous MIMO systems by using hardware and wireless channel features. Simulation analysis revealed the efficiency of the proposed technique. Jiazi *et al.* [237] proposed a novel PLA scheme for HetNets by using AF cooperative relaying. The proposed scheme selects the optimal relay from several AF relays for collaboration between genuine transmitter and intended receiver to mitigate spoofing.

Challenges in PLA: PLA provides several benefits, including minimal computing requirements and network overhead [11, 238, 239], but it also has some drawbacks. The primary reason for this is that most physical layer authentication approaches use static procedures when dealing with the diverse and complex environment of 5G and beyond networks. For instance, the inaccurate estimations and changes of the selected attribute degrade the performance of single attribute-based physical layer authentication methods [240]. It limits the security performance of these methods in many applications, such as device security. Also, the unexpected changes of attributes owing to possible decorrelation at various time instants and device mobility can significantly impact the authentication performance of conventional static authentication systems. As a result, changes in attributes raise the uncertainty for attackers while decreasing the authentication accuracy of legal devices running without learning the various attributes.

AI assisted Physical Layer Authentication: As a consequence of the significant rise in the number of smart devices and to solve the aforementioned issues, AI has already begun to attract a lot of attention in the context of PLA for HetNets. The AI-based PLA methods can detect a large number of UEs simultaneously while retaining good security performance. Fang *et al.* [241] proposed ML based PLA scheme for 5G and beyond HetNets. The ML-based intelligent authentication scheme uses the adaptive authentication process and varying attributes to learn from the dataset without an exact attribute model. Simulation results revealed that the proposed technique provides improved security and continuous adaptive authentication for legitimate UEs. However, the proposed ML based authentication

schems is not suitable for contionous long authentication across low power devices. Liao *et al.* [22] explored multiple DL based PLA techniques for heterogeneous industrial wireless sensor networks (ISWNs). Specifically, three algorithms have been used to implement PHY-layer authentication in IWSNs: DNN-based authentication, the CNN-based authentication, and convolution preprocessing neural network (CPNN)-based authentication of sensor nodes. All the schemes are capable of both lightweight authentication and multiple nodes authentication concurrently. Also, the CPNN based authentication is best among three in terms of authentication and training time performance. However, CPNN based authentication performance is limited in highly dynamic 5G HetNet environments.

Xiaoying *et al.* [242] developed an adaptive neural network to identify anomalies in wireless channel characteristics and assess if an intrusion has happened. It deals with the problem of PLA in time-varying scenarios. Although the proposed scheme provides a good trade-off between computational complexity and performance, there is a high computation cost. Shi *et al.* [243] employed an autoencoder to identify human behavioral characteristics based on daily mobility patterns using Wi-Fi data given by IoT devices. Particularly, the authors build a DL-based user authentication technique to properly identify each user by extracting representative characteristics from CSI values of WiFi signals. The authors claimed to achieve over 91% and 94% authentication accuracy for stationary and walking activities of 11 subjects. However, the autoencoder based techniques require large amount of data for training and have high training time. Xiaofan *et al.* [244] proposed a residual network (ResNet)-based PLA scheme for industrial cyber-physical systems (CPS). Specifically, a CSI based physical layer authentication scheme is designed, where ResNet was used for the classification of CSI measurements. The authentication accuracy is considerably increased, and all of the numerical findings provide important insights on how to enhance such an approach for industrial wireless CPS. However, the training of ResNet-based schemes requires a huge amount of time, making them practically infisible for highly complexed and dynamic practical 5G HetNet scenarios. A relative comparision of the existing PLA techniques for HetNets is shown in Table 2.16.

Lessons Learned: AI based PLA schemes act as the intelligent process for learning and utilizing the sophisticated time-varying environment, thereby improving the robustness and reliability of PLA in wireless networks. AI based PLA schemes are capable of authenticating a large number of UEs and can deliver good security performance. Also, it was observed that AI based schemes can deal effectively with anomalies and intrusion detection for the time-varying scenarios.

Table 2.16: A relative comparision of AI enabled PLA techniques for HetNets

Author(s)	Technique Used	Performance Evaluation Metric	Advantage	Downsides
Fang *et al.* [241]	ML aided PLA	Miss detection rate, mean square error	Cost effective and improved authentication performance for UEs	Limited to continuous long authentication across low-powered UEs
Liao *et al.* [22]	CNN based PLA	Authentication rate, cost value	CPNN provides best improvement in authentication rate and training time	Not ideal for highly dynamic 5G HetNet environment
Xiaoying *et al.* [242]	ML and data compression based PLA	Probability of false alarm, probability of miss detection	Provide good trade-off between performance and complexity	High computation cost
Shi *et al.* [243]	Autoencoder based smart user authentication	Detection accuracy	Resiliant to spoofing attacks	Requires large amount of dataset to train
Xiaofan *et al.* [244]	Residual Network based PLA	Authentication accuracy	Improved authentication accuracy	Requires huge amount of time for training

Table 2.17: A SWOT Analysis of various PLS aware secure data trasnmssion techniques

Techniques	Strength	Weakness	Opportunities	Threats
Security Oriented Beamforming	It assures that the legitimate receiver and the eavesdropper have the best feasible signal quality difference. It enhances the security, energy efficiency and spectral efficiency of wireless networks.	Since numerous antennas and other hardware systems are used, the hardware complexity is increased. In addition to this, time domain beamforming necessitates the use of highly optimised calculation algorithms.	The BF techniques can be further explored with other AI based methods to enhance the PLS of 5G and beyond networks.	Concern on the effectiveness of beamforming for PLS and interference mitigation in ultra-dense HetNets
Cooperative Jamming/Injection of Artificial Noise	Enhances the wireless channel's secrecy rate and secrecy capacity by confusing the eavesdroppers. It can also be used with other PLS techniques to enhance security further.	High jamming power consumption and these methods are based on the assumption that every user equipment (UE) in a wireless system has perfect CSI, which can't be guaranteed for eavesdroppers.	These approaches may be further investigated using edge computing to fulfill data security/hiding needs and user expectations such as minimal latency.	Although these techniques provide an improved secrecy rate, they still have secrecy threats in case of multiple eavesdroppers scenario.
Resource Allocation and Power Control	These techniques are effective in signal interference control and to maximize the secrecy capacity of wireless channels.	The user association issue can have an impact on the design of wireless resource allocation and power management between the macro base station and the small base stations	Due to the optimized use of resources, It looks to be a viable technique for 5G and beyond 5G networks, with the potential to meet future traffic and security demands for various indoor and outdoor communication networks.	Allocation of a large number of resources for dense HetNets remains the biggest issue in these techniques.

Table 2.17: A SWOT Analysis of various PLS aware secure data trasnmssion techniques

Techniques	Strength	Weakness	Opportunities	Threats
Game Theoretic Approaches	When presented with autonomous and competing actors in a strategic context, game theory can assist participants in reaching optimal decision-making. It provides a firm level of security in presence of multiple eavesdroppers.	The notion that participants are aware of their own and others' pay-offs is unrealistic in terms of wireless networks.	Other AI based techniques can be further explored with these techniques to reduce the computational time.	The most significant threat to game theory-based secure data transmission techniques is that they are predicated on the premise that individuals are rational, self-interested and utility-maximizing agents.
Secrecy channel Codes	Ensures reliable and confidential transmission of wireless data.	High encoding complexity.	These techniques must be explored in dense HetNets for 5G and beyond networks.	Vulnerable to timing attacks in post-quantum schemes.
Secure Handover Schemes	Controls which devices are to be remain linked, which aids in mobility management. It prevents the network from being brought down by congestion.	Inefficient handover techniques can lead to issues such as system overload, packet loss, call termination, etc.	Forced handover schemes can be further explored to enhance the security levels in 5G and beyond networks.	Increased risk of increased latency and drop calls.
Physical Layer Authentication	The AI-based PLA techniques can identify a large number of UEs at once while maintaining a high level of security.	Most of the PLA techniques are built using CSI of UEs, which may not be available in all cases.	PLA techniques can be explored with other PLS techniques to enhance the security of wireless networks further.	Concern in its effectiveness with the unexpected changes in attributes

Summary and Insights: AI can help in solving 5G and beyond HetNets security issues in novel ways. The use of AI-enabled PLS techniques has significantly outperformed key-based encryption techniques, as well as traditional PLS techniques. Since it summarises the inherent patterns of hybrid and complex physical layer characteristics, AI is a powerful approach for dealing with complex issues like beamforming optimization, resource allocation and power allocation, physical layer authentication, etc. Furthermore, channel secrecy coding techniques such as polar codes, fountain codes, and LDPC codes have high decoding complexity in current networks. The AI has the potential to dramatically enhance the performance of such techniques. Table 2.17 depicts the SWOT analysis of different secure data transmission techniques in 5G and beyond HetNets.

2.5 Research Gaps

After a detailed analysis of the existing proposals, following research gaps must be considered during the implementation of PLS for 5G HetNets.

2.5.1 Security of Massive MIMO based HetNets

Massive MIMO is one of the "big three" 5G technologies [245], and is considered as a promising solution for transferring large amounts of data efficiently. The use of a large number of antennas at the transmitter and/or receiver (massive MIMO) can considerably improve the wireless network's spectral and energy efficiency. Thus, massive MIMO is critical for addressing many of the technological issues that the 5G HetNet will face, and they can be seamlessly integrated with existing systems and access technologies. Although the use of massive MIMO increases the capability of PLS against passive eavesdroppers, but smart eavesdroppers can take countermeasures. For example, an eavesdropper may position itself near the legitimate user such that the pathways between the legitimate user and the eavesdropper become densely coupled, making secure transmission of data difficult.

2.5.2 Security of NOMA based HetNets

NOMA has been proposed to improve the spectral efficiency of wireless networks by allowing more users to use the same subchannel [51]. The combination of NOMA and HetNet creates a new NOMA-based HetNet capable of high throughput and large connectivity. The fundamental distinction between OFDM-based HetNet and NOMA-based HetNet is that in the NOMA network, the signal interference cancellation (SIC) approach is used at the receivers, allowing several UEs to share a single subchannel for data transmission. In contrast to traditional HetNets, there

are two types of eavesdroppers in NOMA based HetNets, i.e., external (passive) eavesdroppers whose CSI cannot be detected by the transmitter and internal (active) eavesdroppers whose CSI can be recognized by the transmitter. Moreover, the disparity in transmit power and diverse security requirements of UEs are the crucial factors while designing secure data transmission techniques in NOMA based HetNets.

2.5.3 Energy efficiency and Power Control

In the communication system of 5G wireless networks, consumption of power and the subsequent energy-relevant pollution are the major functional and cost-effective concerns. But with an exponential increase in the network traffic and the increase in the number of connected devices make the energy efficiency parameter an essential aspect while dealing with the security of 5G HetNet. Also, to efficiently utilize the resources, power of the transmitter must be controlled. If the devices are randomly deployed under cellular networks, the network's performance gets degraded due to co-channel interference. Also, to fulfill the SNR requirements of cellular users, devices need to limit their power. The power optimization techniques can improve the throughput and energy efficiency of the network. The same can be used to mitigate the co-channel interference. In 5G, especially for a multi-cell scenario, an efficacious power control scheme needs to be addressed.

2.5.4 Privacy of the User Identity and Communications Infrastructure Inheritence

Issues related to the identity of the user have been known since 4G and even previous generations of wireless networks. In 5G, privacy needs to be maintained for the users which can protect the user's identity against different types of cyber-attacks. Further, even though 5G networks are designed to be more secure, 4G information security methods and practices are still supported in 5G networks. If not addressed, the weaknesses in this archaic communication network can be exploited by malicious parties.

2.5.5 Primary and Secondry Authentication

Device and network authentication in HetNets comes under primary authentication. In the primary authentication, the base station must have inbuilt control to know if the device is authentic for a specified network and then decide on data transmission. So, to protect the network and user devices from getting eavesdropped on by unauthorized users, there is a need for primary authentication. Moreover, secondary authentication in HetNets should be designed for the networks which are

not directly linked with user devices but are a part of HetNet, or we can say for secondary networks. Consider the case of the smart grid, where the signals from the smart meters are transmitted to the data concentrator unit by using Wi-Fi as the primary network and other cellular network as secondary. So, here there is an equal need to authenticate the secondary network to protect the data signal from getting jammed or eavesdropped.

2.5.6 Need of Decentralized Security and Security by design

There were less number of physical points of contact in pre-5G networks, making security evaluations and management more effortless. Also, the number of traffic routing points in 5G's vibrant software-based systems is significantly higher. To be completely secure, all of these should be checked. Moreover, in the traditional communication systems, the location of the devices are first identified, authenticated, and then encrypted to carry out the communication. This is a trusted system as it is based on the core network. The security of 5G wireless networks must be ensured from the first phase of the design by embedding the security mechanisms with the architecture of 5G with the focus on access, identity management, integrity controls, storage, and interface encryption, etc.

2.6 Objectives

After analysis of existing proposals, and the research gaps identified, following objectives are finalized by the research commitee:
1. To study and review the existing secure trasmission techniques for 5G-enabled HetNet.
2. To design energy-efficient secure transmission techniques for 5G-enabled HetNet.
3. To verify and validate the proposed techniques by using various performance evaluation metrics.

2.7 Methodology for Objective 1

Various existing physical layer security techniques will be analysed with their limitations. A review will be performed on the issue of possible attacks on the physical layer and existing schemes to mitigate these attacks in 5G enabled HetNet.

2.8 Methodology for Objective 2

This objective aims to design energy-efficient secure data transmission techniques for 5G enabled HetNet. To achieve this objective, we will integrate PLS techniques with

AI to mittigate various physical layer attacks such as eavesdropping and jamming attacks on 5G HetNet scenario. Specifically, we will explore reinforcement learning and deep learning techniques to design new secure data transmission techniques for mmWave and sub 6 GHz scenarios.

2.9 Methodology for Objective 3

To achieve this objective, the proposed schemes will be tested for their accuracy and efficiency using realistic parameters. Since, the target scheme comprises of 5G enabled HetNet, so the proposed scheme will be simulated on the Python programming tool to demonstrate the performance evaluation. On the Python tool, a program will be designed to test the various performance metrics in different scenarios.

Chapter 3

SecBoost: Secrecy-Aware Deep RL based Energy-Efficient Scheme for 5G HetNets

In this chapter, we explore a two-tier HetNet consisting of a sub-6 GHz macrocell and multiple mmWave picocells to maximize pico cells' average secrecy energy-efficiency in the presence of eavesdroppers in realistic time-varying channels. A multi-agent DRL-based secure transmission scheme is proposed to jointly optimize the power control, channel allocation, and beamforming vectors of pico base stations. Table 3.1 and 3.2 depicts the major symbols used in the chapter and relative comparison of the *SecBoost* with existing PLS schemes, respectively.

3.1 Contributions

The key contributions of this chapter are summarized below.

• The secrecy level of the mmWave channel model for multiple picocells consisting of legitimate users and eavesdroppers is investigated. Moreover, we formulate a joint power control, channel allocation, and beamforming optimization problem with an aim of maximizing the secrecy energy-efficiency of picocells.

• A MARL based framework is presented to obtain an optimal policy for joint power control, channel allocation, and beamforming, such that an RL agent (controller) is placed at each pico base station, which try to cooperatively optimize the policy by using Markov decision process (MDP).

• A multi-agent cooperative DRL based scheme, *SecBoost* is proposed to boost the secrecy energy efficiency of picocells. The proposed *SecBoost* scheme exploits the channel allocation, beamforming, and power control domain, along with the dueling structure of dueling double deep Q-network (D3QN).

• Finally, the SEE performance of *SecBoost* is compared with MARL, MA-DQN, and JBF-SEEM schemes.

Table 3.1: Major symbols used

Symbol	Description
h_k^M	Channel vector from MBS to k–th MU
$h_{n,np}^P$	Channel vector from PBS_n to p–th PU of n–th picocell
y_k^M	Recieved signal at k–th MU
w_k, w_{np}	Beamforming vectors for k–th MU and for p–th PU of n–th picocell
s_k, s_{np}	The information signal intended for k–th MU and for p–th PU of n–th picocell
z_k, z_{np}	Additive white Gaussian Noise at k–th MU and at p–th PU of n–th picocell
y_{np}^P	Recieved signal at p–th PU of n–th picocell
γ_{np}^P	SINR of p–th PU of n–th picocell
γ_k^M	SINR of k–th MU
y_{nE}^P	Signal recieved at Eve, intended for p–th PU of n–th picocell
$h_{n,nE}^P$, h_{nE}^M	channel vectors from n–th PBS to Eve and from MBS to Eve
γ_{nE}^P	SINR of Eve
P_N^{Total}	Total power consumption of the picocells
R_{np}^P	Data transmission rate from PBS_n to PU_{np}
R_{nE}^P	Eavesdropping rate at PU_{np}
$R_{np,sec}^P$	Achievable secrecy rate at PU_{np}
$R_{np,sec}^{P,min}$	Minimum required secrecy rate at PU_{np}
ϖ	Discount Factor
λ	Learning Rate
N_I	Number of sub-channels

Table 3.2: Comparision of *SecBoost* with pre-existing HetNets' secrecy optimization schemes

References	Power Control	Resource Allocaton	Energy-Efficient	Consideration of the mmWave	Consideration of the Sub-6 GHz	Contionous Action Space
[162]	-	-	-	-	-	-
[174]	-	-	✓	-	-	-
[129]	-	-	-	-	-	-
[246]	-	-	-	-	-	-
[57]	-	✓	-	-	-	-
[178]	-	-	-	✓	✓	-
[247]	-	-	✓	-	-	-
[179]	-	✓	✓	-	-	-
[180]	-	-	-	-	-	-
[248]	-	-	✓	-	-	-
Proposed *SecBoost* scheme	✓	✓	✓	✓	✓	✓

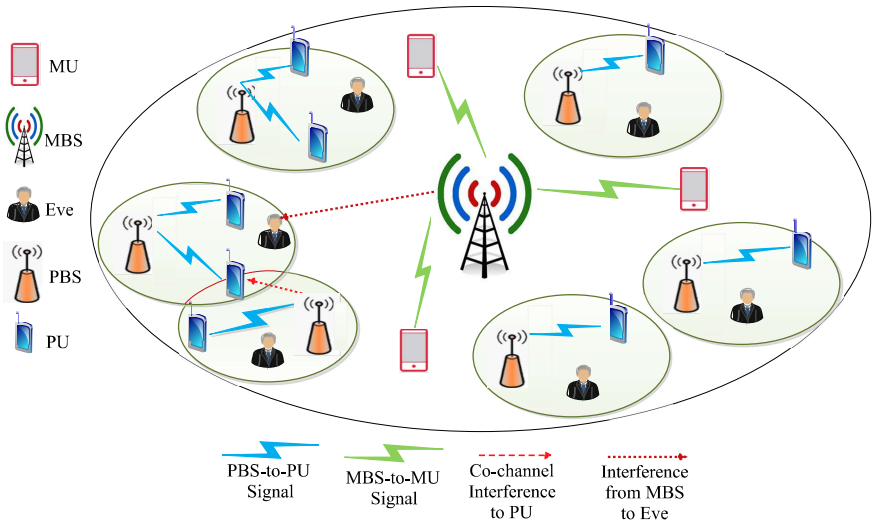

Figure 3.1: System Architecture

3.2 System Model

We consider a downlink two-tier heterogeneous network as shown in Fig. 3.1, where first tier consists of a N_M-antenna macro base station (MBS) and the set of $\mathcal{K} = \{1, 2, \ldots, K\}(\forall k \in \mathcal{K})$ macrocell users (MUs). The second tier comprises of $\mathcal{N} = \{1, 2, \ldots, N\}(\forall n \in \mathcal{N})$ picocells and each picocell consists of a N_P-antenna pico base station (PBS) and the set of $\mathcal{P} = \{1, 2, \ldots, P\}(\forall p \in \mathcal{P})$ picocell users (PUs). The PBS communicates with the PUs using mmWave. Also, to limit the effect of co-channel interference, the transmit power P_p of all the PBSs is considered to be equal. The transmit power of MBS is assumed to be P_M. Also, each picocell has an eavesdropper which tries to detect the desired signal for PUs. Orthogonal frequency division multiplexing (OFDM) is used for resource allocation.

3.2.1 Sub-6 GHz Channel Model for Macrocell

Sub-6 GHz channel model is considered for macrocell in this study. In Sub-6 GHz bands, the frequency directed from a base station is less than 6 GHz [249]. Also, we consider that the sub-6 GHz channel experiences an independent and identically distributed Rayleigh fading. The channel vector of MBS to k–th MU can be mathematically represented as

$$h_k^M = \sqrt{x_k^M} u_k^M \tag{3.1}$$

where u_k^M is the small scale fading vector, and x_k^M is the path loss represented as

$$x_k^M = \left(\frac{c}{4\pi f_c}\right)^2 (R_k^M)^{-\alpha_\mu} \tag{3.2}$$

where c is the light speed (m/s) in vacuum, f_c represents carrier frequency, R_k^M is distance of MBS to k–th MU, and α_μ is the exponent of path loss.

3.2.2 Millimeter Wave Channel Model for Picocells

mmWave channel model is used for picocells in this study. The mmWave have frequency range from 24 to 100 GHz [249]. The channel vector of PBS_n to p–th PU of n–th picocell, similar to [250, 251] is represented as

$$h_{n,np}^P = \sqrt{X_{n,np}^P} a_s(\phi_{n,np}^P) a_u(\beta_{n,np}^P) \tag{3.3}$$

where $a_s(\phi_{n,np}^P)$ and $a_u(\beta_{n,np}^P)$ are the steering vectors, $\phi_{n,np}^P$ and $\beta_{n,np}^P$ are the angles of arrival and departure at PBS_n and PU_{np}, and $X_{n,np}^P$ is the path loss between PBS_n and PU_{np} which can be formulated as follows.

$$X_{n,np}^P = \begin{cases} c_l (R_{n,np}^P)^{-\alpha_l}, & \text{for line-of-sight (LoS)} \\ c_n (R_{n,np}^P)^{-\alpha_n}, & \text{otherwise} \end{cases} \tag{3.4}$$

where c_l, c_n, α_l, and α_l are the path loss intercepts and path loss exponents for line-of-sight (LoS) and non-LoS communication links, respectively. $R_{n,np}^P$ denotes distance between p–th PU and n–th PBS.

3.2.3 Signal Model

Signal received at k–th MU, p–th PU and Eve of n–th picocell is given as follows.

$$y_k^M = h_k^M w_k s_k + h_k^M \sum_{j=1, j\neq k}^{K} w_j s_j \tag{3.5}$$

$$\begin{aligned} y_{np}^P = h_{n,np}^P w_{np} s_{np} + h_{n,np}^P \sum_{i=1, j\neq p}^{P} w_{ni} s_{ni} \\ + \sum_{l=1, l\neq n}^{N} \sum_{i=1}^{P} h_{l,np}^P w_{li} s_{li} + z_{np} \end{aligned} \tag{3.6}$$

$$y_{nE}^{P} = h_{n,nE}^{P} w_{np} s_{np} + h_{n,nE}^{P} \sum_{i=1, j\neq p}^{P} w_{ni} s_{ni} + \sum_{l=1, l\neq n}^{N} \sum_{i=1}^{P} h_{l,nE}^{P} w_{li} s_{li} + h_{nE}^{M} \sum_{k=1}^{K} w_k s_k + z_{nE} \tag{3.7}$$

where h_k^M shows the channel vector of MBS to k–th MU. w_k and s_k represent the beamforming vector and information signal intended for k–th MU. Similarly, w_{np} and s_{np} are the beamforming vector and information signal intended for the p–th PU of n–th picocell. $z_k \sim \mathcal{CN}(0, \delta_k^2)$ is the additive white Gaussian noise (AWGN) at the k–th MU. $h_{n,np}^P$ is the channel vector from n–th PBS to PU_{np}. w_{np} and s_{np} represent the beamforming vector and information signal intended for PU_{np}. $z_{np} \sim \mathcal{CN}(0, \delta_{np}^2)$ represents the AWGN at the PU_{np}. $h_{n,nE}^P$ and h_{nE}^M represents the channel vectors from n–th PBS to Eve and from MBS to Eve, respectively. $z_{nE} \sim \mathcal{CN}(0, \delta_{nE}^2)$ is the AWGN at the Eve placed in n–th picocell. Also, beamforming vectors for pico cell users assure power constraint of $\sum_{p=1}^{P} w_{np} = P_p$ and the beamforming vectors for macro cell users satisfy power constraint of $\sum_{k=1}^{K} w_k = P_M$.

The achievable secrecy rate at PU_{np} is formulated as follows.

$$R_{np,sec}^{P} = \{log_2(1 + \gamma_{np}^{P}) - \log_2(1 + \gamma_{nE}^{P})\}^{+} \tag{3.8}$$

where $\{z\}^+$ denotes $\max\{z, 0\}$. γ_{np}^P and γ_{nE}^P are the signal-to-interference plus noise ratio (SINR) at p–th PU and Eve of n–th picocell, respectively, given as follows.

$$\gamma_{np}^{P} = \frac{|h_{n,np}^{P} w_{np}|^2}{I_C + I_D + \delta_{np}^2} \tag{3.9}$$

$$\text{where} \quad I_C = \sum_{i=1, i\neq p}^{P} |h_{n,np}^{P} w_{ni}|^2$$

$$\text{and} \quad I_D = \sum_{l=1, l\neq n}^{N} \sum_{i=1}^{P} |h_{l,np}^{P} w_{li}|^2$$

$$\gamma_{nE}^{P} = \frac{|h_{n,nE}^{P} w_{np}|^2}{I_F + I_G + I_H + \delta_{nE}^2} \tag{3.10}$$

$$\text{where} \quad I_F = \sum_{i=1, i\neq p}^{P} |h_{n,nE}^{P} w_{ni}|^2$$

$$I_G = \sum_{l=1,l\neq n}^{N}\sum_{i=1}^{P}|h_{l,nE}^{P}w_{li}|^2$$

$$\text{and } I_H = \sum_{k=1}^{K}|h_{nE}^{M}w_k|^2$$

Moreover, the total secrecy rate of pico cells is defined as follows.

$$R_{Total,sec}^{P} = \sum_{n=1}^{N}\sum_{p=1}^{P}\{log_2(1+\gamma_{np}^{P}) - \log_2(1+\gamma_{nE}^{P})\}^{+} \tag{3.11}$$

Also, the total power utilization of pico cells is given as follows.

$$P_{\mathcal{N}}^{Total} = \zeta\Big[\sum_{k=1}^{K}\sum_{p=1}^{P}\parallel w_{np}\parallel^2 + \sum_{l=1,l\neq n}^{N}\sum_{i=1}^{P}\parallel w_{li}\parallel^2\Big] + \mathcal{N}(N_P P_a + P_b) \tag{3.12}$$

where ζ is power amplifier coefficient, P_a shows the power consumption by each PBS antenna, and P_b depicts the basic power consumed by each PBS.

The objective of this chapter is to maximize the average secrecy energy efficiency (SEE) of pico cells by jointly optimizing power control, channel allocation, and beamforming vectors of pico cells while satisfying the SINR and minimum secrecy rate requirements. Thus, the optimization problem is mathematically formulated as follows:

$$\begin{aligned}
\mathcal{P}.\mathcal{F}. &: \max_{w_k,w_{np}} \frac{R_{Total,sec}^{P}}{P_{\mathcal{N}}^{Total}} \\
s.t.\ \ C1 &: \sum_{k=1}^{K}\|w_k\|^2 \leq P_M \\
C2 &: \sum_{p=1}^{P}\|w_{np}\|^2 \leq P_p \\
C3 &: \gamma_k^M \geq \gamma_k^{M,\min} \\
C4 &: \gamma_{np}^P \geq \gamma_{np}^{P,\min} \\
C5 &: R_{np,sec}^P \geq R_{np,sec}^{P,\min}
\end{aligned} \tag{3.13}$$

where $\gamma_k^{M,\min}$ and $\gamma_{np}^{P,\min}$ are the minimum SINR requirements of MU_k and PU_{np}. $R_{np,sec}^{P,\min}$ is the minimum secrecy rate requirement of PUs. Constraints $C1$ and $C2$ states that the beamforming vectors for MUs and PUs satisfy the power constraints of $\sum_{p=1}^{P} w_{np} = P_p$ and $\sum_{k=1}^{K} w_k = P_M$. Constraints $C3$ and $C4$ satisfy the SINR requirements of macro and picocell users. Constraint $C5$ states that the secrecy rate

of picocell users should be greater than the minimum secrecy rate requirement.

Due to the presence of continuous variables and interference terms in γ_k^M and γ_{np}^P, the optimization problem given in (3.13) is non-convex. Also, in practical HetNet systems, wireless channel quality and the capacity of MUs and PUs alter dynamically. Thus, to obtain the optimum solution, we translate the above-mentioned optimization problem into a MARL problem using model-free RL approach. Model-free RL is a dynamic programming technique for finding an optimum solution to the decision-making problems in dynamic contexts [252].

3.3 SecBoost: Multi-Agent Cooperative DRL based SEE Maximization

In this section, first of all, we discuss the rationale behind the use of reinforcement learning. Then, we convert the optimization problem formulated in (3.13) as a MARL problem using MDP. Also, we investigate the multi-agent Q-learning method to jointly optimize the power control, channel allocation, and beamforming vectors to maximize the SEE of picocells. Finally, we have proposed *SecBoost*, a multi-agent cooperative DRL-based SEE maximization scheme for 5G HetNets.

3.3.1 Rationale behind the use of Reinforcement Learning

As discussed in Section III, the problem formulated in (3.13) is non-convex optimization problem with five constraint conditions. Also, in a realistic HetNet environment, the number of users, channel quality, CSI, data rate, and secrecy rate are all time-varying. Thus, applying classic optimization techniques that transform a dynamic environment into a static environment can degrade the secrecy performance and may not be able to satisfy all the constraint conditions. Therefore, we solve the optimization problem formulated in (3.13) using DRL, which is very well known for its effectiveness in dynamic environments [253]. DRL uses the previously stored experiences in the replay buffer to update the network parameters in real-time while also keeping track of the heterogeneous dynamic environment. Moreover, it is difficult to obtain instantaneous global CSI of dynamic HetNet, since global CSI often varies frequently [254]. The proposed scheme does not require prior knowledge of instantaneous global CSI. It can be estimated during the learning process of DRL. Thus in view of the above discussion, we propose a multi-agent DRL-based scheme, i.e., *SecBoost*, to solve the stated optimization problem.

3.3.2 Markov Decision Process

We formulate the MDP as a tuple $(\mathcal{S}, \mathcal{A}, r, \mathcal{T}_{ss'}, \lambda)$, where $\mathcal{S}$ represents the state space set, $\mathcal{A}$ denotes the set of actions, r is the reward function, $\mathcal{T}_{ss'}$ shows the transition probability from state s to s', and $\varpi \in (0, 1)$ is the discount factor. Here, the controller at every PBS acts as an agent with respect to the HetNet environment. The key elements of the MDP are as follows:

3.3.2.1 Agent

We consider the RL agent (controller) at each PBS capable of obtaining adequate environmental information and has enough memory and resources to support deep learning. Define $c \in \{c_1, c_2, \ldots, c_N\}(\forall n \in \mathcal{N})$, where c_n is controller at PBS_n. The agent acquires experience from the wireless HetNet environment through the trial and error method.

3.3.2.2 State Space

The state space $s \in \mathcal{S}$ contains the channel information of FUs and MUs, interference to the agents, minimum secrecy rate required, and secrecy rate of the last time slot, which is given as follows:

$$s = \{h_k^M, h_{n,np}^P, h_{n,nE}^P, I_c, R_{np,sec}^P, R_{np,sec}^{P,min}\} \tag{3.14}$$

where h_k^M, $h_{n,np}^P$, $h_{n,nE}^P$ are the channel vectors of k–th MU, p–th PU of n–th picocell, and eavesdropper of n–th picocell. I_c denotes the interference to the agents, and $R_{np,sec}^P$ represents the secrecy rate. In this work, each agent (controller) at PBSs concurrently observes the unknown environment. Specifically, using reinforcement learning, at each time step, each agent at PBSs recieves observation of the local environment, which consists of the pilot signal from PUs, through which the local CSI is estimated. Further, each agent at PBS shares their local information with other agents in order to collect global CSI. Here, the prior geometric information of the HetNet environment is used to obtain the approximate location of Eve and the corresponding channel information. Specifically, the agent at PBS explore the prior geometric information of HetNet environment to estimate the possible location of Eve, and thus estimate the approximate channel vector of the eavesdropping channel.

3.3.2.3 Action Space

According to the observed state s, the agent chooses the subchannel N_i, the transmit power level of PBS, and the beamforming vector for each PU. Therefore, the action

$a_{c_n} \in \mathcal{A}$ of c_n–th agent is defined as follows.

$$\begin{aligned} a_{c_n} = \{\{&P_{p\{1,c_n\}}, P_{p\{2,c_n\}}, \ldots, P_{p\{P,c_n\}}\}, \\ &\{N_{\{1,c_n\}}, N_{\{2,c_n\}}, \ldots, N_{\{P,c_n\}}\}, \\ &\{w_{n\{1,c_n\}}, w_{n\{2,c_n\}}, \ldots, w_{n\{P,c_n\}}\}\} \end{aligned} \tag{3.15}$$

3.3.2.4 Transition Probability

$\mathcal{T}(s'|s,a)$ is the probability of transition from current state $s \in \mathcal{S}$ to the next state $s' \in \mathcal{S}$, after executing joint action of all agents.

3.3.2.5 Reward Function

When the agent performs an action in the current state, the reward function serves as a signal to verify the goodness of the policy. Also, MARL can be classified into two types: MARL with centralized rewards and MARL with decentralized rewards. All agents in MARL with centralized reward receive a shared (common) reward [255]. While, in MARL with decentralized rewards, each agent receives a unique reward [256]. In this study, we have selected MARL with a centralized reward. We train a single policy using the collective experiences of all the agents, which is followed by all the agents at FBSs. Moreover, we have used a centralized reward calculator to calculate a centralized reward based on the joint actions of all the RL agents at FBSs, which is emitted to all the FBS agents [257]. Also, the reward function reflects the optimization goal, and our goal is to maximize the secrecy rate of PUs while satisfying all the constraints. The common reward function of all the agents is expressed as follows.

$$r = R^{P}_{np,sec} \tag{3.16}$$

In the MDP model, an agent observes a state s_t at time t and executes an action a_t in response, depending on the policy π.

3.3.3 Multi-Agent Cooperative Q-learning

In multi-agent Q-learning, the goal of every agent is to choose an optimal policy $\pi^*_{c_n}$ to maximize the expected discounted reward by choosing an action $a_{c_n,t} : \pi^*(s_t) : \mathcal{S} \rightarrow \mathcal{A}_{c_n}$. At state s_t and policy π_{c_n}, the discounted cumulative reward function is given as

$$D^{\pi_{c_n}}_{c_n}(s_t) = \sum_{t=1}^{\infty} \varpi_t r_t(s_t, a_{c_n,t}|s_0 = s_t, \pi_{c_n}) \tag{3.17}$$

where $\varpi \in (0, 1\}$ is discount factor. Furthermore, the optimal value of discounted cumulative reward is given as follows:

$$D^*_{c_n}(s_t) = \max_{\pi_{c_n}} P^{\pi_{c_n}}(s_t) \tag{3.18}$$

The learning agent's goal is to determine the best policy $\pi^*_{c_n}$. Here, Q-learning is used to learn the $\pi^*_{c_n}$. The Q function, which is utilized in Q-learning, is defined as

$$Q^{\pi_{c_n}}_{c_n}(s_t, a_{c_n,t}) = \mathbb{E}\Bigg[\varpi \sum_{s_{t+1}} \mathcal{T}\varpi_t(s_{t+1}|s_t, a_{c_n,t}) \sum_{a_{c_n,t+1}} \pi_j(s_{t+1}, a_{c_n,t+1}) Q^{\pi_{c_n}}(s_{t+1}, a_{c_n,t+1}) + r_t\Bigg] \tag{3.19}$$

and the optimal value of $Q(s_t, a_t)$ is given as

$$Q^*_{c_n}(s_t, a_t) = \varpi \max_{a_{c_n,t+1}} Q^*(s_{t+1}, a_{c_n,t+1}) + r_t \tag{3.20}$$

Also, the Q-value can be updated as:

$$Q_{(c_n,t+1)}(s_t, a_{c_n,t}) = Q_{(c_n,t)}(s_t, a_{c_n,t}) + \lambda_t[r_{(t+1)} + \varpi \max_{(a_{c_n,t+1})} Q_{(c_n,t)}(s_{t+1}, a_{(c_n,t+1)}) - Q_{(c_n,t)}(s_t, a_{(c_n,t)})] \tag{3.21}$$

where $\lambda_t \in (0, 1]$ denotes the learning rate.

Finally, when the optimal Q function $Q^*_{c_n}(s_t, a_{c_n,t})$ is obtained, the optimal policy can be obtained by

$$\pi^*_{c_n}(s_t) = \arg\max_{a_{c_n,t}} Q^*_{c_n}(s_t, a_{c_n,t}) \tag{3.22}$$

Algorithm 1 describes the multi-agent cooperative Q-learning for joint power control, channel allocation, and beamforming optimization problem. More specifically, agents apply the multi-agent RL method of Q-learning to obtain an optimal policy according to the current state and Q-function. During the training process, the proposed RL based technique acquires the Q-value for each policy. Further, agents observe the next state s_{t+1} to update the Q-function at t time slot. Also, the exploration-exploitation tradeoff must be included in the action selection mechanism during the training process [258]. We have used a greedy approach in the action selection mechanism to balance the exploitation of the present optimal Q-value function. In this, the action a_{c_n} is selected with ϵ probability and the best action $a^*_{c_n}$ is selected with $\epsilon - 1$ probability. Every episode of training lasts T steps, and the execution action a_{c_n} is chosen with the greedy policy from the projected Q-value at each step of an episode.

Algorithm 1 Multi-agent cooperative Q-learning for SEE maximization

Input:
1) ϖ;
2) λ;
3) Minimum Secrecy rate required $R_{np,sec}^{P,min}$.
Output: Optimal set of actions

Each agent initializes $Q_j(s_t, a_{c_n,t})$ to 0.
for $(Episode = 1; Episode \leq N; Episode++)$ **do**
    Initialize s_t.
    for $(t = 1; t \leq T; t++)$ **do**
        Select action $a_{c_n,t}$, according to current policy, .
        Compute the value of SINRs of the macro user, pico user, and Eve, i.e., γ_k^M,γ_{np}^P,γ_{nE}^P.
        Compute Secrecy rate of pico users, i.e., $R_{np,sec}^P$.
        Compute average SEE of pico cells.
        Calculate the reward r_t.
        Each agent gets new state s_{t+1}, set $s_t \rightarrow s_{t+1}$.
        Each agent updates Q values.
    end for
end for

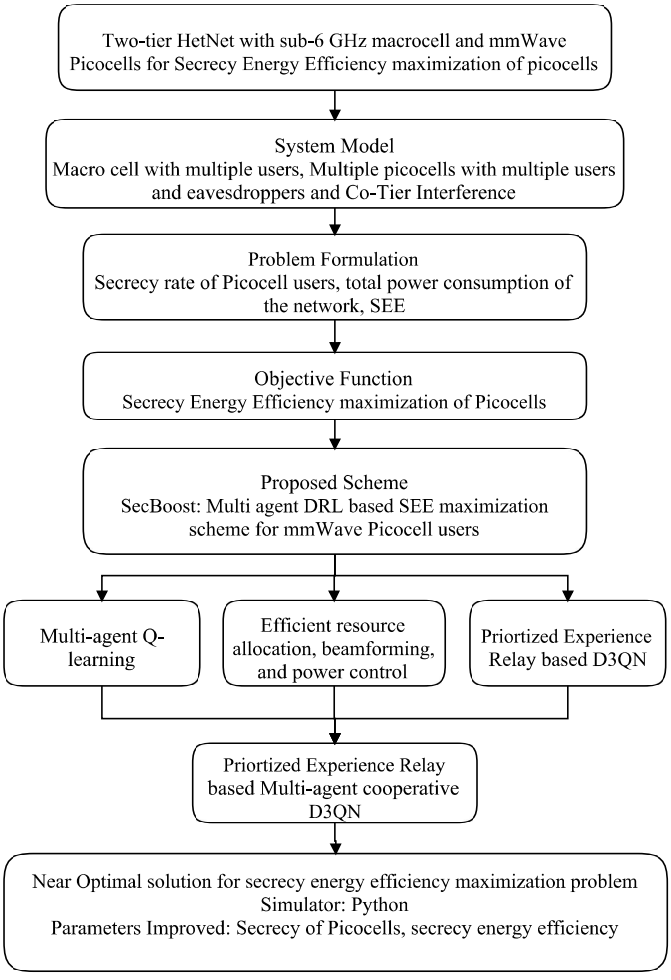

Figure 3.2: Progression of steps in SecBoost scheme

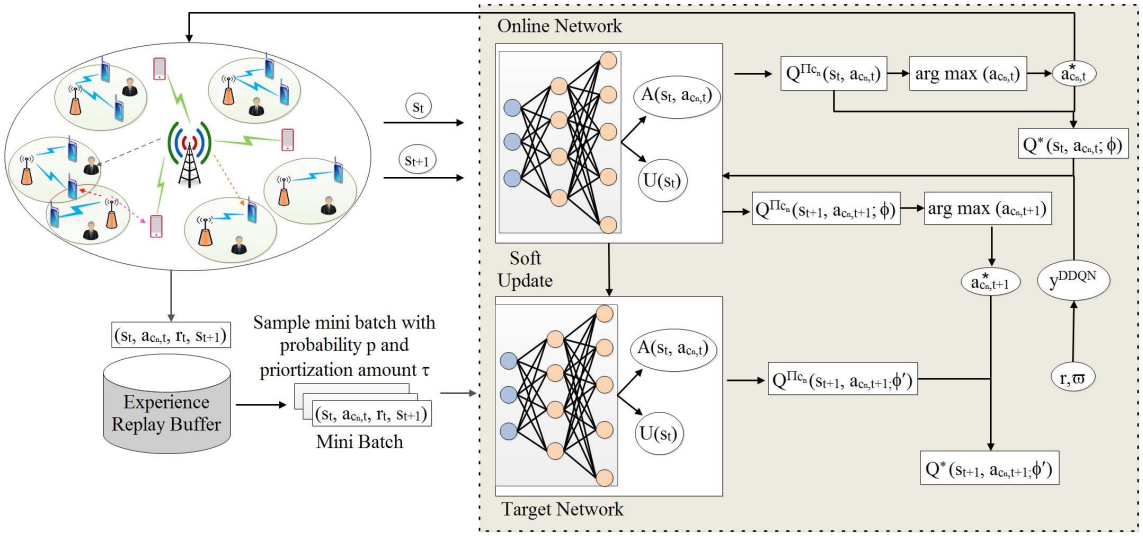

Figure 3.3: Architecture of SecBoost

Algorithm 2 Multi-agent cooperative DQN for SEE Maximization of Picocells

Input:
1) ϖ;
2) λ;
3) Batch size $\flat$;
4) Minimum Secrecy rate required $R_{np,sec}^{P,min}$;

Output: Optimal set of actions

Initialize replay memory having size V
Initialize online Q network $Q_{c_n}(s, a_{c_n}; \phi)$ for all controllers (agents).
Initialize target network parameters for all controllers.
for $(Episode = 1; Episode \leq N; Episode++)$ **do**
    Initialize s_t.
    for $(t = 1; t \leq T; t++)$ **do**
        Select action $a_{c_n,t}$ at state s_t from $Q_{c_n}(s, a_{c_n}; \phi)$ according to current ϵ-greedy policy.
        All controllers take joint action.
        Compute value of SINRs of the macro user, pico user, and Eve, i.e., γ_k^M,γ_{np}^P,γ_{nE}^P.
        Compute Secrecy Rate of pico user, i.e., $R_{np,sec}^P$.
        Compute average SEE of pico cells.
        Compute the reward r_t.
        Each controller obtains new state s_{t+1}, set $s_t \rightarrow s_{t+1}$.
        Store a tuple of $(s_t, a_{c_n,t}, r_t, s_{t+1})$ in replay memory V.
        Select a small batch of size $\flat$ uniformly from V with the help of stochastic gradient for minimizing the loss between online and target network using $L_{c_n}(\phi) = \mathbb{E}[(y_{c_n}^{DQN} - Q_{c_n}(s_t, a_{c_n,t}; \phi))^2]$
        Update the weights of target network
    end for
end for

3.3.4 Multi-Agent Cooperative DRL based SEE Maximization

In general, the MARL method works smoothly with considerably limited and discrete state and action spaces [259,260]. However, if the state spaces and action spaces are continuous or have a large dimension, using a Q table to store all state-action pairs is unfeasible. In the HetNet scenario, with a rising number of state-action spaces, size of Q-table also increases significantly. We also have continuous action spaces due to continuous variables such as power.

Algorithm 3 SecBoost: A Multi-Agent Cooperative DRL based SEE Maximization Scheme

Input:
1) ϖ;
2) λ;
3) $\flat$;
4) Minimum Secrecy rate required $R_{np,sec}^{P,min}$;
Output: Optimal set of actions

Initialize replay memory
Initialize online Q network $Q_{c_n}(s, a_{c_n}; \phi)$ for all controllers (agents).
Initialize target network parameters for all controllers.
for $(Episode = 1; Episode \leq N; Episode++)$ **do**
    Initialize s_t.
    for $(t = 1; t \leq T; t++)$ **do**
        Select action $a_{c_n,t}$ at state s_t from $Q_{c_n}(s, a_{c_n}; \phi)$ according to current ϵ-greedy policy.
        All controllers take joint action.
        Compute the SINRs of the macro user, pico user, and Eve, i.e., γ_k^M,γ_{np}^P,γ_{nE}^P.
        Compute Secrecy Rate of pico user, i.e., $R_{np,sec}^P$.
        Compute average SEE of pico cells.
        Compute the reward r_t.
        Each controller obtains new state s_{t+1}, set $s_t \rightarrow s_{t+1}$.
        Store a tuple of $(s_t, a_{c_n,t}, r_t, s_{t+1})$ in replay memory of size V.
        Select a small batch transition of size $\flat$ with probability $p = \frac{|\delta|^\tau}{\sum_{j'} |\delta(j')|^\tau}$, where τ is the prioritization amount.
        Determine the TD error using $\delta = y_{c_n}^{DDQN} - Q_{c_n}(s_t, a_{c_n,t}; \phi)$.
        Update the mini-batch priority using δ.
        Each controller performs gradient descent step on $\mathbb{E}[(y_{c_n}^{DDQN} - Q_{c_n}(s_t, a_{c_n,t}; \phi))^2]$.
        Update the weights of target network
    end for
end for

Therefore, to overcome above-mentioned issues, multi-agent cooperative DQN is proposed to approximate the Q function $Q_{c_n}(s, a_{c_n}; \phi)$ by pairing multi-agent Q learning with Deep Neural Network (DNN), where ϕ denotes its weights. Since huge storage space is required for state-action pairs (Q-values), the DRL agent only keeps weights in its local memory, thus reducing the computation complexity. In this, a

neural network function approximator $Q_{c_n}(s, a_{c_n}; \phi) \approx Q^*_{c_n}(s, a_{c_n})$ is used as an online Q network. The multi-agent cooperative DQN uses a target network in addition to the online network to keep the total network performance stable. Each agent c_n has a DQN in a multi-agent DRL configuration that accepts the current state s_t as an input and Q-value function as the output for all actions. Every agent captures and saves their experiences as a tuple $(s_t, a_{c_n,t}, r_t, s_{t+1})$. A mini-batch of the data is uniformly sampled from memory and utilized to update the online network weights ϕ in each iteration.

Using a variant of stochastic gradient descent, the Q network is trained by minimizing the loss function given by

$$L_{c_n}(\phi) = \mathbb{E}[(y^{DQN}_{c_n} - Q_{c_n}(s_t, a_{c_n,t}; \phi))^2] \tag{3.23}$$

where $y^{DQN}_{c_n}$ is the target value obtained by target network given as follows:

$$y^{DQN}_{c_n} = \varpi \max_{a_{(c_n,t+1)} \in \mathcal{A}} Q_{c_t}(s_{(t+1)}, a_{(c_n,t+1)}, \phi^-) + r_t \tag{3.24}$$

where ϕ^- indicates the target network weights. The pseudo-code of the training process of multi-agent DQN for secrecy energy efficient maximization is shown in Algorithm 2. Also, because the same values are used in the multi-agent DQN approach to choose and assess actions, the Q-value function may be over-optimistically calculated [261]. For instance, if action $a_{c_n,t}$ is more valuable in some states than action $a_{c_n,t+1}$, the agents would still take action $a_{c_n,t}$ in such states. Now, if action $a_{c_n,t+1}$ has become better choice for such memory experiences, it will be difficult for the neural network to comprehend that action $a_{c_n,t+1}$ is superior to action $a_{c_n,t}$ in such circumstances, since the DNN has been prepared to assign a significantly higher value to action $a_{c_n,t}$. Thus, the preceding issue is mitigated by using multi-agent double DQN (DDQN), which replaces the target $y^{DQN}_{c_n}$ with the following target $y^{DDQN}_{c_n}$ defined as

$$\begin{aligned} y^{DDQN}_{c_n} = \varpi Q_{c_t}(s_{t+1}, \max_{a_{(c_n,t+1)} \in \mathcal{A}} Q_{c_t}(s_{(t+1)}, a_{(c_n,t+1)}; \phi); \\ \phi^-) + r_{c_n,t} \end{aligned} \tag{3.25}$$

Also, the Q network in DDQN can be trained by minimizing the loss function defined as

$$L_{c_n}(\phi) = \mathbb{E}[(y^{DDQN}_{c_n} - Q_{c_n}(s_t, a_{c_n,t}; \phi))^2] \tag{3.26}$$

In DDQN, both the online Q network and the target Q network use the next state s_{t+1} to calculate the optimal Q value $Q_{c_n}(s_{(t+1)}, a_{(c_n,t+1)}; \phi)$. In addition, because the Q-value function explains how advantageous an action $a_{c_n,t}$ is made at a given state s_t, the duelling NN [262] is used for the estimation of advantage function

$A(s_t; a_{c_n,t}) = Q_{c_n}(s_t; a_{c_n,t}) - U(s)$ and a value function $U(s)$. Here, the avantage function $A(s_t; a_{c_n,t})$ shows the advantage of an action $a_{c_n,t}$ over various other actions.

Although multi-agent D3QN has the ability to perform well in policy learning with large and continuous state space, it may train ineffectively and produce divergence due to the sample correlations. Multi-agent D3QN samples each mini-batch $(s_t, a_{c_n,t}, r_t, s_{t+1})$ from the experience replay uniformly, which may have an unknown or bad influence on learning a better policy. This is because sampling every mini-batch transition equally may result in the wasteful use of meaningful mini-batch transitions. Therefore, to solve this issue and improve sampling efficiency, a prioritized experience replay-based multi-agent D3QN approach is proposed, where the priority of mini-batch transitions is decided by the values of temporal difference (TD) error. In prioritized experience replay-based multi-agent D3QN, a mini-batch transition with a higher absolute TD error has a higher priority because the action-value function is corrected more aggressively.

The learning model of the proposed scheme *SecBoost* is loaded for the implementation step when it has received appropriate training, as shown by the pseudo-code mentioned in Algorithm 3. The computational complexity of *SecBoost* is $\mathcal{O}(NT(j_0 j_k + \sum_{k=1}^{K-1} j_k j_{k+1}))$, where N is number of episodes, T is number of time steps, j_0 is the size of input layer of DNN, and j_k denotes the neurons in the k^{th} layer of DNN. In the implementation phase, each controller employs the trained learning model of *SecBoost* to output its chosen action $a_{c_n,t}$, by combining the observed states with the multi-agent D3QN parameter ϕ.

The explaination of proposed *SecBoost* is described below.

Step 1: Each agent initializes the replay memory of size V and online Q network $Q_{c_n}(s, a_{c_n}; \phi)$, and observes the current system state s_t including all the parameters such as channel vectors, interference to the agents, minimum secrecy rate requirements, previously predicted secrecy rate of PUs.

Step 2: Initialize the target Q network parameters and provide the state vector $s_t = \{h_k^M, h_{n,np}^P, h_{n,nE}^P, I_c, R_{np,sec}^P, R_{np,sec}^{P,min}\}$ as an input to D3QN to train the model.

Step 3: Evergy agent selects joint action a_{c_n} having maximum reward with probabiity $1 - \epsilon$ using ϵ-greedy technique.

Step 4: Compute the SINRs, of macrocell users, picocell users, and Eve, i.e., γ_k^M,γ_{np}^P,γ_{nE}^P, secrecy rate of picocell user and average SEE of picocells.

Step 5: Evergy agent gets a common reward r based on the choosen action a_{c_n}.

Step 6: Based on the obtained reward r, every agent observes transition of state from s_t to s_{t+1} and store a tuple of $(s_t, a_{c_n,t}, r_t, s_{t+1})$ in replay memory.

Step 7: Select a small batch transition with probability $p = \frac{|\delta|^\tau}{\sum_{j'} |\delta(j')|^\tau}$ using prioritized experience replay.

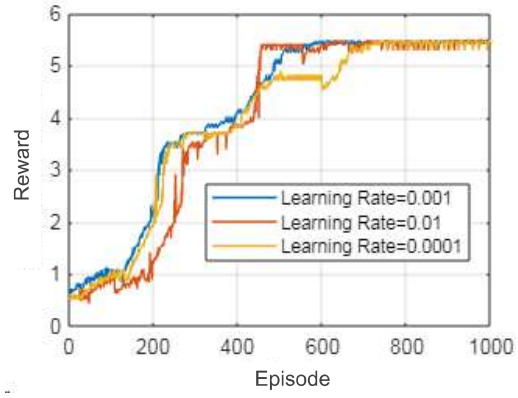

Figure 3.4: Reward performance analysis

Step 8: Calculate the TD error using $\delta = y_{c_n}^{DDQN} - Q_{c_n}(s_t, a_{c_n,t}; \phi)$ and update the mini batch transition priority using the value of TD.

Step 9: Every agent performs gradient descent step and updates the weights of target network.

Theorem 1: The learning process of the proposed scheme *Secboost* achieves its convergence, i.e., optimal $Q_{c_n}^*(s_t, a_t)$ of Markov decision process with the probability of 1, if the sequence of learning rate $\lambda_t \in (0, 1]$ confirms the following conditions: $\sum_{t=0}^{\infty} \lambda_t = \infty$ and $\sum_{t=0}^{\infty} \lambda_t^2 < \infty$.

Proof: The Q function and its associated policy have their optimal values with a probability of 1 if every action is carried out with an unbounded number of learning steps at every state of MDP [263].

Theorem 2: The proposed scheme *Secboost* given in Algorithm 3 achieves its optimal policy $\pi_{c_n}^*$ as

$$\pi_{c_n}^* = \arg\max_{w_{np}} h_{n,np}^P w_{np} \tag{3.27}$$

Proof: According to (3.13), the optimal SEE performance can be achieved if $h_{n,np}^P w_{np}$ is maximized or $h_{n,np}^E w_{np}$ is reduced to 0. However, $h_{n,np}^E w_{np} = 0$ is only possible if the geometric location of Eve is known [263]. Thus, optimal SEE performance of *SecBoost* is achieved if $h_{n,np}^P w_{np}$ is maximized.

3.4 Performance Evaluation

In this section, we evaluate the SEE performance of *SecBoost* and compare it with the following approaches:

- The classical MARL based SEE maximization scheme, in which Q learning is used to find an optimal policy.
- The MA-DQN based SEE maximization scheme, which uses DNN to find out

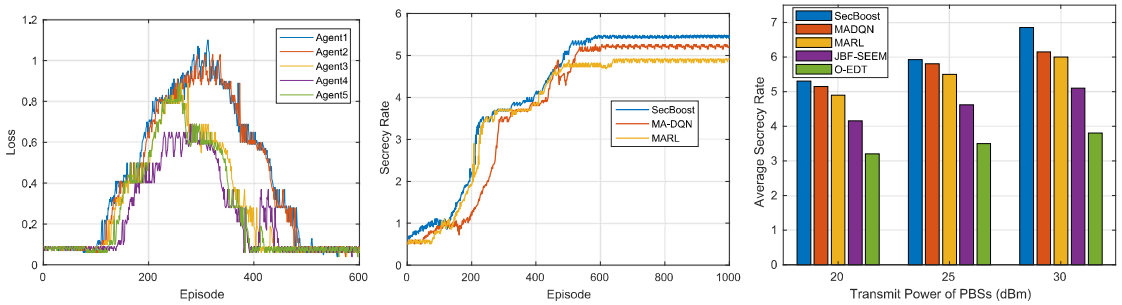

Figure 3.5: Comparative Simulation Analysis (a) Loss in Training (b) Secrecy rate performance (c) Average secrecy rate v/s Transmit power of PBSs.

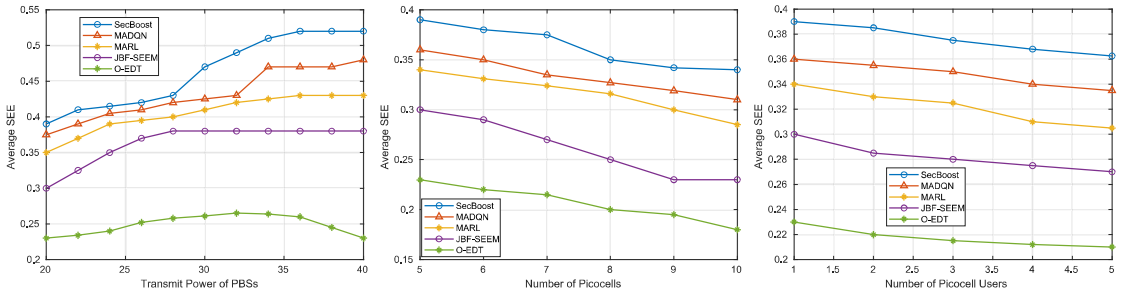

Figure 3.6: Comparative Simulation Analysis (a) Average SEE v/s Transmit power of PBSs (b) Average SEE v/s Number of Picocells (c) Average SEE v/s Number of Picocell Users

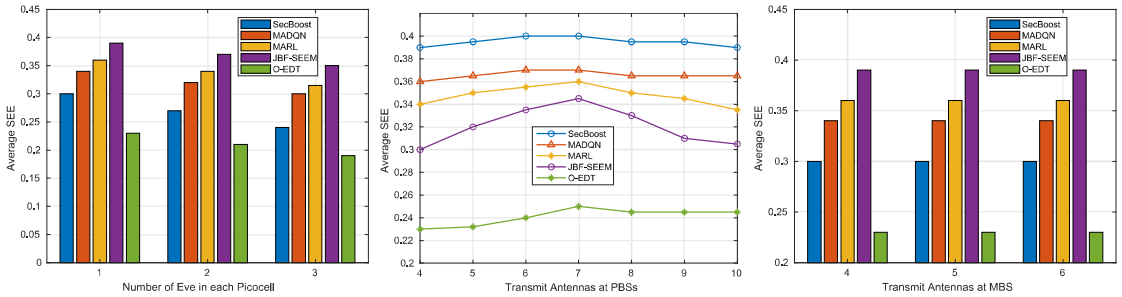

Figure 3.7: Comparative Simulation Analysis (a) Average SEE v/s Total Eve in each Picocell (b) Average SEE v/s number of PBS's transmit antennas (c) Average SEE v/s MBS's transmit antennas

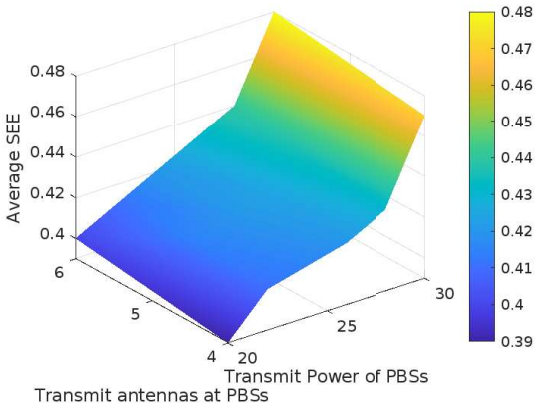

Figure 3.8: SEE region v/s Transmit power PBSs v/s Transmit Anteenas at PBSs

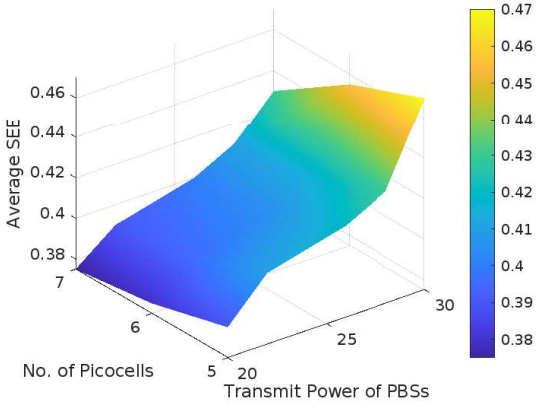

Figure 3.9: SEE region v/s Transmit power PBSs v/s No. of Picocells

Table 3.3: Values of simulation parameters

Parameters	Values
Radius of macro cell	500m
Radius of each pico cell	250m
Number of PBS	5
No. of RBs	10
No. of MUs	2
No. of PUs in each pico cell	1∼8
MBS transmit power	40dBm
Noise power spectrum density	-174dBm
Learning Rate	0.001
Discount Factor	0.9
Starting exploration	1
Exploration in final	0.01
Total exploratory steps	1000
Replay buffer size	1000
Layers of hidden layers in DNN	3
Neurons in hidden layers	64,32,32
Mini-batch size	32
Weights of reward	1,1
Power discretization level	10
Weights updation interval	10
Used activation function	ReLu
Optimizer used	Adam

the optimal policy of power control, channel allocation, and beamforming corresponding to the optimal Q-value.

• Joint beamforming based secrecy energy efficiency maximization (JBF-SEEM) [248] scheme which jointly optimizes the beamforming and artificial noise vectors at MBS and small base station (SBS). Using Dinkelbach's method, firstly, the non-convex optimization problem is translated into a subtractive optimization problem and further approximated as a convex problem using the difference of convex functions (D.C.) technique.

• One-time pad based encrypted data transmission (O-EDT) scheme. One-time pad is a post-quantum cryptography technique which uses a pre-shared key (single use) and have value larger than the message to be sent. Channel state information, i.e., the channel vectors, are used for the key generation. Since the key generation depends on the common randomness between PBSs and PUs, we have evaluated both downlink and uplink channels. We have allocated $\mathcal{Y}$ time slots for the estimation of channels and the remaining $\mathcal{Z} - \mathcal{Y}$ time slots for key generation and data transmission. Based on [264], the key generation rate (KGR) for femtocells is given as follows.

$$R_{kg} = H(h^P_{n,np}|h^P_{n,nE}, h^P_{np,nE}) - H(h^P_{n,np}|h^P_{np,n}, h^P_{n,nE}, h^P_{np,nE}) \tag{3.28}$$

where $h^P_{np,nE}$ is the channel vectors from p–th PU of n–th picocell to Eve of n–th picocell, and $h^P_{np,n}$ is the channel vector from p–th PU to n–th PBS. Moreover, based

on [265], the secrecy rate of O-EDT is given as follows.

$$R^{P}_{np,sec(O-EDT)} = \log_2(1+\gamma^{P}_{np})(\mathcal{Z}-\mathcal{Y})/\mathcal{Z} \tag{3.29}$$

3.4.1 Numerical Settings

This simulation considers a two-tier downlink HetNet consisting of a central MBS and uniformly distributed PBSs, MUs, and PUs. We assume the radius of the MBS and PBS is 500m and 250m, respectively. The MBS and PBS antenna configurations are $N_M = 4$ and $N_P = 4$, respectively, in all simulations. The number of MUs is $\mathcal{K} = 2$, and the number of PUs for each PBS is $\mathcal{P} = 1$. Also, according to [266], the power amplification coefficient is assumed to be $\zeta = 2.6$, power consumed by each antenna of PBS is $P_a = 30$ dBm, and each PBS's basic power consumption is $P_b = 40$ dBm. The minimum average SEE requirement of picocells is set to be 0.05 (bit/Joule/Hz). The parameters used in the simulation are shown in Table 3.3.

Remark: It is to be noted that the proposed scheme does not require prior knowledge of instantaneous global CSI of HetNet. It is learnt from the continuous learning process of DRL, which satisfies our rationale behind the use of DRL in the absence of instantaneous CSI of HetNets.

3.4.2 SecBoost Performance Analysis

Since the convergence of the techniques influences system performance, we evaluate the convergence of *SecBoost* in Fig. 3.4. The selection of network parameters determines the efficacy and speed of the learning convergence. Here, we have used the learning rate as the network parameter to show the relevance of network parameter selection. Learning rate is a user-configurable hyper-parameter that determines how quickly the model adapts to the task. The numerical configuration for this analysis is $\mathcal{P} = 5$, $N_M = N_P = 4$, $P_M = 40$ dBm and $P_p = 20$ dBm. Fig. 3.4 depicts the system reward versus training episodes at various learning rates i.e., $\lambda = \{0.01, 0.001, 0.0001\}$. It can be observed that a high learning rate (λ =0.01) causes rapid convergence and substantial fluctuations in reward value, which can lead to unstable training or even divergence. Further, if we use a too-low learning rate (λ =0.0001), it takes longer to reach convergence. In terms of secrecy rate, the modest learning rate value (λ =0.001) yields the richest and most consonant reward. As a result, we pick a suitable learning rate (λ =0.001) that is neither too large nor too small for our simulations.

The loss function readings of the proposed *SecBoost* scheme during training are displayed in Fig. 3.5(a). All five agents' loss function readings reach their peak after roughly 300 episodes and then decline as the agents exploit improved actions. As the reward value converges, the loss function readings of all agents gradually

decrease until it reaches the minimum value, which validates the accurate Q value prediction of *SecBoost*. In *SecBoost*, agents are PBSs that try to maximize their rewards by selecting the optimal policy. We have designed the reward function in a manner that all agents receive the same reward. Thus, allowing agents to choose actions that raise the cumulative value of the reward.

3.4.3 Secrecy Rate Analysis

Fig. 3.5(b) shows the convergence of the secrecy rate with respect to the number of episodes. The numerical settings for this analysis is $\lambda = 0.001$, $\mathcal{P} = 5$, $N_M = N_P = 4$, $P_M = 40$ dBm and $P_p = 20$ dBm. As the number of episodes increases, the secrecy rate of *SecBoost*, MA-DQN, MARL reaches its convergence and achieves their respective maximum secrecy rate floors after around 600 episodes. Also, the proposed *SecBoost* scheme achieves a better secrecy rate than MA-DQN and MARL schemes. It is to be noted that, even with a low transmit power of PBSs, i.e., $P_p = 20$ dBm, the proposed scheme achieves a better secrecy rate. This also satisfies the rationale behind the use of prioritized experience replay with D3QN in the proposed SecBoost scheme, which outperforms MARL and MA-DQN schemes in terms of achievable secrecy rate performance.

3.4.4 Impact of Transmit Power on Secrecy Rate and SEE

Fig. 3.5(c) depicts the average secrecy rate of picocell users versus the transmit power of PBSs. The numerical configuration for this analysis is $\mathcal{P} = 5$, $N_M = N_P = 4$. It can be seen that the average secrecy rate of picocells increases with an increase in the transmit power level of PBSs. The reason is that when P_p increases, the received SINR of PUs also increases as compared to Eve. Also, it can be observed from Fig. 3.5(c) that the proposed *SecBoost* scheme outperforms MA-DQN, MARL, JBF-SEEM, and O-EDT schemes in terms of the achievable secrecy rate of picocells.

Remark: It is evident from Fig. 3.5(c) that the JBF-SEEM scheme utilizes high transmit power to broadcast artificial noise to improve the secrecy level of 5G HetNet. On the other hand, the proposed scheme jointly optimizes the beamforming and power allocation vectors efficiently, thus achieving better secrecy levels even at the low transmit power values. Thus, the proposed scheme address the issue of high transmit power required to broadcast artificial noise, discussed in Subsection B of Section I.

Fig. 3.6(a) shows the average SEE of picocells versus the transmit power level of every PBS i.e., $P_p = \{20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40\}$. It depicts the comparative analysis of the SEE performance of *SecBoost*, MA-DQN, MARL, JBF-SEEM, and O-EDT schemes with $P_M = 40$ dBm, $\mathcal{P} = 5$, $N_M = N_P = 4$. It can be observed

that the average SEE of the *SecBoost*, MA-DQN, MARL, JBF-SEEM, and O-EDT schemes increases as the transmit power of each PBS is increased, implying that all schemes achieve the maximum average SEE with the given transmit power level of each PBS. This is because as the value of P_p rises, the received SINR and average secrecy rate at PUs rise, resulting in improved SEE performance. Further, with a continuous increase in the transmit power of PBSs, the average SEE of all the schemes converges to their respective average SEE levels. This is because all the schemes will stop consuming more transmit power to improve the SEE of all picocells after achieving the maximum average SEE level. Moreover, we observe that the proposed scheme *SecBoost* achieves 14.7%, 8.33%, 30%, and 69% better average SEE than MARL, MA-DQN, JBF-SEEM, and O-EDT schemes, respectively, thus showing its superiority to achieve maximum SEE. Thus, it is evident from Fig. 3.5(c) and 3.6(a) that applying the classical optimization technique degrades the secrecy and SEE performance of the 5G HetNet, since the secrecy and SEE performance of JBS-SEEM is less than other RL based approaches, which proves the rationale behind the use of RL. This also addresses the limitation of applying classical optimization problems in dynamic environments, mentioned in Subsection B of section I.

3.4.5 Impact of Number of Picocells

Fig. 3.6(b) demonstrates the average SEE of picocells versus the number of picocells $\mathcal{P}$ for *SecBoost*, MA-DQN, MARL, JBF-SEEM, and O-EDT with $P_P = 20$ dBm, and $N_M = N_P = 4$. It can be observed that the proposed scheme, *SecBoost* outperforms MA-DQN, MARL, JBF-SEEM, and O-EDT schemes in terms of average SEE of picocells in two-tier downlink HetNet. As shown in Fig. 3.6(b), with an increase in picocells $\mathcal{P}$, the average SEE of all schemes decreases. This is because, as the picocells increases, the average secrecy rate of picocells also increases with increase in the total power consumption of picocells. However, the pace of increase in average secrecy rate is slower than the rate of increase in total power consumption of picocells, resulting in a decreased average SEE performance.

3.4.6 Impact of Number of Picocell Users

Fig. 3.6(c) shows the comparative analysis of the average SEE versus the number of legitimate picocell users for *SecBoost*, MA-DQN, MARL, JBF-SEEM, and O-EDT schemes. The numerical settings for this analysis is $\mathcal{P} = 5$, $N_M = N_M = 4$, $P_P = 20$ dBm. As the number of PUs increases, the average secrecy rate of picocells decreases due to increased co-channel interference with more PUs connected to the same PBS. However, as the transmit power level of PBS remains the same, the average SEE of

picocells also decreases with increasing picocell users. Also, it can be noticed from Fig. 3.6(c), that the proposed scheme *SecBoost* performs better than MA-DQN, MARL, JBF-SEEM, and O-EDT schemes in terms of SEE.

3.4.7 Impact of Number of Eavesdroppers

Fig. 3.7(a) shows the effect of multiple eavesdropping nodes on the achievable average SEE of picocells in two-tier downlink HetNet. The numerical configuration for this evaluation is $\mathcal{P} = 5$, $N_M = N_P = 4$, $P_P = 20$ dBm. It has been observed that with an increase in the number of eavesdropping nodes, the average SEE of picocells decreases. It is because more eavesdroppers in each picocell cause a significant decrease in the achievable data rate of PUs, which reduces the achievable secrecy rate of picocells. Thus, the average SEE of picocells decreases in all the schemes.

3.4.8 Impact of Number of PBS's Transmit Antennas

Fig. 3.7(b) depicts the comparative analysis of the average SEE performance of *SecBoost*, MA-DQN, MARL, JBF-SEEM, and O-EDT schemes for the varying number of transmit antennas of PBSs. The numerical configuration for this analysis is $\mathcal{P} = 5$, $N_M = N_P = 4$, $P_P = 20$ dBm. It is to be noted from Fig. 3.7 (b) that the average SEE of all the schemes firstly increases and then decreases with a continuous increase in the number of transmit antennas of PBSs. This is because increasing the quantity of PBSs transmit antennas enhances the average secrecy rate of picocells, it also increases the system power consumption. Further, the *SecBoost* outperforms MA-DQN, MARL, and JBF-SEEM schemes in terms of the average SEE performance of picocells.

Remark: It is evident from Fig. 3.5(c), 3.6(a), and 3.7(b) that the proposed scheme consumes less transmit power to achieve better secrecy and SEE levels, which leads to reduction in the overall power expenses of the system. Thus, the proposed scheme addresses one of the limitations, i.e., high power expenses, mentioned in subsection B of section I.

3.4.9 Impact of Number of MBS's Transmit Antennas

In Fig. 3.7(c), we evaluate the SEE of picocells versus the number of MBS's transmit antennas for *SecBoost*, MA-DQN, MARL, JBF-SEEM, and O-EDT schemes with $\mathcal{P} = 5$, $N_P = 4$, $P_P = 20$ dBm. It can be observed that increasing the number of transmit antennas of MBS has nil impact on the SEE of picocells. It is because there is no effect of interference from MBS to Eve on the achievable secrecy rate of PUs. Also, the proposed *SecBoost* scheme performs better than MA-DQN, MARL, JBF-SEEM, and O-EDT schemes in terms of achieving SEE of picocells.

3.4.10 SEE Region of Picocells

To understand the behavior of *SecBoost* more effectively, we exhibit the average SEE of pico cell users versus the transmit power and transmit antennas of PBSs in a three-dimensional (3D) figure in Fig. 3.8, where we can see the variation pattern of average SEE more effectively with different transmit power levels and transmit antennas of PBSs. It is evident from Fig. 3.9 that the average SEE of picocells increases with an increase in the level of transmit power and transmit antennas of PBSs. Fig. 3.9 shows the SEE region of picocell users acquired via the proposed *SecBoost* scheme with variation in the number of picocells and transmit power of PBSs.

3.5 Summary

In this chapter, we maximize the secrecy energy efficiency of mmWave enabled 5G HetNets while maintaining the other constrain requirments of UEs. To achieve this goal, a joint optimization problem of beamforming, power and channel allocation is formulated. To solve the non-convex optimization problem, we have used Markov decision process. We translate the problem into multi-agent RL problem. Further, we have proposed a mult-agent DRL scheme to enhance the SEE of 5G enabled HetNet. The simulation results demonstrate the effectiveness of the proposed scheme in achieving better SEE performance than other state-of-the-art schemes.

Chapter 4

Mitigating Jamming Attack in 5G HetNets: A Federated DRL Approach

In this chapter, we propose a federated DRL based joint optimization of beamforming vectors and power allocation at FBSs with an aim to maximize the achievable rate at FUs in the presence of jammers. Table 4.1 shows the comparison of the proposed federated DRL scheme with various pre-existing anti-jamming techniques.

4.1 Major Contributions

The major contributions of the chapter are as follows.

- The achievable rate for multiple mmWave femtocells having FUs and jammers is investigated in this chapter. Moreover, we have formulated a joint optimization problem of beamforming and power allocation at FBSs to maximize the achievable rate at FUs.
- A multi-agent reinforcement learning (MARL) problem is formulated using the Markov decision process (MDP) to obtain an optimal strategy for joint beamforming and power allocation optimization.
- A federated DRL scheme is proposed to maximize the achievable rate of mmWave-enabled FUs in a two-tier downlink 5G HetNet.
- Using extensive simulations, we have compared the achievable rate of proposed scheme with double deep Q network (DDQN) and deep Q network (DQN) schemes.

Table 4.1: Comparision of the proposed scheme with pre-existing anti-jamming techniques

Reference	Description	Model	Power Control	Beamforming	Energy-Efficient	Use of mmWave	Use of Sub-6 GHz
[190]	A bimatrix game framework is proposed to model the interactions between jammer and transmitter against jamming attack	FH based wireless communication consisting of a transmitter, receiver and a jammer	×	×	×	×	×
[191]	A Bayesian game is formulated between a jammer and base station (BS) to minimize the transmission cost while ensuring QoS requirements	Downlink massive MIMO system	×	✓	×	×	×
[175]	A resource allocation strategy based on transmit power control was proposed against jamming	IoT network with jammer	✓	×	×	×	×
[150]	A cooperative relay beamforming based scheme was proposed to mitigate the jamming attack in wireless vehicular networks	Vehicular wireless network	×	✓	×	×	×
[176]	A jamming-resistant receiver technique was developed to enhance the jamming resistance of massive	Massive MIMO uplink system	✓	×	✓	×	×

Table 4.1: Comparision of the proposed scheme with pre-existing anti-jamming techniques

Reference	Description	Model	Power Control	Beamforming	Energy-Efficient	Use of mmWave	Use of Sub-6 GHz
[177]	To prevent jamming attacks in wireless HetNet, a Stackelberg based game is modeled for power domains	HetNet with malicious jammer	✓	×	×	×	×
[151]	A beam domain based anti-jamming scheme was proposed for the uplink massive MIMO enabled system	Uplink massive MIMO system	×	✓	×	×	×
[267]	RL and IRS based hill-climbing (fast-policy) learning scheme was proposed against jamming attacks	IRS assisted communication system	✓	✓	✓	×	×
[268]	Unmanned Aerial Vehicles (UAVs) were used to relay the messages of onboard units to improve the anti-jamming performance	UAV assisted VANETs	×	×	✓	×	×
[269]	A RL based power control strategy was introduced to enhance the average SINR of UEs	mmWave massive MIMO system	×	✓	✓	✓	×
Proposed scheme	Federated DRL based joint optimization of the beamforming and power allocation against jamming for the mmWave femtocells	Downlink two-tier HetNet	✓	✓	✓	✓	✓

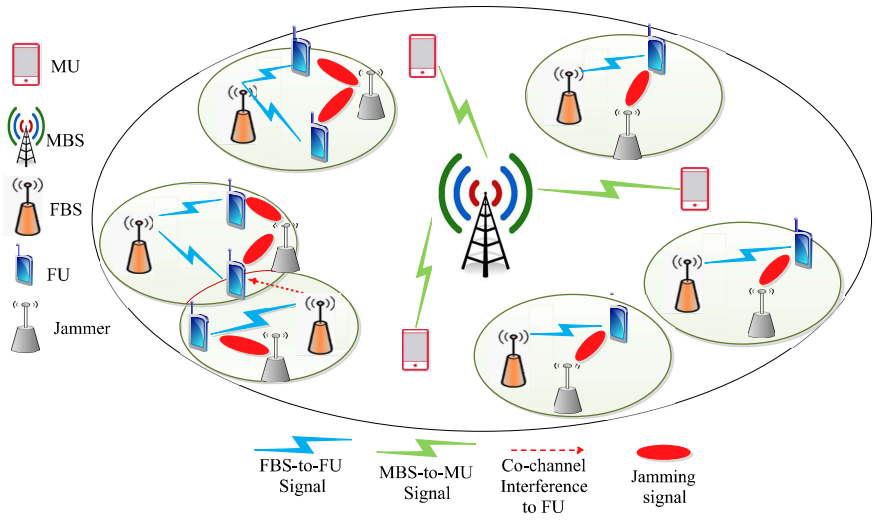

Figure 4.1: System Architecture

4.2 System Model

We have considered a two-tier heterogeneous network consisting of a macrocell and a set of $\mathcal{Q} = \{1, 2, \ldots, Q\}(\forall q \in \mathcal{Q})$ femtocells, as shown in Fig. 4.1. The macrocell comprises of a N_m-antenna macro base station (MBS), and the set of $\mathcal{M} = \{1, 2, \ldots, M\}(\forall m \in \mathcal{M})$ single antenna macrocell users (MUs). Each femtocell consists of a N_f-antenna femto base station (FBS), and a set of single antenna $\mathcal{F} = \{1, 2, \ldots, F\}(\forall f \in \mathcal{F})$ femto users (FUs). The FBSs communicate with the FUs using mmWave. Transmit power for all the FBSs is considered to be equal to limit the effect of co-channel interference (CCI) and is represented by P_f. The transmit power of MBS is represented by P_m. Also, a reactive jammer having N_J-antennas is placed in each femtocell, which tries to jam the downlink signals from FBS to FUs. The reactive jammer has the capability to listen to the wireless channel while remaining in an idle state. Further, it immediately transmits the jamming signals to jam the active transmission when senses any channel activity, irrespective of the frequency/channel in use. The set of jammers is defined as $\mathcal{J} = \{1, 2, \ldots, J\}(\forall j \in \mathcal{J})$. The transmit power of jammers is represented by P_j. The major symbols used in the chapter are shown in Table 4.2.

4.2.1 mmWave Tier

We have used mmWave channel model for femtocells in this work. The frequency range spectrum of mmWave ranges from 24 GHz to 100 GHz [249]. Similar to [250,251], the mmWave channel vector from f–th FU of q–th femtocell is represented as follows.

$$h_{q,qf} = \sqrt{X_{q,qf}} a_s(\phi_{q,qf}) a_u(\beta_{q,qf}) \tag{4.1}$$

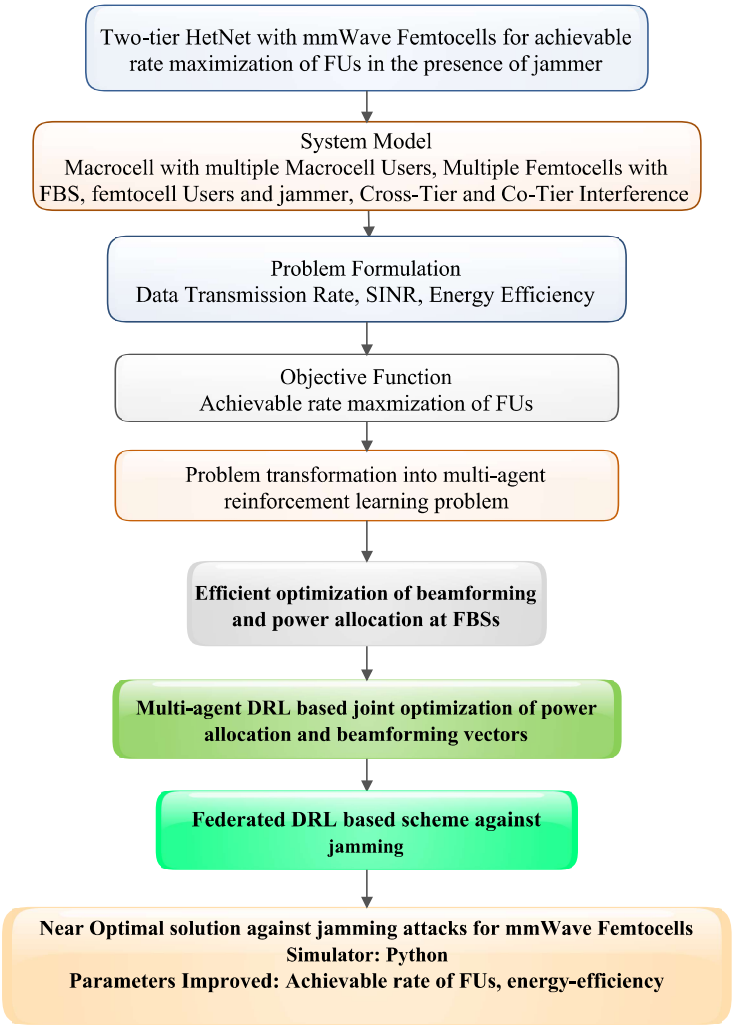

Figure 4.2: Sequence of steps involved in the proposed work

Table 4.2: List of major symbols used

Symbol	Description
$\mathcal{Q}$	Total number of femtocells
$\mathcal{M}$	Number of macrocell users
$\mathcal{F}$	Number of femto users in each femtocell
N_m	Number of the transmit antennas of MBS
N_f	Number of transmit antennas of each FBS
$\mathcal{J}$	Number of jammers
N_J	Number of antennas at jammer
$h_{q,qf}$	Channel vector from FBS_q to f–th FU of q–th femtocell
y_m	Received signal at m–th MU
w_m	Beamforming vector from MBS to m–th MU
w_{qf}	Beamforming vector from q–th FBS to f–th FU of q–th femtocell
s_m	The information signal intended for m–th MU
s_{qf}	The information signal intended for f–th FU of q–th femtocell
n_{qf}	AWGN at m–th MU
n_{qf}	Additive white Gaussian Noise f–th FU of q–th femtocell
y_{qf}	Received signal at f–th FU of q–th femtocell
ψ_{qf}	SINR of f–th FU of q–th femtocell
ψ_m	SINR of m–th MU
$h_{j,qf}$	channel vectors from j–th jammer to f–th FU of q–th femtocell and from MBS to jammer
R_{qf}	Achievable rate at f–th FU of q–th femtocell
R_{qf}^{min}	Minimum achievable rate requirement at f–th FU of q–th femtocell
λ	Learning Rate
β	Discount Factor
N	Total number of episodes

where $a_s(\phi_{q,qf})$ and $a_u(\beta_{q,qf})$ represents the steering vectors, $\phi_{q,qf}$ and $\beta_{q,qf}$ denotes the angle of arrival and the angle of departure at FBS_q and FU_{qf}, respectively. $X_{q,qf}$ is the path loss between FBS_q and FU_{qf} which is formulated as follows.

$$X_{q,qf} = \begin{cases} c_l(d_{q,qf})^{-\alpha_l}, & \text{for line-of-sight (LoS)} \\ c_n(d_{q,qf})^{-\alpha_n}, & \text{otherwise} \end{cases} \tag{4.2}$$

where c_l and c_n denotes the path loss of LoS and non-LoS links, $d_{q,qf}$ is the distance between q–th FBS to f–th FU, α_l and α_n are path loss exponents for the LoS and non LoS wireless links.

4.2.2 Sub-6 GHz Tier

In this chapter, a sub-6 GHz channel model is used for macrocell. The frequency emitted by a cellular base station in Sub-6 GHz bands is less than 6 GHz [249]. The channel vector from MBS to m–th MU is given as follows.

$$h_m = \sqrt{x_m u_m} \tag{4.3}$$

where u_m represents the fading vector, and x_m is the path loss from MBS to m–th MU defined as follows.

$$x_m = \left(\frac{c}{4\pi f_c}\right)^2 (D_m)^{-\alpha_p} \tag{4.4}$$

where c denotes the speed of the light in vacuum (m/s), f_c is the carrier frequency, D_m is distance from MBS to m–th MU, and α_p shows the path loss exponent.

4.2.3 Downlink Signal Model

The signal received at m–th MU of macrocell is defined as follows.

$$y_m = h_m w_m s_m + \sum_{i=1, i\neq m}^{M} h_i w_i s_i + n_m \tag{4.5}$$

where h_m denotes the channel vector from MBS to the m–th MU. w_m and s_m denotes beamforming vector and information signal from MBS to the m–th MU, where $w_m \in \mathcal{C}^{N_m \times 1}$ is the continuous linear precoding [270]. $n_m \sim \mathcal{CN}(0, \delta_m^2)$ shows the additive white Gaussian noise (AWGN) of m–th MU. Also, the beamforming vectors for MUs satisfy the power constraint of $\sum_{m=1}^{M} w_m = P_m$.

Also, the signal received at f–th FU of q–th femtocell is given as follows.

$$\begin{aligned} y_{qf} = {} & h_{q,qf} w_{qf} s_{qf} + \sum_{t=1, t\neq f}^{F} h_{q,qf} w_{qt} s_{qt} + h_{j,qf} z_{j,qf} \\ & + \sum_{l=1, l\neq q}^{Q} \sum_{t=1}^{F} h_{l,qf} w_{lt} s_{lt} + \sum_{u=1, u\neq j}^{J} h_{j,qf} z_{j,qu} + n_{qf} \end{aligned} \tag{4.6}$$

where $h_{q,qf}$ and $h_{j,qf}$ represents the channel vector from q–th FBS to FU_{qf}, and channel vector from j–th jammer to FU_{qf}, respectively. s_{qf} and w_{qf} denote the information signal and beamforming vector for f–th FU from q–th FBS, where $w_{qf} \in \mathcal{C}^{N_f \times 1}$ is the continuous linear precoding. Also, the beamforming vectors for FUs satisfy the power constraint of $\sum_{f=1}^{F} w_{qf} = P_f$. $n_{qf} \sim \mathcal{CN}(0, \delta_{qf}^2)$ is the AWGN at the FU_{qf}. Moreover, $z_{j,qf} \in \mathcal{C}^{N_j \times 1}$ denotes the jamming vector from j–th jammer to FU_{qf} and $\sum_{f=1}^{F} z_{j,qf} = P_j$.

4.2.4 Signal-to-Interference-Noise Ratio (SINR) Calculation

The SINR of the m–th MU of macrocell is defined as follows.

$$\psi_m = \frac{|h_m w_m|^2}{\sum_{i=1,i\neq m}^{M} |h_i w_i|^2 + \delta_m^2} \tag{4.7}$$

Also, SINR at f–th FU of q–th femtocell is defined as follows.

$$\psi_{qf} = \frac{|h_{q,qf} w_{qf}|^2}{\sum_{t=1,t\neq f}^{F} |h_{q,qf} w_{qt}|^2 + I_y + \delta_{qf}^2} \tag{4.8}$$

$$\text{where} \quad I_y = |h_{j,qf} z_{j,qf}|^2 + \sum_{l=1,l\neq q}^{Q} \sum_{t=1}^{F} |h_{l,qf} w_{lt}|^2 + \sum_{u=1,u\neq j}^{J} |h_{j,qf} z_{j,qu}|^2 \tag{4.9}$$

4.2.5 Downlink Data Transmission Rate Calculation

The downlink data transmission rate from MBS to m–th MU is represented as follows.

$$R_m = \log_2(1+\psi_m) = \log_2\left(1 + \frac{|h_m w_m|^2}{\sum_{i=1,i\neq m}^{M} |h_i w_i|^2 + \delta_m^2}\right) \tag{4.10}$$

Also, the downlink data transmission rate from q–th FBS to f–th FU is expressed as follows.

$$R_{qf} = \log_2\left(1 + \frac{|h_{q,qf} w_{qf}|^2}{\sum_{t=1,t\neq f}^{F} |h_{q,qf} w_{qt}|^2 + I_y + \delta_{qf}^2}\right) \tag{4.11}$$

Also, the total sum rate of femtocells can be defined as

$$R_{\mathcal{Q}}^{Total} = \sum_{q=1}^{Q} \sum_{f=1}^{F} R_{qf} \tag{4.12}$$

4.2.6 Total Power Consumption at Femtocells

The total power consumption of femtocells is given as follows.

$$P_{\mathcal{Q}}^{Total} = \zeta \left[\sum_{q=1}^{Q} \sum_{f=1}^{F} \parallel w_{qf} \parallel^2 \right] + \mathcal{Q}(N_f P_a + P_b) \tag{4.13}$$

where ζ denotes the power amplification coefficient, P_a is the power consumption by antenna of FBS, and P_b represents the basic power consumption of FBS.

4.2.7 Energy Efficiency Calculation

The energy efficiency of femtocells is mathematically defined as follows.

$$E_{\mathcal{Q}} = \frac{R_{\mathcal{Q}}^{Total}}{P_{\mathcal{Q}}^{Total}} \tag{4.14}$$

4.2.8 Problem Formulation

The objective of problem defined in this chapter is to maximize the achievable rate of FUs by jointly optimizing the beamforming vector and power allocation at FBSs. The objective function is mathematically formulated as follows.

$$\mathcal{P}.\mathcal{F}. \quad : \quad \max_{q\in\mathcal{Q}, f\in\mathcal{F}} R_{qf} \tag{4.15}$$

$$\begin{aligned} s.t. \quad C1 \quad &: \quad \sum_{m=1}^{M} \|w_m\|^2 \leq P_m \\ C2 \quad &: \quad \sum_{f=1}^{F} \|w_{qf}\|^2 \leq P_f \\ C3 \quad &: \quad R_{qf} \geq R_{qf}^{\min} \end{aligned} \tag{4.16}$$

where $R_{qf}^{\min}$ is the minimum achievable rate requirement at f–th FU of q–th femtocell, P_m is MBSs' transmit power (transmission cost), and P_f is FBSs' transmit power (transmission cost). The explanation of various constraints defined above is given as follows. $C1$ states that the sum of beamforming vectors for MUs does not exceed the maximum transmission power of MBS, i.e., P_m. $C2$ ensures that the sum of beamforming vectors for FUs from FBS does not exceed the maximum transmission power of FBS, i.e., P_f. $C3$ implies that the achievable rate at FUs should be greater than minimum achievable rate requirement at FUs.

The mathematical problem formulated in (4.15) is a non-convex optimization problem due to the existence of interference terms in ψ_m and ψ_{qf}. Also, the objective function is non-concave in nature over the beamforming vectors w_m and w_{qf}. Moreover, the maximum transmit power levels P_m and P_f and the beamforming vectors w_m and w_{qf} are complexly coupled inside the main objective function, which makes the joint optimization problem difficult to be solved. In a realistic HetNet scenario, the quality of wireless quality, and capabilities of MUs and FUs change dynamically. The model-free RL is a dynamic programming tool which can be used to learn the optimum policy for solving decision-making problems in dynamic environments [252]. Thus, we used the model-free RL approach to solve the given optimization problem.

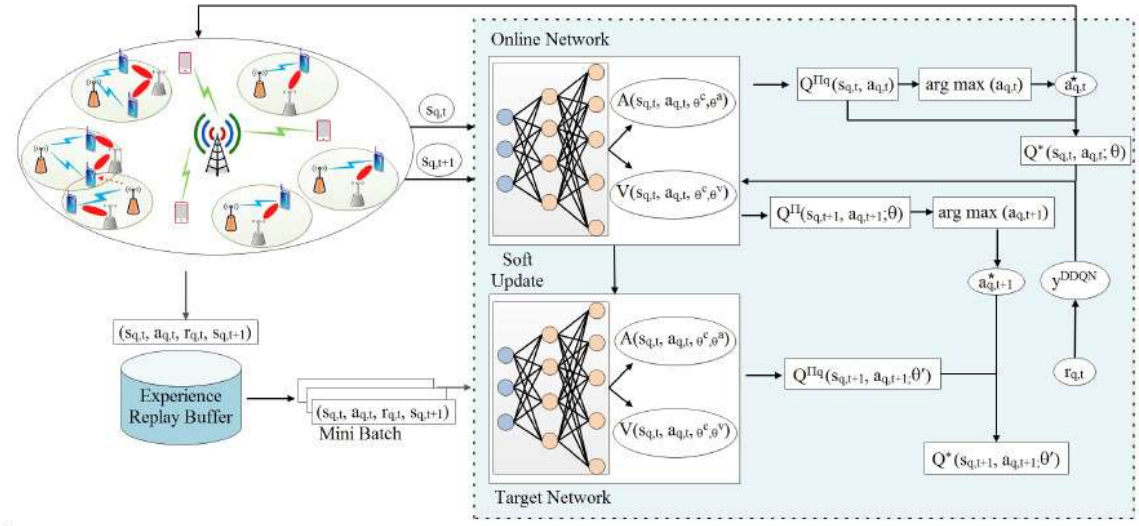

Figure 4.3: Architecture of multi-agent DRL based anti-jamming transmission

4.3 Proposed Scheme

This section firstly transforms the optimization problem formulated in (4.15) as a multi-agent RL problem using MDP. Then, we propose a federated DRL based anti-jamming scheme to maximize the achievable rate of mmWave-enabled FUs.

4.3.1 Markov Decision Process (MDP)

We define the MDP as $(\mathcal{S}, \mathcal{A}, \mathcal{T}_{ss'}, \mathcal{R}, \beta)$, where $\mathcal{S}$ denotes the set of state space, $\mathcal{A}$ represents the set of action space, $\mathcal{T}_{ss'}$ depicts the transition probability from state s_t to state s_{t+1}, $\mathcal{R}$ is reward function, and $\beta \in (0, 1\}$ is discount factor. We assume that the controller at each FBS acts as an agent. The complete detail of the MDP is defined as follows.

4.3.1.1 Agent(s)

Controller at FBSs.

4.3.1.2 State Space

The present state $s \in \mathcal{S}$ contains the estimated channel information and the achievable rate of previous time slot. For each controller, the observed state space is defined as $s = \{h_m, h_{q,qf}, h_{j,qf}, I_c, R_{qf}, P_j\}$, where $s \in \mathcal{S}$, h_m is the channel vector from MBS to MU_m, $h_{q,qf}$ is the channel vector from FBS to FUs, $h_{j,qf}$ shows the channel vector from jammer to FUs, and I_c is the interference to the agents (controllers).

4.3.1.3 Action Space

As per the observed state space, each controller selects the beamforming vector $\{w_{q,f}\}_{q\in\mathcal{Q},f\in\mathcal{F}}$ at FBS and transmit power $\{P_{q,f}\}_{q\in\mathcal{Q},f\in\mathcal{F}}$ for FU. The action set for q–th controller is defined as follows.

$$a = \{\{w_{q,f}\}_{q\in\mathcal{Q},f\in\mathcal{F}}, \{P_{q,f}\}_{q\in\mathcal{Q},f\in\mathcal{F}}\} \tag{4.17}$$

where $a \in \mathcal{A}$, and $\{P_{q,f}\}_{q\in\mathcal{Q},f\in\mathcal{F}}$ is the set of discrete power levels between 0 and P_f.

4.3.1.4 Transition Probability

$\mathcal{T}_{ss'}$ denotes the probability of transition. The probability of transition from existing state $s_t \in \mathcal{S}$ to the next state $s_{t+1} \in \mathcal{S}$ after executing an action a_t at time t is given as follows.

$$\mathcal{T}_{ss'} = \mathcal{T}(s_{t+1} = s'|s_t = s, a_t) \tag{4.18}$$

The state transition probability reflects the system's dynamics in such a way that the probability of next state s_{t+1} depends on the action a_t and existing state s_t only. Further, the transition probability must sastisfy the following condition.

$$\sum_{s\in\mathcal{S}}\sum_{r\in\mathcal{R}}\mathcal{T}(s', r|s, a) = 1 \tag{4.19}$$

4.3.1.5 Reward Function

Learning process of RL is based on an immediate value of reward function. As a result, each agent selects an action to maximize its reward based on its interactions with the environment.

In this study, the controller at q–th FBS takes action $a_{q,t}$ upon state $s_{q,t}$ at time t to obtain the feedback reward. Here, the reward function reflects the correlation between the optimization objective and the goal is the maximization of achievable rate at FUs. Therefore, reward function at time slot t is represented as follows.

$$r_{q,t} = \sum_{f=1}^{F} R_{qf,t} \tag{4.20}$$

where $r_{q,t} \in \mathcal{R}$ is the reward obtained by q–th FBS's controller.

In MDP, an agent observes the state s_t at time slot t and responds with an action a_t, which is chosen based on the policy π. So, the existing state of environment transits to a new state s_{t+1} with the reward r_t.

Here, the controller at each FBS chooses the policy π to optimize the accumulative reward function by selecting an action $a_{q,t} : \pi(s_{q,t}) : \mathcal{S}\ \mathcal{A}$. The accumulative

reward function at policy π and state $s_{q,t}$ is given as follows.

$$P^{\pi}_{r_{q,t}}(s_{q,t}) = \sum_{t=1}^{\infty} \beta_t r_{q,t}(s_{q,t}, a_{q,t} | s_{q,t-1} = s_{q,t}, \pi) \tag{4.21}$$

Also, the best discounted accumulative reward function is given as follows.

$$P^{*}_{r_{q,t}}(s_{q,t}) = \max_{\pi} P^{\pi}_{r_{q,t}}(s_{q,t}) \tag{4.22}$$

The goal of RL agent is to search for the optimal policy π^* which maximizes the discounted accumulative reward function. In Q-learning, the state-action function of an agent at the pair of state-action $(s_{q,t}, a_{q,t})$ is given as follows.

$$Q^{\pi}(s_{q,t}, a_{q,t}) = \mathbb{E}\Bigg[\beta \sum_{s_{q,t+1}} \mathcal{T}_{ss'} \sum_{a_{q,t+1}} \pi(s_{q,t+1}, a_{q,t+1}) \\ Q^{\pi}(s_{q,t+1}, a_{q,t+1}) + r_{q,t}\Bigg] \tag{4.23}$$

and the optimal value is given by

$$Q^{*}(s_{q,t}, a_{q,t}) = \max_{\pi} Q^{\pi}(s_{q,t}, a_{q,t}) \tag{4.24}$$

Moreover, the Q-value is updated as follows:

$$Q_{q,t+1}(s_{q,t}, a_{q,t}) = Q_t(s_{q,t}, a_{q,t}) + \lambda[r_{q,t+1} \\ +\beta \max_{a_{q,t+1}} Q_t(s_{q,t+1}, a_{q,t+1}) - Q_t(s_{q,t}, a_{q,t})] \tag{4.25}$$

where $\lambda \in (0, 1\}$ shows the learning rate.

The optimal policy π^* at state $s_{q,t}$ is given as follows.

$$\pi^{*}(s_{q,t}) = \arg\max_{a_{q,t}} Q^{*}(s_{q,t}, a_{q,t}) \tag{4.26}$$

4.3.2 Multi-Agent DRL based joint optimization of beamforming and power allocation

Multi-agent Q-learning (MAQL) method works well for limited set of state and action spaces [259, 260]. However, it becomes difficult for the large-dimension state and action spaces to store all the state-action pairs in Q-table. In a practical heterogeneous network scenario, the size of the Q-table increases with a rise in the number of state-action spaces. Therefore, to solve the MDP with large state-action pairs and to obtain an optimal policy, we introduce a DRL based scheme. It uses a multi-agent DQN, that approximates the Q function $Q(s_q, a_q; \theta)$ by combining MAQL with Deep

Algorithm 4 Multi-agent DRL based joint optimization of beamforming vectors and transmit power

Input:
1) Discount Factor β;
2) Learning Rate λ;
3) Batch size $\flat$;
Output: Optimal sequence of actions

Initialize the replay buffer with size V
Initialize the train network $Q(s_{q,t}, a_{q,t}; \theta)$ for all FBS controllers.
Initialize the parameters of target Q network.
for $(Ep = 1; Ep \leq N; Ep++)$ **do**
    Initialize $s_{q,t}$.
    for $(t = 1; t \leq T; t++)$ **do**
        Select action $a_{q,t}$ at state $s_{q,t}$ from $Q(s_{q,t}, a_{q,t}; \theta)$ by using the ϵ-greedy policy.
        Compute the achievable rate of FUs, i.e., R_{qf}.
        Compute the value of SINRs for macrocell users and femtocell users, i.e., ψ_m,ψ_{qf}.
        Compute the total power consumption at femtocells, i.e., P_Q^{Total}.
        Calculate immediate reward $r_{q,t}$.
        Each controller transits to a new state $s_{q,t+1}$, set $s_{q,t} \rightarrow s_{q,t+1}$.
        Store a tuple of $(s_{q,t}, a_{q,t}, r_{q,t}, s_{q,t+1})$ in replay memory of size V.
        Determine the least square loss using $L_{q,t}(\theta) = \mathbb{E}[(y_q^{DDQN} - Q_q(s_{q,t}, a_{q,t}; \theta))^2]$.
        Each controller updates target network weights ϕ^-.
    end for
end for

Neural Network (DNN), where θ represents the DQN parameters.

In DQN, the function $Q(s_q, a_q; \theta)$ is determined by using θ. Thus, the challenge to find the best Q-function in an extremely large function space can be reduced for finding the best θ of finite dimension. Also, the DQN agent only stores weights in its local memory, reducing computation complexity. We define two DQNs, as suggested by the "quasi-static target network" technique [271], the target DQN with the parameters θ' and the train DQN with the parameters θ. Every controller captures and store their experiences in replay buffer as a tuple $(s_{q,t}, a_{q,t}, r_{q,t}, s_{q,t+1})$.Also, a mini-batch sample is sampled from the experience replay buffer and utilized for updating the parameters of the train DQN network in each iteration.

The least square loss of train DQN at time t is expressed as follows.

$$L_{q,t}(\theta) = \mathbb{E}[(y_{q,t}^{DQN} - Q_q(s_{q,t}, a_{q,t}; \theta))^2] \tag{4.27}$$

where $y_{q,t}^{DQN}$ is the target value defined as follows:

$$y_{q,t}^{DQN} = \varpi \max_{a_{(q,t+1)} \in \mathcal{A}} Q_q(s_{(q,t+1)}, a_{(q,t+1)}, \theta') + r_{q,t} \tag{4.28}$$

where ϕ^- indicates the target network parameters. Also, we assume that the sampled mini-batch is trained using a stochastic gradient descent approach that minimizes

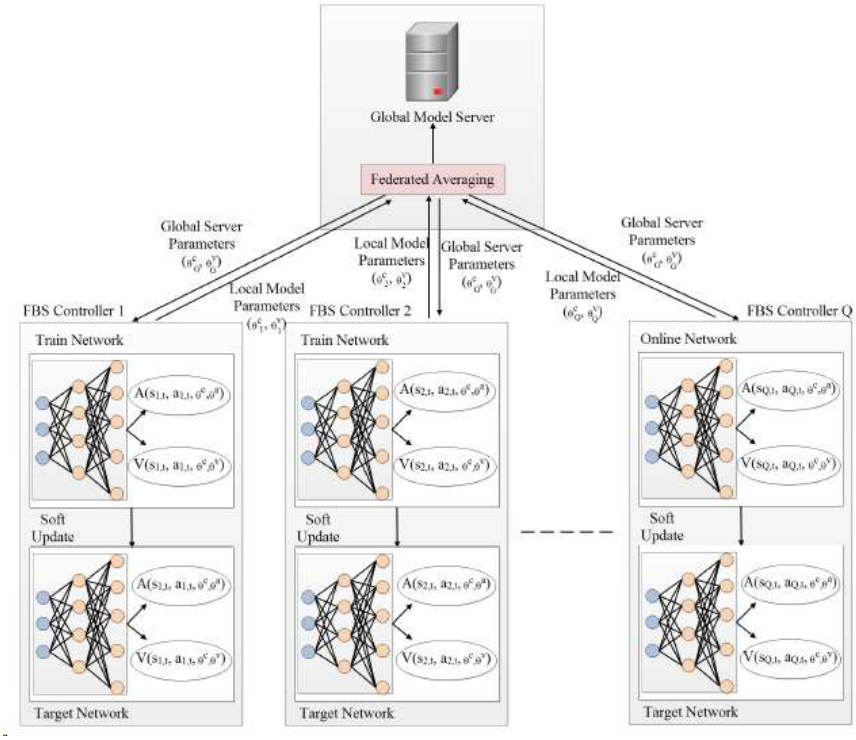

Figure 4.4: Proposed federated DRL architecture

the loss function at each time step.

However, as the same values have been used to choose and analyze the actions in the multi-agent DQN technique, the Q-value may be overestimated [261]. For example, if an action $a_{q,t}$ has a higher value than action $a_{q,t+1}$ in some states, the agents will always choose action $a_{q,t}$. Now, if the action $a_{q,t+1}$ becomes the superior option for some memory experience, the neural network finds it difficult to learn that action $a_{q,t+1}$ is better than action $a_{q,t}$ for these instances, because the neural network has been trained to give higher value for action $a_{q,t}$. Thus, multi-agent double DQN (DDQN) is considered to resolve the preceding issue, which uses the train Q network to select the actions and the target Q network is used for evaluating the action. Multi-agent DDQN replaces the target $y_{q,t}^{DQN}$ with the target $y_{q,t}^{DDQN}$, which is defined as follows.

$$y_{q,t}^{DDQN} = \varpi Q_q(s_{(q,t+1)}, \max_{a_{(q,t+1)} \in \mathcal{A}} Q_q(s_{(q,t+1)}, a_{(q,t+1)}; \phi); \phi^-) + r_{q,t} \tag{4.29}$$

Also, the least square loss for DDQN is expressed as follows.

$$L_{q,t}(\theta) = \mathbb{E}[(y_q^{DDQN} - Q_q(s_{q,t}, a_{q,t}; \theta))^2] \tag{4.30}$$

Both the train and target Q networks in DDQN use the next state $s_{(q,t+1)}$ to determine the optimal Q value $Q_q(s_{(q,t+1)}, a_{(q,t+1)}; \theta)$. Further, since the Q-value function describes how beneficial an action $a_{q,t}$ can be at a particular state $s_{q,t}$, we have used the duelling NN architecture [262] to estimate an advantage function $A(s_{q,t}, a_{q,t}; \theta^c, \theta^a) = Q_q(s_{q,t}; a_{q,t}; \theta^c, \theta^v, \theta^a) - V(s_{q,t}; \theta^c, \theta^v)$, where $V(s_{q,t}; \theta^c, \theta^v)$ represents the value function, and $\theta^c, \theta^a, \theta^v$ are parameters of common network, advantage function parameters, and value function parameters. In D3QN, the advantage function $A(s_{q,t}, a_{q,t}; \theta^c, \theta^a)$ denotes the advantage of an action $a_{q,t}$ over other actions. Thus, the final layer of multi-agent DDQN is divided into two different subnetworks in multi-agent D3QN to estimate the advantage and value functions. The architecture and pseudo-code of multi-agent DRL based anti-jamming transmission scheme is presented in Fig. 4.3 and Algorithm 4, respectively.

Algorithm 5 Proposed federated DRL based scheme against jamming

Input:
1) Discount Factor β;
2) Learning Rate λ;
3) Batch size $\flat$;

Output: Optimal sequence of actions.

1: Each controller initializes two D3QNs, i.e., $Q(s_q, a_q; \theta)$ and $Q(s_q, a_q; \theta^-)$ and a replay memory.
2: Initialize the global DQN.
3: **for** $(Episode = 1; Episode \leq N; Episode++)$ **do**
4: Initialize $s_{q,t}$
5: **for** $(t = 1; t \leq T; t++)$ **do**
6: Each agent visits the initialized state $s_{q,t}$ and perform action $a_{q,t}$ according to current policy $\pi(a_{q,t}|s_{q,t})$.
7: Compute the SINRs of macrocell user and femtocell user, i.e., ψ_m and ψ_{qf}.
8: Compute the achievable rate of femtocell users, i.e., R_{qf}.
9: Compute the reward r_t
10: Each agent obtains new state $s_{q,t+1}$, i.e., set $s_t \rightarrow s_{q,t+1}$.
11: Store the experience $(s_{q,t}, a_{q,t}, r_{q,t}, s_{q,t+1})$ in a replay memory.
12: Global model server selects FBS controllers from the set of the $\mathcal{Q} = \{1, 2, \ldots, Q\}(\forall q \in \mathcal{Q})$ controllers.
13: Determine the least square loss using $L_{q,t}(\theta) = \mathbb{E}[(y_q^{DDQN} - Q_q(s_{q,t}, a_{q,t}; \theta))^2]$.
14: The selected controllers samples the mini-batch transition from its replay buffer and update its local D3QN parameters.
15: The selected controllers sample a mini-batch transition from its replay buffer and update its local D3QN parameters.
16: Each selected controller provides parameters θ^v and θ^c to the global server, and the global server aggregates them using $\theta_G^v = \frac{1}{|\nu_t|} \sum_{q \in \nu_t} \theta_q^v$ and $\theta_G^c = \frac{1}{|\nu_t|} \sum_{q \in \nu_t} \theta_q^c$, where ν_t denotes the set of selected controllers.
17: Global server passes the new global parameters to all the controllers.
18: Each controller combines its locally trained parameters with obtained global parameters.
19: **end for**
20: **end for**

4.3.3 Federated DRL based scheme

Although DRL can determine the best strategy efficiently, it also requires a lot of computational resources. DRL agent training has the following limitations [272]: i) Multi-agent DRL takes significant amount of the time to properly train each agent. ii) Although the training data can be modified to secure privacy, but the received agent data is less relevant and directed among the UEs.

To mitigate the above-mentioned issues, we further propose a federated DRL based scheme against jamming for 5G HetNets. Since the goal is to enhance the overall security performance, so FBS controllers should consider the environment's common knowledge and local information, represented by the global parameters of DQN. Also, to train the global DQN parameters as efficiently as possible, the shared knowledge of environment must be expressed without taking controllers' local preferences in action space. Here, we have considered partial D3QN parameters of each FBS controller to depict the controllers' consensus and trained some local parameters to include the local preferences.

In the proposed federated DRL based anti-jamming technique, each controller takes decisions depending on the trained D3QN parameters $\{\theta^a_{q,t}, \theta^c_{G,t}, \theta^v_{G,t}\}$, trained by federated learning based training algorithm. Also, to obtain long-term performance goals, global model server is placed for selecting FBS controllers in order to perform the local training, and further aggregating the parameters acquired by the chosen controllers for updating the global D3QN parameters. The proposed federated learning based approach consists of initialization and the training phase. In initialization phase, firstly, each controller establish two D3QNs, i.e., the train D3QN with $Q(s_{q,t}, a_{q,t}, \theta)$ and the target D3QN with $Q(s_{q,t}, a_{q,t}, \theta^-)$, and a replay memory to store their experience in form of a tuple $(s_{q,t}, a_{q,t}, r_{q,t}, s_{q,t+1})$.

In training phase, at every time slot t, controller q visits a state $s_{q,t}$ and makes its beamforming and power allocation decision based on trained D3QN. Each controller gets an immediate reward and moves to a new state after executing the selected action. Then, the global server selects multiple controllers to undertake local D3QN training using beamforming and power allocation strategy. Also, each chosen controller samples a mini-batch transition of experiences randomly from the replay memory. Each selected controller then provides its value function parameter (θ^v) and the common network parameter (θ^c) to the global model server. Then, the global model server aggregates these parameters to achieve value function parameter (θ^v_G) and also the common network parameter (θ^c_G). Finally, the global parameters are passed to all the controllers, where these parameters are merged with local advantage function parameter (θ^a) to obtain new local D3QN parameters. The architecture and pseudo-code of the proposed federated DRL based scheme

are presented in Fig. 4.4 and Algorithm 5, respectively.

4.3.4 Computational Complexity Analysis

The computational complexity of the proposed scheme mainly depends on the DNN model and its learning process. In the DNN training process, let l_t, i_k, and i_0, denotes the number of training layers in DNN, neurons in k^{th} layer, and the size of the input layer proportional to the total possible states. For each agent, the computational complexity of each time step is $\mathcal{O}(i_0 i_k + \sum_{k=1}^{K-1} i_k i_{k+1})$. Further, in the learning phase of the proposed scheme, each mini-batch transition has N episodes, and each episode has T time steps. Thus, the computational complexity of proposed scheme is $\mathcal{O}(NT(i_0 i_k + \sum_{k=1}^{K-1} i_k i_{k+1}))$.

4.4 Performance Evaluation

In this section, we have evaluated the performance of the proposed federated DRL scheme in comparison to DDQN and DQN.

4.4.1 Numerical Settings

We have considered a downlink HetNet consisting of MBS, FBSs, and jammers in the simulation settings. MUs and FUs are placed inside the radius of macrocell and femtocells, respectively. The radius of the macrocell and femtocell is assumed to be 500m and 50m, respectively. Also, a reactive jammer is placed inside the radius of every femtocell. Unless otherwise mentioned in the following subsections, the MBS, FBSs, and jammers are equipped with $N_M = 8$, $N_f = 8$, and $N_J = 8$ antennas, respectively. The number of FUs in each femtocell is $\mathcal{F} = 4$. Also, the transmit power of MBS, FBSs, and jammers is taken as $P_m = 40$ dBm, $P_f = 20$ dBm, and $P_j = 20$ dBm. It has been observed that RL based algorithms do not require prior knowledge of the jammer, i.e., jamming power, jamming strategy, and the channel activity [273]. Major simulation parameters used are presented in Table 4.3.

4.4.2 Convergence Analysis

In this section, we have compared the reward convergence of proposed federated DRL scheme using different learning rates, i.e., $\lambda = \{0.01, 0.001, 0.001\}$. The learning rate is a tunable hyper-parameter which regulates how quickly the model adjusts to its new task. It has been observed that a high learning rate of 0.01 creates significant variations and relatively fast convergence in reward value for the proposed scheme, as shown in Fig. 4.5. A significant change in reward values can lead to unstable training or divergence, even though it delivers rapid convergence. Moreover, the

Table 4.3: Major simulation parameters used

Parameters	Values
Number of FBSs	3
No. of MUs	6
No. of FUs in each femtocell	4∼6
Transmission power of MBS	40dBm
Noise power spectrum density	-174dBm
Carrier frequency of mmWave Tier	1 28 GHz
Bandwidth of mmWave Tier	1 GHz
FBS to FU path loss exponent	3.75
Path loss constant for LoS, i.e., c_l	-61.4 dB
Path loss constant for non-LoS, i.e., c_n	-72 dB
FBS to FU path loss exponent	3.75
Jammer path loss exponent	2.5
Discount Factor	0.9
Starting Exploration	1
Exploration at final	0.01
Total exploratory steps	1000
Replay buffer capacity	1000
Mini-batch size	32
Weights of reward function	1,1
Updated weight intervals	10
Activation function used	ReLu
Model optimizer	Adam

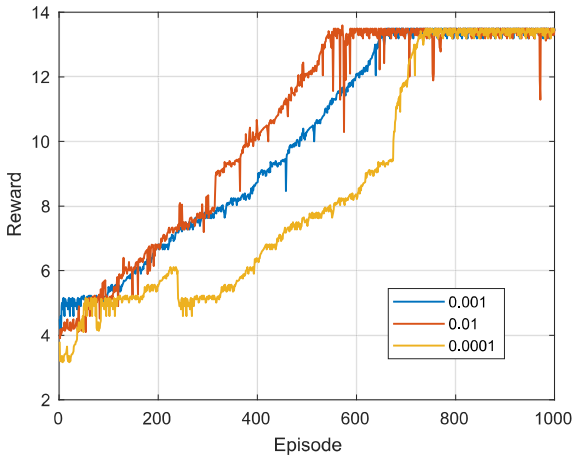

Figure 4.5: Reward performance with different learning rates

proposed scheme takes more time and episodes to reach convergence at a shallow learning rate of 0.0001. However, the moderate learning rate, i.e., $\lambda = 0.001$ provides the most consistent reward values in the proposed scheme. Therefore, we set the learning rate of $\lambda = 0.001$ for all the simulations, neither too high nor too low.

4.4.3 Achievable Rate Analysis

In Fig. 4.6(a), the convergence analysis of the total achievable rate per femtocell for federated DRL, DDQN, and DQN schemes is analyzed with respect to the number of episodes. It has been observed that with an increase in the number of episodes, the total achievable rate per femtocell of federated DRL, DDQN, and DQN schemes reaches to its maximum floors and converges roughly after 600 episodes. Also, it has

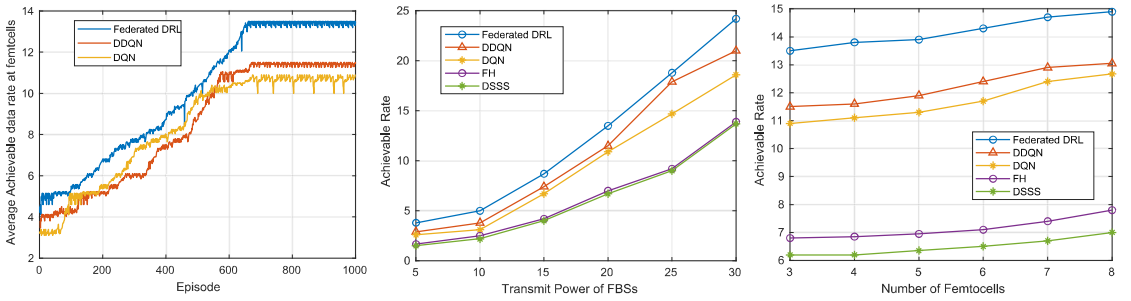

Figure 4.6: Comparative Analysis: (a) Total Achievable Rate per femtocell (bits/s/Hz) v/s Number of Episodes (b) Total Achievable Rate per femtocell v/s P_f (dBm) (c) Total Achievable Rate per femtocell v/s N_f

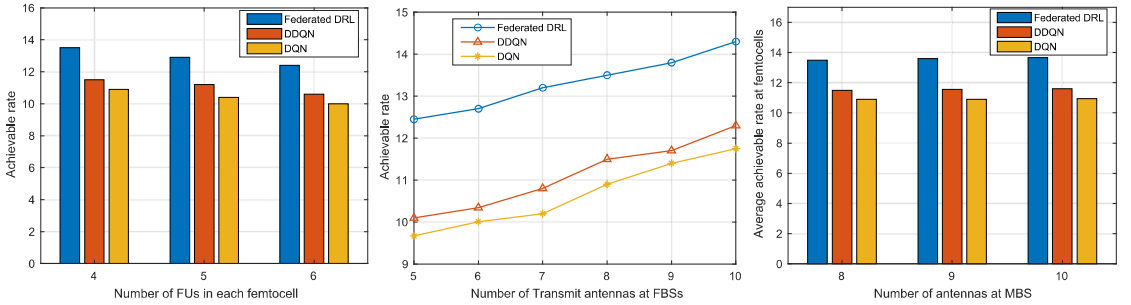

Figure 4.7: Comparative Analysis: (a) Total Achievable Rate per femtocell (bits/s/Hz) v/s Number of FUs in each femtocell (b) Total Achievable Rate per femtocell v/s N_f (c) Total Achievable Rate per femtocell v/s N_M

been observed that proposed federated DRL scheme reaches the convergence faster and also outperforms DDQN and DQN schemes in terms of maximum achievable data rates at femtocells. During the training, the proposed scheme takes about 105 minutes to achieve convergence as compared to DDQN (150 minutes) and DQN (170 minutes).

4.4.3.1 Impact of Transmit Power

Fig. 4.6(b) shows the total achievable rate per femtocell with respect to varying transmit power levels of FBSs, i.e., $P_f = \{5, 10, 15, 20, 25, 30\}$ dBm. It has been observed that the achievable rate of federated DRL, DDQN, DQN, FH, and DSSS schemes increases with an increase in the transmit power level of FBSs. It is because as P_f rises, the received SINR at FUs also increases, resulting in an increase in the FUs' achievable data rates. Also, at $P_f = 20$ dBm, the proposed federated DRL scheme achieves 17.39% and 23.85% better total achievable rate per femtocell in comparison to the DDQN and DQN schemes, respectively. Also, the proposed scheme outperforms conventional FH and DSSS techniques significantly.

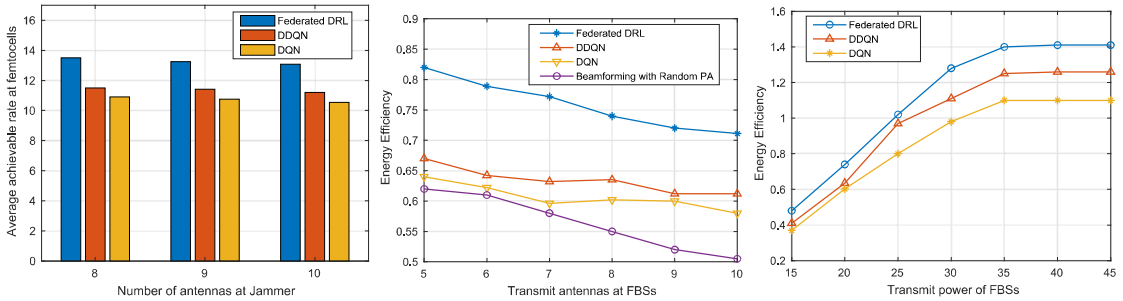

Figure 4.8: Comparative Analysis: (a) Total Achievable Rate per femtocell (bits/s/Hz) v/s Number of antennas at Jammers (b) Energy Efficiency (bits/s/Hz/W) v/s N_f (c) Energy Efficiency (bits/s/Hz/W) v/s P_f

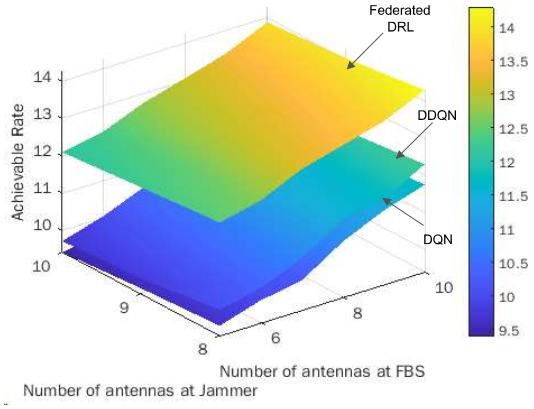

Figure 4.9: Total Achievable Rate per femtocell v/s number of transmit antennas at jammer v/s number of transmit antennas at FBSs

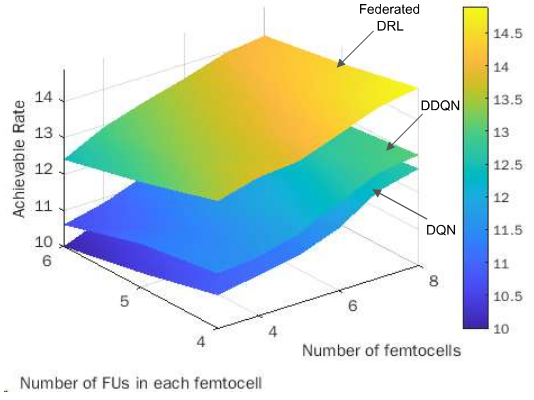

Figure 4.10: Total Achievable Rate per femtocell v/s number of FUs in each femtocell v/s number of femtocells

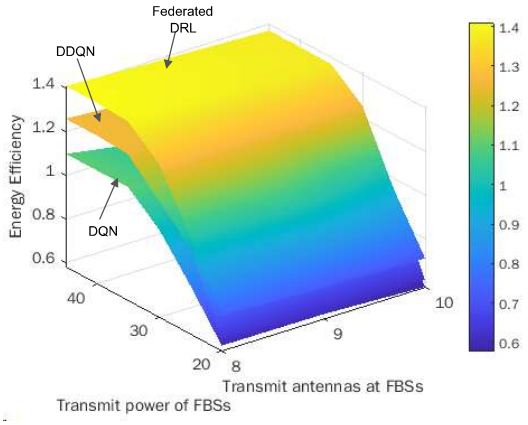

Figure 4.11: Energy Efficiency v/s Transmit power of FBSs v/s Transmit antennas at FBSs

4.4.3.2 Impact of number of Femtocells

Fig. 4.6(c) shows the total achievable rate per femtocell versus the total number of femtocells, i.e., $\mathcal{Q} = \{3, 4, 5, 6, 7, 8\}$. It has been observed that the proposed federated DRL scheme outperforms DDQN, DQN, FH and DSSS schemes in terms of the achievable rate performance of FUs. As observed from Fig. 4.6(c), with an increase in the total number of femtocells, the achievable rates of all the schemes increase. It is because when the number of femtocells increases, the SINR at the FUs of all femtocells also increases significantly, resulting in improved achievable rate performance.

4.4.3.3 Impact of number of FUs

In Fig. 4.7(a), the total achievable rate per femtocell of federated DRL, DDQN, and DQN schemes is explored with respect to different numbers of FUs at each femtocell, i.e., $\mathcal{F} = \{1, 2, 3, 4, 5, 6\}$. It is evident from Fig. 4.7(a) that as the number of FUs in each femtocell increases, the total achievable rate per femtocell decreases. This is because as the number of FUs connected to the same FBS increases, the co-channel interference also increases, resulting in a decrease in the average achievable rate performance. Also, it has been observed that the proposed federated DRL scheme outperforms DDQN and DQN schemes in terms of the achievable rate performance.

4.4.3.4 Impact of FBS transmit antennas

Fig. 4.7(b) shows the total achievable rate per femtocell versus the transmit antennas at FBSs, i.e., $N_f = \{5, 6, 7, 8, 9, 10\}$ for federated DRL, DDQN, and DQN schemes. It has been observed that with an increase in the number of antennas at each FBS, the total achievable rate per femtocell of federated DRL, DDQN, and DQN schemes

increases. It is because as the number of transmit antennas at FBSs increases, the signal strength also increases, resulting in an improved achievable rate performance. Also, the proposed federated DRL scheme performs better than DDQN and DQN schemes in terms of achievable rate.

4.4.3.5 Impact of MBS transmit antennas

Fig. 4.7(c) depicts the total achievable rate per femtocell with respect to different number of transmit antennas at MBS, i.e., $N_M = \{8, 9, 10\}$. It has been observed that increasing the number of N_M has no effect on FUs' achievable rate. It is due to the presence of different channel models in macrocell and femtocells. Also, it has been observed that the proposed federated DRL scheme outperforms DDQN and DQN schemes.

Table 4.4: Relative Comparison of the Simulation Results

Parameters	**Our Proposed Scheme**	**DDQN**	**DQN**	**DSSS**	**FH**
Training Time	105 minutes	150 minutes	170 minutes	-	-
Convergence Reached (episodes)	after 620 episodes	after 650 episodes	after 700 episodes	-	-
Achievable Data Rate at FUs	13.5 bits/s/Hz	11.5 bits/s/Hz	10.9 bits/s/Hz	6.7 bits/s/Hz	7 bits/s/Hz
Energy-efficiency	0.735 bits/s/Hz/W	0.62 bits/s/Hz/W	0.6 bits/s/Hz/W	0.361 bits/s/Hz/W	0.378 bits/s/Hz/W

4.4.3.6 Impact of number of antennas at jammers

Fig. 4.8(a) shows the total achievable rate per femtocell versus the number of jammers transmit antennas, i.e., $N_J = \{8, 9, 10\}$. It has been observed that as number of N_J increases, the total achievable rate per femtocell decreases. The reason behind this is with an increase in the number of jammers transmit antennas, the jamming capability of the jammer also increases, which results in a decrease in the achievable rate of FUs. Also, it has been observed that the federated DRL scheme performs better than DDQN and DQN schemes in terms of total achievable rate per femtocell performance.

4.4.4 Energy-Efficiency Analysis

In this subsection, we evaluated the energy efficiency of the proposed federated DRL scheme with DDQN and DQN schemes. Similar to [274] and [266], the numerical setting for this simulation is $\zeta = 2.6$, $P_a = 30$ dBm, and $P_b = 40$ dBm.

4.4.4.1 Impact of FBSs' transmit antennas on energy-efficiency

It has been observed from Fig. 4.8(b) that the energy-efficiency of femtocells slightly decreases with an increase in N_f. This is because with an increase in the number of FBS's transmit antennas, the total power consumption at each FBS also increases; thus, the energy-efficiency of femtocells decreases. Also, it has been observed that the proposed federated DRL scheme outperforms DDQN, DQN, and beamforming with random power allocation schemes in terms of energy-efficiency.

4.4.4.2 Impact of transmit power of FBSs on energy-efficiency

In Fig. 4.8(c), we analyze the energy efficiency of the proposed scheme with respect to transmit power of FBSs. It has been observed that the energy-efficiency first increases and then converge to its maximum level as P_f increases. It firstly increases due to the fact that large transmit power can obtain higher achievable rates, but the rate of increase in achievable rate slows down after a certain transmit power level. Also, the proposed scheme achieves higher energy-efficiency as compared to DDQN and DQN schemes.

4.4.5 Achievable Rate Region of Femtocells

To understand the behavior of the proposed federated DRL scheme, we evaluated the achievable rate area of femtocells versus N_f and N_J in a three-dimensional (3D) figure in Fig. 4.9. It is evident from Fig. 4.9 that the total achievable rate per femtocell increases when we jointly increase the values of N_f and N_J. Although the

increase in the number of jammer's transmit antennas decreases the achievable rate. Still, this loss is overshadowed by the increase in achievable rate with respect to an increase in the number of FBS transmit antennas. Also, the proposed federated DRL scheme outperforms the DDQN and DQN schemes in terms of the acquired achievable rate region of femtocells.

Moreover, in Fig. 4.10, we have compared the achievable rate region of femtocell with respect to increasing number of femtocells, i.e., $\mathcal{Q} = \{3, 4, 5, 6, 7, 8\}$ and FUs $\mathcal{F} = \{4, 5, 6\}$. It has been observed that jointly increasing the number of femtocells and FUs leads to a slow and steady increase in the total achievable rate per femtocell. Also, it is evident from Fig. 4.10 that the proposed scheme performs better than DDQN and DQN schemes.

4.4.6 Energy-Efficient Region of Femtocells

To clearly understand the effect of FBSs' transmit power and antennas on the energy efficiency, we have demonstrated the energy-efficiency region of femtocells acquired by the proposed federated DRL scheme in Fig. 4.11. It has been observed that the energy-efficiency of femtocells increases with an increase in the value of P_f, and then reaches convergence after reaching the maximum values. While the energy-efficiency decreases a little with an increase in the transmit antennas at FBS. The proposed scheme outperforms DDQN and DQN schemes in terms of efficient energy-efficiency of femtocells. The comparison of the simulation results between the proposed scheme and other existing anti-jamming techniques is presented in Table 4.4.

4.5 Summary

In this chapter, we maximize the achievable data rate and reduce the energy consumption of mmWave enabled 5G HetNets in the presence of malicious jammers. To achieve this goal, a joint optimization problem of beamforming and power allocation is formulated. We translate the non-convex optimization problem into multi-agent RL problem using MDP. Further, we have proposed a mult-agent federated DRL scheme to enhance the data rate of 5G enabled HetNet. The simulation results demonstrate the effectiveness of the proposed federated DRL scheme in achieving better achievable data rate performance than other state-of-the-art schemes in the presence of malicious jammers.

Chapter 5

Conclusion and Future Scope

This chapter gives the concluding remarks on the research work related to applying AI enabled PLS techniques in 5G heterogeneous networks for enhancing the security levels. Moreover, it also points out the future scope of the proposed AI enabled PLS techniques for 5G HetNets which an be taken as future works in these domains.

It has been predicted that by the year 2030, 5G and beyond 5G (B5G) networks are expected to provide hundreds of trillions of gigabytes of data for various emerging applications such as augmented, mixed, and virtual reality (AR/MR/VR), wireless computer-brain interfaces (WCBI), connected robotics and autonomous systems. Most of these applications share data with each other using an open channel, i.e., the Internet. The open and broadcast nature of wireless channel makes the communication susceptible to various types of attacks (e.g., eavesdropping, jamming). Thus, there is a strong requirement to enhance the secrecy of wireless channel to maintain the privacy and confidentiality of transmitted data. Physical layer security (PLS) has evolved as a novel concept and robust alternative to cryptographybased techniques, which have a number of drawbacks and practical issues for 5G and beyond networks.

Firstly, we present an in-depth survey on conventional PLS techniques and various AI-enabled secure data transmission techniques to understand the conceptual and practical challenges linked to 5G and beyond HetNets. We provide an in-depth analysis of the use of AI in various PLS applications such as security oriented beamforming, cooperative jamming, resource allocation and power control, etc. SWOT analysis of these secure data transmission techniques is also performed.

Then, we proposed SecBoost, a secrecy-aware joint optimization of power control, channel allocation, and beamforming by utilizing multi-agent cooperative DRL to maximize the SEE of picocell users in 5G HetNets. Firstly, the optimization problem is transformed into MARL problem using MDP. Then, to address the issue of large state-action spaces of MDP in 5G HetNets, we have presented a dueling architecture of multi-agent D3QN to determine the optimal policy of MDP. Further, to

improve the sampling efficiency of mini-batch transitions from the replay buffer, we have used prioritized experience replay. The numerical results demonstrate that the proposed SEE maximization scheme, SecBoost outperforms the MA-DQN, MARL, and the existing state-of-the-art scheme JBF-SEEM in terms of average SEE performance of picocell users. In the future, the scalability of SecBoost will be explored with multiple eavesdroppers in each picocell for the k-tier heterogeneous networks.

Further, we propose a federated DRL-based transmission scheme to maximize the achievable rate of FUs in the presence of jammers in 5G HetNet. Firstly, we formulate an anti-jamming joint optimization problem of beamforming and power allocation to achieve the maximum data rate for FUs. Then, we have used the MDP to convert the non-convex optimization problem into multi-agent RL problem. Also, a federated DRL-based optimization of beamforming vectors and power allocation is proposed to solve the Markov decision process with huge state and action spaces. Numerical analysis shows that the proposed federated DRL scheme outperforms DDQN and DQN in terms of the achievable rate performance of mmWave enabled femtocells. In future, the scalability of the proposed scheme will be explored for k-tier HetNets, and multiple eavesdroppers and jammers in each femtocell.

Bibliography

[1] M. Series, "Imt vision–framework and overall objectives of the future development of imt for 2020 and beyond," *Recommendation ITU*, vol. 2083, no. 0, 2015.

[2] D. Liu, W. Hong, T. S. Rappaport, C. Luxey, and W. Hong, "What will 5g antennas and propagation be?" *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 12, pp. 6205–6212, 2017.

[3] A. Ghosh, N. Mangalvedhe, R. Ratasuk, B. Mondal, M. Cudak, E. Visotsky, T. A. Thomas, J. G. Andrews, P. Xia, H. S. Jo *et al.*, "Heterogeneous cellular networks: From theory to practice," *IEEE communications magazine*, vol. 50, no. 6, pp. 54–64, 2012.

[4] Global heterogeneous networks (hetnets) industry, 2021. Accessed: 2021-09-20. [Online]. Available: https://www.globenewswire.com/en/news-release/2020/07/10/2060645/0/en/Global-Heterogeneous-Networks-HetNets-Industry.html

[5] Y. Xu, G. Gui, H. Gacanin, and F. Adachi, "A survey on resource allocation for 5G heterogeneous networks: Current research, future trends, and challenges," *IEEE Commun. Surv. Tut.*, vol. 23, no. 2, pp. 668–695, Feb. 2021.

[6] Z. Hasan, H. Boostanimehr, and V. K. Bhargava, "Green cellular networks: A survey, some research issues and challenges," *IEEE Communications surveys & tutorials*, vol. 13, no. 4, pp. 524–540, 2011.

[7] A. Shaik, R. Borgaonkar, N. Asokan, V. Niemi, and J.-P. Seifert, "Practical attacks against privacy and availability in 4g/lte mobile communication systems," 01 2016.

[8] D. F. Kune, J. Koelndorfer, N. Hopper, and Y. Kim, "Location leaks on the gsm air interface," *ISOC NDSS (Feb 2012)*, 2012.

[9] K. K. Sehra and M. Dave, "Privacy preserving data aggregation in wireless body sensor network," in *Proceedings of the 2nd International Conference on IoT, Social, Mobile, Analytics & Cloud in Computational Vision & Bio-Engineering (ISMAC-CVB 2020)*, 2020.

[10] I. Ahmad, S. Shahabuddin, T. Kumar, J. Okwuibe, A. Gurtov, and M. Ylianttila, "Security for 5g and beyond," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3682–3722, 2019.

[11] Y. Liu, H.-H. Chen, and L. Wang, "Physical layer security for next generation wireless networks: Theories, technologies, and challenges," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 1, pp. 347–376, 2016.

[12] J. M. Hamamreh, H. M. Furqan, and H. Arslan, "Classifications and applications of physical layer security techniques for confidentiality: A comprehensive survey," *IEEE Communications Surveys Tutorials*, vol. 21, no. 2, pp. 1773–1828, 2019.

[13] Y. Wu, A. Khisti, C. Xiao, G. Caire, K.-K. Wong, and X. Gao, "A survey of physical layer security techniques for 5g wireless networks and challenges ahead," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 4, pp. 679–695, 2018.

[14] H. Yang, H. Liu, C. Luo, Y. Wu, W. Li, A. Y. Zomaya, L. Song, and W. Xu, "Vehicle-key: A secret key establishment scheme for lora-enabled iov communications," in *2022 IEEE 42nd International Conference on Distributed Computing Systems (ICDCS)*. IEEE, 2022, pp. 787–797.

[15] A. D. Wyner, "The wire-tap channel," *Bell system technical journal*, vol. 54, no. 8, pp. 1355–1387, 1975.

[16] I. Csiszár and J. Korner, "Broadcast channels with confidential messages," *IEEE transactions on information theory*, vol. 24, no. 3, pp. 339–348, 1978.

[17] F. Zhu, F. Gao, H. Lin, S. Jin, J. Zhao, and G. Qian, "Robust beamforming for physical layer security in bdma massive mimo," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 4, pp. 775–787, 2018.

[18] Y. Huo, X. Fan, L. Ma, X. Cheng, Z. Tian, and D. Chen, "Secure communications in tiered 5g wireless networks with cooperative jamming," *IEEE Transactions on Wireless Communications*, vol. 18, no. 6, pp. 3265–3280, 2019.

[19] T. Ma, F. Hu, and M. Ma, "Fast and efficient physical layer authentication for 5g hetnet handover," in *2017 27th International Telecommunication Networks and Applications Conference (ITNAC)*. IEEE, 2017, pp. 1–3.

[20] Z. Lv, A. K. Singh, and J. Li, "Deep learning for security problems in 5g heterogeneous networks," *IEEE Network*, vol. 35, no. 2, pp. 67–73, 2021.

[21] Y. Huo, Y. Wu, R. Li, Q. Gao, and X. Luo, "A learning-aided intermittent cooperative jamming scheme for non-slotted wireless transmission in an iot system," *IEEE Internet of Things Journal*, 2021.

[22] R.-F. Liao, H. Wen, J. Wu, F. Pan, A. Xu, Y. Jiang, F. Xie, and M. Cao, "Deep-learning-based physical layer authentication for industrial wireless sensor networks," *sensors*, vol. 19, no. 11, p. 2440, 2019.

[23] N. M. Alotaibi and S. S. Alwakeel, "A neural network based handover management strategy for heterogeneous networks," in *2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA)*, 2015, pp. 1210–1214.

[24] T. M. Hoang, T. Q. Duong, H. D. Tuan, S. Lambotharan, and L. Hanzo, "Physical layer security: Detection of active eavesdropping attacks by support vector machines," *IEEE Access*, vol. 9, pp. 31 595–31 607, 2021.

[25] X. Wang and F. Liu, "Data-driven relay selection for physical-layer security: A decision tree approach," *IEEE Access*, vol. 8, pp. 12 105–12 116, 2020.

[26] J. Wang, Y. Zou, and J. Ding, "Ads-b spoofing attack detection method based on lstm," *EURASIP Journal on Wireless Communications and Networking*, vol. 2020, no. 1, pp. 1–12, 2020.

[27] R. Liao, H. Wen, F. Pan, H. Song, A. Xu, and Y. Jiang, "A novel physical layer authentication method with convolutional neural network," in *2019 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)*, 2019, pp. 231–235.

[28] X. Xiao, B. Vasić, R. Tandon, and S. Lin, "Designing finite alphabet iterative decoders of ldpc codes via recurrent quantized neural networks," *IEEE Transactions on Communications*, vol. 68, no. 7, pp. 3963–3974, 2020.

[29] T. Marchioro, N. Laurenti, and D. Gündüz, "Adversarial networks for secure wireless communications," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 8748–8752.

[30] X. Tan, S. Su, Z. Zuo, X. Guo, and X. Sun, "Intrusion detection of uavs based on the deep belief network optimized by pso," *Sensors*, vol. 19, no. 24, p. 5529, 2019.

[31] T. Erpek, T. J. O'Shea, Y. E. Sagduyu, Y. Shi, and T. C. Clancy, "Deep learning for wireless communications," in *Development and Analysis of Deep Learning Architectures*. Springer, 2020, pp. 223–266.

[32] Y. Hu, L. Li, J. Yin, H. Zhang, W. Liang, A. Gao, and Z. Han, "Optimal transmit antenna selection strategy for mimo wiretap channel based on deep reinforcement learning," in *2018 IEEE/CIC International Conference on Communications in China (ICCC)*, 2018, pp. 803–807.

[33] X. Zhang and S. Sun, "Dynamic optimization for secure mimo beamforming using large-scale reinforcement learning," in *2019 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2019, pp. 1–6.

[34] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski *et al.*, "Human-level control through deep reinforcement learning," *nature*, vol. 518, no. 7540, pp. 529–533, 2015.

[35] Y. Hu, L. Li, J. Yin, H. Zhang, W. Liang, A. Gao, and Z. Han, "Optimal transmit antenna selection strategy for mimo wiretap channel based on deep reinforcement learning," in *2018 IEEE/CIC International Conference on Communications in China (ICCC)*. IEEE, 2018, pp. 803–807.

[36] L. Vailshery, "Number of internet of things (iot) connected devices worldwide in 2018, 2025 and 2030," *Date accessed: September*, 2021.

[37] Y. Zhang, W. He, X. Li, H. Peng, K. Rabie, G. Nauryzbayev, B. M. ElHalawany, and M. Zhu, "Covert communication in downlink noma systems with channel uncertainty," *IEEE Sensors Journal*, vol. 22, no. 19, pp. 19 101–19 112, 2022.

[38] F. Irram, M. Ali, M. Naeem, and S. Mumtaz, "Physical layer security for beyond 5g/6g networks: Emerging technologies and future directions," *Journal of Network and Computer Applications*, p. 103431, 2022.

[39] A. Kakkar, "A survey on secure communication techniques for 5g wireless heterogeneous networks," *Information Fusion*, vol. 62, pp. 89–109, 2020.

[40] G. Karopoulos, G. Kambourakis, and S. Gritzalis, "Survey of secure handoff optimization schemes for multimedia services over all-ip wireless heterogeneous networks," *IEEE Communications Surveys Tutorials*, vol. 9, no. 3, pp. 18–28, 2007.

[41] L. Bai, L. Zhu, J. Liu, J. Choi, and W. Zhang, "Physical layer authentication in wireless communication networks: A survey," *Journal of Communications and Information Networks*, vol. 5, no. 3, pp. 237–264, 2020.

[42] N. Haider, M. Z. Baig, and M. Imran, "Artificial intelligence and machine learning in 5g network security: Opportunities, advantages, and future research trends," *arXiv preprint arXiv:2007.04490*, 2020.

[43] A. Mukherjee, S. A. A. Fakoorian, J. Huang, and A. L. Swindlehurst, "Principles of physical layer security in multiuser wireless networks: A survey," *IEEE Communications Surveys Tutorials*, vol. 16, no. 3, pp. 1550–1573, 2014.

[44] Y. Xu, G. Gui, H. Gacanin, and F. Adachi, "A survey on resource allocation for 5g heterogeneous networks: Current research, future trends and challenges," *IEEE Communications Surveys & Tutorials*, 2021.

[45] J. D. V. Sánchez, L. Urquiza-Aguiar, M. C. P. Paredes, and D. P. M. Osorio, "Survey on physical layer security for 5g wireless networks," 2020.

[46] J. Cao, M. Ma, Y. Fu, H. Li, and Y. Zhang, "Cppha: Capability-based privacy-protection handover authentication mechanism for sdn-based 5g hetnets," *IEEE transactions on dependable and secure computing*, vol. 18, no. 3, pp. 1182–1195, 2019.

[47] B. ZHANG and K. HUANG, "Robust secure transmission scheme based on artificial noise-aided for heterogeneous networks with simultaneous wireless information and power transfer," *Journal of Electronics*, pp. 1–8, 2019.

[48] S. Ghosh, M. R. Bhatnagar, A. Singh, and B. K. Panigrahi, "Secrecy capacity in crn with malicious energy harvester using game theoretic techniques," *IEEE Transactions on Cognitive Communications and Networking*, vol. 3, no. 3, pp. 343–360, 2017.

[49] Y. Arjoune and S. Faruque, "Smart jamming attacks in 5G new radio: A review," in *Proc. IEEE Annu. Comput. Commun. Wrkshp. Conf. (CCWC)*, Las Vegas, NV, USA, Mar. 2020, pp. 1010–1015.

[50] J. D. V. Sánchez, L. Urquiza-Aguiar, M. C. P. Paredes, and D. P. M. Osorio, "Survey on physical layer security for 5g wireless networks," *Annals of Telecommunications*, vol. 76, no. 3, pp. 155–174, 2021.

[51] R. Miller and W. Trappe, "On the vulnerabilities of csi in mimo wireless communication systems," *IEEE Transactions on Mobile Computing*, vol. 11, no. 8, pp. 1386–1398, 2012.

[52] Y. Wu, Y. Ma, H.-N. Dai, and H. Wang, "Deep learning for privacy preservation in autonomous moving platforms enhanced 5g heterogeneous networks," *Computer Networks*, vol. 185, p. 107743, 2021.

[53] A. Shaik and R. Borgaonkar, "New vulnerabilities in 5g networks. black hat 2019," 2019.

[54] S. R. Islam, N. Avazov, O. A. Dobre, and K.-S. Kwak, "Power-domain non-orthogonal multiple access (noma) in 5g systems: Potentials and challenges," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 721–742, 2016.

[55] T. Dragičević, P. Siano, and S. S. Prabaharan, "Future generation 5g wireless networks for smart grid: A comprehensive review," *Energies*, vol. 12, no. 11, p. 2140, 2019.

[56] Y. Liu, Z. Qin, M. Elkashlan, Y. Gao, and L. Hanzo, "Enhancing the physical layer security of non-orthogonal multiple access in large-scale networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1656–1672, 2017.

[57] Y. Zhang, H.-M. Wang, Q. Yang, and Z. Ding, "Secrecy sum rate maximization in non-orthogonal multiple access," *IEEE Communications Letters*, vol. 20, no. 5, pp. 930–933, 2016.

[58] Z. Lin, M. Lin, J.-B. Wang, T. de Cola, and J. Wang, "Joint beamforming and power allocation for satellite-terrestrial integrated networks with non-orthogonal multiple access," *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 3, pp. 657–670, 2019.

[59] W. Li, M. Ghogho, B. Chen, and C. Xiong, "Secure communication via sending artificial noise by the receiver: Outage secrecy capacity/region analysis," *IEEE Communications Letters*, vol. 16, no. 10, pp. 1628–1631, 2012.

[60] G. Zheng, I. Krikidis, J. Li, A. P. Petropulu, and B. Ottersten, "Improving physical layer secrecy using full-duplex jamming receivers," *IEEE Transactions on Signal Processing*, vol. 61, no. 20, pp. 4962–4974, 2013.

[61] B. Akgun, O. O. Koyluoglu, and M. Krunz, "Exploiting full-duplex receivers for achieving secret communications in multiuser miso networks," *IEEE Transactions on Communications*, vol. 65, no. 2, pp. 956–968, 2016.

[62] Y. Cao, N. Zhao, Y. Chen, M. Jin, L. Fan, Z. Ding, and F. R. Yu, "Privacy preservation via beamforming for noma," *IEEE Transactions on Wireless Communications*, vol. 18, no. 7, pp. 3599–3612, 2019.

[63] Y. Hao and T. Lv, "Swipt-aided secure beamforming design for downlink cooperative noma systems," in *2018 Global Wireless Summit (GWS)*. IEEE, 2018, pp. 364–369.

[64] N. Zhao, W. Wang, J. Wang, Y. Chen, Y. Lin, Z. Ding, and N. C. Beaulieu, "Joint beamforming and jamming optimization for secure transmission in miso-noma networks," *IEEE Transactions on Communications*, vol. 67, no. 3, pp. 2294–2305, 2018.

[65] Y. Deng, Q. Li, Q. Zhang, L. Yang, and J. Qin, "Secure beamforming design in mimo noma networks for internet of things with perfect and imperfect csi," *Computer Networks*, vol. 187, p. 107839, 2021.

[66] Y. Li, M. Jiang, Q. Zhang, Q. Li, and J. Qin, "Secure beamforming in downlink miso nonorthogonal multiple access systems," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 8, pp. 7563–7567, 2017.

[67] L. Jiang, C. Qin, X. Zhang, and H. Tian, "Secure beamforming design for swipt in cooperative d2d communications," *China Communications*, vol. 14, no. 1, pp. 20–33, 2017.

[68] Y. Feng, S. Yan, Z. Yang, N. Yang, and J. Yuan, "Beamforming design and power allocation for secure transmission with noma," *IEEE Transactions on Wireless Communications*, vol. 18, no. 5, pp. 2639–2651, 2019.

[69] C. Yin and L. Yan, "Secure beamforming design for the uav-enabled transmission over noma networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2020, no. 1, pp. 1–11, 2020.

[70] F. Zhu, F. Gao, M. Yao, and H. Zou, "Joint information-and jamming-beamforming for physical layer security with full duplex base station," *IEEE Transactions on Signal Processing*, vol. 62, no. 24, pp. 6391–6401, 2014.

[71] B. K. Chalise, Q. Li, and W.-K. Ma, "Full-duplex secure relay beamforming design for systems with perfect and partial csi," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 6, pp. 5570–5584, 2019.

[72] J. Seo and J. H. Lee, "Energy beamforming for full-duplex wireless powered communication in presence of eavesdropper," in *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*. IEEE, 2017, pp. 1–5.

[73] W. Wu, B. Wang, Y. Zeng, H. Zhang, Z. Yang, and Z. Deng, "Robust secure beamforming for wireless powered full-duplex systems with self-energy recycling," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 11, pp. 10 055–10 069, 2017.

[74] Z. Kong, S. Yang, D. Wang, and L. Hanzo, "Robust beamforming and jamming for enhancing the physical layer security of full duplex radios," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 12, pp. 3151–3159, 2019.

[75] Q. Li, W.-K. Ma, and D. Han, "Sum secrecy rate maximization for full-duplex two-way relay networks using alamouti-based rank-two beamforming," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 8, pp. 1359–1374, 2016.

[76] Ö. Cepheli, S. Tedik, and G. K. Kurt, "A high data rate wireless communication system with improved secrecy: Full duplex beamforming," *IEEE communications letters*, vol. 18, no. 6, pp. 1075–1078, 2014.

[77] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, "What will 5g be?" *IEEE Journal on selected areas in communications*, vol. 32, no. 6, pp. 1065–1082, 2014.

[78] J. Zhu, R. Schober, and V. K. Bhargava, "Secure transmission in multicell massive mimo systems," *IEEE Transactions on Wireless Communications*, vol. 13, no. 9, pp. 4766–4781, 2014.

[79] ——, "Linear precoding of data and artificial noise in secure massive mimo systems," *IEEE Transactions on Wireless Communications*, vol. 15, no. 3, pp. 2245–2261, 2015.

[80] J. Wang, J. Lee, F. Wang, and T. Q. Quek, "Jamming-aided secure communication in massive mimo rician channels," *IEEE Transactions on Wireless Communications*, vol. 14, no. 12, pp. 6854–6868, 2015.

[81] J. Zhu, W. Xu, and N. Wang, "Secure massive mimo systems with limited rf chains," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 6, pp. 5455–5460, 2016.

[82] J. Zhu, D. W. K. Ng, N. Wang, R. Schober, and V. K. Bhargava, "Analysis and design of secure massive mimo systems in the presence of hardware impairments," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 2001–2016, 2017.

[83] B. Chen, C. Zhu, W. Li, J. Wei, V. C. Leung, and L. T. Yang, "Original symbol phase rotated secure transmission against powerful massive mimo eavesdropper," *IEEE Access*, vol. 4, pp. 3016–3025, 2016.

[84] B. Chen, C. Zhu, L. Shu, M. Su, J. Wei, V. C. Leung, and J. J. Rodrigues, "Securing uplink transmission for lightweight single-antenna ues in the presence of a massive mimo eavesdropper," *IEEE Access*, vol. 4, pp. 5374–5384, 2016.

[85] H. Wei, D. Wang, X. Hou, Y. Zhu, and J. Zhu, "Secrecy analysis for massive mimo systems with internal eavesdroppers," in *2015 IEEE 82nd vehicular technology conference (VTC2015-Fall)*. IEEE, 2015, pp. 1–5.

[86] Q. Shi, W. Xu, J. Wu, E. Song, and Y. Wang, "Secure beamforming for mimo broadcasting with wireless information and power transfer," *IEEE Transactions on Wireless Communications*, vol. 14, no. 5, pp. 2841–2853, 2015.

[87] A. Mukherjee and A. L. Swindlehurst, "Utility of beamforming strategies for secrecy in multiuser mimo wiretap channels," in *2009 47th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2009, pp. 1134–1141.

[88] K. Jayasinghe, P. Jayasinghe, N. Rajatheva, and M. Latva-Aho, "Secure beamforming design for physical layer network coding based mimo two-way relaying," *IEEE communications letters*, vol. 18, no. 7, pp. 1270–1273, 2014.

[89] A. Kaushik, E. Vlachos, C. Tsinos, J. Thompson, and S. Chatzinotas, "Joint bit allocation and hybrid beamforming optimization for energy efficient millimeter wave mimo systems," *IEEE Transactions on Green Communications and Networking*, vol. 5, no. 1, pp. 119–132, 2020.

[90] X. Tian, Z. Wang, H. Li, M. Li, and Z. Sun, "Secure hybrid beamforming with low-resolution phase shifters in mmwave mimo systems," in *2019 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2019, pp. 1–6.

[91] S. Zhao, J. Liu, X. Li, Y. Shen, and X. Jiang, "Secure beamforming for full-duplex mimo two-way communication via untrusted relaying," in *2017 IEEE Globecom Workshops (GC Wkshps)*. IEEE, 2017, pp. 1–6.

[92] N. Romero-Zurita, M. Ghogho, and D. McLernon, "Physical layer security of mimo–ofdm systems by beamforming and artificial noise generation," *Physical Communication*, vol. 4, no. 4, pp. 313–321, 2011.

[93] B. A. Fette, *Cognitive radio technology*. Elsevier, 2006.

[94] C.-I. Badoi, N. Prasad, V. Croitoru, and R. Prasad, "5g based on cognitive radio." *Wireless Personal Communications*, vol. 57, no. 3, 2011.

[95] D. H. Tashman and W. Hamouda, "An overview and future directions on physical-layer security for cognitive radio networks," *IEEE Network*, vol. 35, no. 3, pp. 205–211, 2020.

[96] Z. Lin, M. Lin, W.-P. Zhu, J.-B. Wang, and J. Cheng, "Robust secure beamforming for wireless powered cognitive satellite-terrestrial networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 7, no. 2, pp. 567–580, 2020.

[97] Z. Lin, M. Lin, B. Champagne, W.-P. Zhu, and N. Al-Dhahir, "Secure and energy efficient transmission for rsma-based cognitive satellite-terrestrial networks," *IEEE Wireless Communications Letters*, vol. 10, no. 2, pp. 251–255, 2020.

[98] F. Khoramnejad, M. Rasti, H. Pedram, and E. Hossain, "On resource management in load-coupled ofdma networks," *IEEE Transactions on Communications*, vol. 66, no. 5, pp. 2295–2311, 2018.

[99] F. Zhu and M. Yao, "Improving physical-layer security for crns using sinr-based cooperative beamforming," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 3, pp. 1835–1841, 2015.

[100] L. Jiang, H. Tian, C. Qin, S. Gjessing, and Y. Zhang, "Secure beamforming in wireless-powered cooperative cognitive radio networks," *IEEE Communications Letters*, vol. 20, no. 3, pp. 522–525, 2016.

[101] K. Tang, R. Shi, H. Shi, M. Z. A. Bhuiyan, and E. Luo, "Secure beamforming for cognitive cyber-physical systems based on cognitive radio with wireless energy harvesting," *Ad Hoc Networks*, vol. 81, pp. 174–182, 2018.

[102] H. S. M. Antony and T. Lakshmanan, "Secure beamforming in 5g-based cognitive radio network," *Symmetry*, vol. 11, no. 10, p. 1260, 2019.

[103] Y. Wu, X. Chen, and X. Chen, "Secure beamforming for cognitive radio networks with artificial noise," in *2015 International Conference on Wireless Communications & Signal Processing (WCSP)*. IEEE, 2015, pp. 1–5.

[104] Y. Pei, Y.-C. Liang, L. Zhang, K. C. Teh, and K. H. Li, "Secure communication over miso cognitive radio channels," *IEEE Transactions on Wireless Communications*, vol. 9, no. 4, pp. 1494–1502, 2010.

[105] B. Rankov and A. Wittneben, "Spectral efficient protocols for half-duplex fading relay channels," *IEEE Journal on selected Areas in Communications*, vol. 25, no. 2, pp. 379–389, 2007.

[106] I. Krikidis, H. A. Suraweera, P. J. Smith, and C. Yuen, "Full-duplex relay selection for amplify-and-forward cooperative networks," *IEEE Transactions on Wireless Communications*, vol. 11, no. 12, pp. 4381–4393, 2012.

[107] L. Dong, Z. Han, A. P. Petropulu, and H. V. Poor, "Improving wireless physical layer security via cooperating relays," *IEEE transactions on signal processing*, vol. 58, no. 3, pp. 1875–1888, 2009.

[108] J. Mo, M. Tao, Y. Liu, and R. Wang, "Secure beamforming for mimo two-way communications with an untrusted relay," *IEEE Transactions on Signal Processing*, vol. 62, no. 9, pp. 2185–2199, 2014.

[109] X. He and A. Yener, "End-to-end secure multi-hop communication with untrusted relays," *IEEE transactions on wireless communications*, vol. 12, no. 1, pp. 1–11, 2012.

[110] J. Huang, A. Mukherjee, and A. L. Swindlehurst, "Secure communication via an untrusted non-regenerative relay in fading channels," *IEEE Transactions on Signal Processing*, vol. 61, no. 10, pp. 2536–2550, 2013.

[111] R. Zhang and C. K. Ho, "Mimo broadcasting for simultaneous wireless information and power transfer," *IEEE Transactions on Wireless Communications*, vol. 12, no. 5, pp. 1989–2001, 2013.

[112] E. Boshkovska, D. W. K. Ng, N. Zlatanov, A. Koelpin, and R. Schober, "Robust resource allocation for mimo wireless powered communication networks based on a non-linear eh model," *IEEE Transactions on Communications*, vol. 65, no. 5, pp. 1984–1999, 2017.

[113] Z. Chu, Z. Zhu, M. Johnston, and S. Y. Le Goff, "Simultaneous wireless information power transfer for miso secrecy channel," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 9, pp. 6913–6925, 2015.

[114] Z. Zhu, Z. Chu, Z. Wang, and I. Lee, "Outage constrained robust beamforming for secure broadcasting systems with energy harvesting," *IEEE Transactions on Wireless Communications*, vol. 15, no. 11, pp. 7610–7620, 2016.

[115] M. R. Khandaker and K.-K. Wong, "Masked beamforming in the presence of energy-harvesting eavesdroppers," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 1, pp. 40–54, 2014.

[116] C. Zhang, H. Gao, T. Lv, Y. Lu, and X. Su, "Beamforming for secure two-way relay networks with physical layer network coding," in *2014 IEEE Global Communications Conference*. IEEE, 2014, pp. 1734–1739.

[117] J. Zhang and M. C. Gursoy, "Collaborative relay beamforming for secrecy," in *2010 IEEE International Conference on Communications*. IEEE, 2010, pp. 1–5.

[118] Y. Yang, Q. Li, W.-K. Ma, J. Ge, and P. Ching, "Cooperative secure beamforming for af relay networks with multiple eavesdroppers," *IEEE Signal Processing Letters*, vol. 20, no. 1, pp. 35–38, 2012.

[119] H.-M. Wang, Q. Yin, and X.-G. Xia, "Distributed beamforming for physical-layer security of two-way relay networks," *IEEE Transactions on Signal Processing*, vol. 60, no. 7, pp. 3532–3545, 2012.

[120] Q. Li and L. Yang, "Beamforming for cooperative secure transmission in cognitive two-way relay networks," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 130–143, 2019.

[121] C. Gu and C. Zhang, "Adaptive distributed beamforming and jamming in df relay networks for physical layer secrecy," in *2016 IEEE International Conference on Communication Systems (ICCS)*. IEEE, 2016, pp. 1–5.

[122] B. Zhu, J. Ge, Y. Huang, Y. Yang, and M. Lin, "Rank-two beamformed secure multicasting for wireless information and power transfer," *IEEE Signal Processing Letters*, vol. 21, no. 2, pp. 199–203, 2014.

[123] D. W. K. Ng, E. S. Lo, and R. Schober, "Robust beamforming for secure communication in systems with wireless information and power transfer," *IEEE Transactions on Wireless Communications*, vol. 13, no. 8, pp. 4599–4615, 2014.

[124] Y. Dong, A. El Shafie, M. J. Hossain, J. Cheng, N. Al-Dhahir, and V. C. Leung, "Secure beamforming in full-duplex swipt systems with loopback self-interference cancellation," in *2018 IEEE International conference on communications (ICC)*. IEEE, 2018, pp. 1–6.

[125] Z. Chu, T. A. Le, H. X. Nguyen, M. Karamanoglu, Z. Zhu, A. Nallanathan, E. Ever, and A. Yazici, "Robust design for miso swipt system with artificial noise and cooperative jamming," in *GLOBECOM 2017-2017 IEEE Global Communications Conference*. IEEE, 2017, pp. 1–6.

[126] Q. Li, Q. Zhang, and J. Qin, "Secure relay beamforming for simultaneous wireless information and power transfer in nonregenerative relay networks," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 5, pp. 2462–2467, 2014.

[127] Z. Zhu, Z. Chu, N. Wang, S. Huang, Z. Wang, and I. Lee, "Beamforming and power splitting designs for an-aided secure multi-user mimo swipt systems," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 12, pp. 2861–2874, 2017.

[128] W. Wang, K. C. Teh, and K. H. Li, "Artificial noise aided physical layer security in multi-antenna small-cell networks," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 6, pp. 1470–1482, 2017.

[129] T. Lv, H. Gao, and S. Yang, "Secrecy transmit beamforming for heterogeneous networks," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 6, pp. 1154–1170, 2015.

[130] H. Wu, X. Tao, N. Li, and J. Xu, "Secrecy outage probability in multi-rat heterogeneous networks," *IEEE Communications Letters*, vol. 20, no. 1, pp. 53–56, 2015.

[131] M. Xu, X. Tao, F. Yang, and H. Wu, "Enhancing secured coverage with comp transmission in heterogeneous cellular networks," *IEEE Communications Letters*, vol. 20, no. 11, pp. 2272–2275, 2016.

[132] M. Jahandideh, P. Azmi, N. Mokari, and M. Forouzesh, "Secure beamforming in relay-aided hetnet with interference nulling," 2019.

[133] B. Li, Z. Fei, and Z. Chu, "Optimal transmit beamforming for secure swipt in a two-tier hetnet," *IEEE Communications Letters*, vol. 21, no. 11, pp. 2476–2479, 2017.

[134] C. Li, X. Zhang, W. Lin, and S. Sun, "Secure transmission based on cooperative jamming relay in heterogeneous massive mimo system," in *2016 IEEE International Conference on Network Infrastructure and Digital Content (IC-NIDC)*. IEEE, 2016, pp. 268–272.

[135] A. Kumari, R. Gupta, S. Tanwar, and N. Kumar, "A taxonomy of blockchain-enabled softwarization for secure uav network," *Computer Communications*, vol. 161, pp. 304–323, 2020.

[136] B. Li, Z. Fei, C. Zhou, and Y. Zhang, "Physical-layer security in space information networks: A survey," *IEEE Internet of things journal*, vol. 7, no. 1, pp. 33–52, 2019.

[137] B. Li, Z. Fei, Y. Zhang, and M. Guizani, "Secure uav communication networks over 5g," *IEEE Wireless Communications*, vol. 26, no. 5, pp. 114–120, 2019.

[138] B. Li, Z. Fei, Z. Chu, F. Zhou, K.-K. Wong, and P. Xiao, "Robust chance-constrained secure transmission for cognitive satellite–terrestrial networks," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 5, pp. 4208–4219, 2018.

[139] U. M. Maurer, "Secret key agreement by public discussion from common information," *IEEE transactions on information theory*, vol. 39, no. 3, pp. 733–742, 1993.

[140] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography. i. secret sharing," *IEEE Transactions on Information Theory*, vol. 39, no. 4, pp. 1121–1132, 1993.

[141] R. Wilson, D. Tse, and R. A. Scholtz, "Channel identification: Secret sharing using reciprocity in ultrawideband channels," *IEEE Transactions on Information Forensics and Security*, vol. 2, no. 3, pp. 364–375, 2007.

[142] M. Wilhelm, I. Martinovic, and J. B. Schmitt, "Secure key generation in sensor networks based on frequency-selective channels," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 1779–1790, 2013.

[143] N. Patwari, J. Croft, S. Jana, and S. K. Kasera, "High-rate uncorrelated bit extraction for shared secret key generation from channel measurements," *IEEE Transactions on Mobile Computing*, vol. 9, no. 1, pp. 17–30, 2009.

[144] C. Zenger, H. Vogt, J. Zimmer, A. Sezgin, and C. Paar, "The passive eavesdropper affects my channel: Secret-key rates under real-world conditions," in *2016 IEEE Globecom Workshops (GC Wkshps)*. IEEE, 2016, pp. 1–6.

[145] L. Mohjazi, L. Bariah, S. Muhaidat, and M. A. Imran, "Performance of reconfigurable intelligent surfaces in the presence of generalized gaussian noise," *IEEE Communications Letters*, vol. 26, no. 4, pp. 773–777, 2022.

[146] Z. Lin, H. Niu, K. An, Y. Wang, G. Zheng, S. Chatzinotas, and Y. Hu, "Refracting ris-aided hybrid satellite-terrestrial relay networks: Joint beamforming design and optimization," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 58, no. 4, pp. 3717–3724, 2022.

[147] Z. Deng, Q. Li, Q. Zhang, L. Yang, and J. Qin, "Beamforming design for physical layer security in a two-way cognitive radio iot network with swipt," *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 10 786–10 798, 2019.

[148] T.-X. Zheng, H.-W. Liu, N. Zhang, Z. Ding, and V. C. M. Leung, "Secure content delivery in two-tier cache-enabled mmwave heterogeneous networks," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 1640–1654, 2021.

[149] M. M. Sande, S. Hamouda, and B. T. Maharaj, "Fast converging robust beamforming for massive mimo in heterogeneous networks," *IEEE Access*, vol. 6, pp. 23 918–23 928, 2018.

[150] P. Gu, C. Hua, W. Xu, R. Khatoun, Y. Wu, and A. Serhrouchni, "Control channel anti-jamming in vehicular networks via cooperative relay beamforming," *IEEE Internet Things J.*, vol. 7, no. 6, pp. 5064–5077, Feb. 2020.

[151] Z. Shen, K. Xu, X. Xia, W. Xie, and D. Zhang, "Spatial sparsity based secure transmission strategy for massive MIMO systems against simultaneous jamming and eavesdropping," *IEEE Trans. Inf. Forensics Secur.*, vol. 15, pp. 3760–3774, Jun. 2020.

[152] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave mimo systems," *IEEE transactions on wireless communications*, vol. 13, no. 3, pp. 1499–1513, 2014.

[153] W. Xia, G. Zheng, Y. Zhu, J. Zhang, J. Wang, and A. P. Petropulu, "A deep learning framework for optimization of miso downlink beamforming," *IEEE Transactions on Communications*, vol. 68, no. 3, pp. 1866–1880, 2019.

[154] H. Huang, Y. Peng, J. Yang, W. Xia, and G. Gui, "Fast beamforming design via deep learning," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 1, pp. 1065–1069, 2019.

[155] J. Kim, H. Lee, S.-E. Hong, and S.-H. Park, "Deep learning methods for universal miso beamforming," *IEEE Wireless Communications Letters*, vol. 9, no. 11, pp. 1894–1898, 2020.

[156] Q. Wang, K. Feng, X. Li, and S. Jin, "Precodernet: Hybrid beamforming for millimeter wave systems with deep reinforcement learning," *IEEE Wireless Communications Letters*, vol. 9, no. 10, pp. 1677–1681, 2020.

[157] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, 2015.

[158] X. Liu, Y. Liu, and Y. Chen, "Machine learning empowered trajectory and passive beamforming design in uav-ris wireless networks," *IEEE Journal on Selected Areas in Communications*, 2020.

[159] R. Negi and S. Goel, "Secret communication using artificial noise," in *IEEE vehicular technology conference*, vol. 62, no. 3. Citeseer, 2005, p. 1906.

[160] J. Fan and *et al.*, "A cooperative jamming based secure uplink transmission scheme for heterogeneous networks supporting d2d communications," in *Wireless Algorithms, Systems, and Applications", year="2018.* Cham": Springer International Publishing, pp. 103–114.

[161] S. Yan, Y. Shang, X. Zhang, D. Li, and X. Li, "An artificial noise scheme for secure communication in heterogeneous d2d and cellular networks," in *2016 IEEE 84th Vehicular Technology Conference (VTC-Fall)*, 2016, pp. 1–5.

[162] S. Gong, S. Ma, C. Xing, Y. Li, and L. Hanzo, "Multi-antenna aided secrecy beamforming optimization for wirelessly powered hetnets," *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5261–5277, May 2020.

[163] Y. Huo, X. Fan, L. Ma, X. Cheng, Z. Tian, and D. Chen, "Secure Communications in Tiered 5G Wireless Networks With Cooperative Jamming," *IEEE Transactions on Wireless Communications*, vol. 18, no. 6, pp. 3265–3280, June 2019.

[164] M. Cao, L. Wang, H. Xu, D. Chen, C. Lou, N. Zhang, Y. Zhu, and Z. Qin, "Sec-D2D: A Secure and Lightweight D2D Communication System With Multiple Sensors," *IEEE Access*, vol. 7, pp. 33 759–33 770, Mar. 2019.

[165] S. Yun, J.-M. Kang, I.-M. Kim, and J. Ha, "Deep artificial noise: Deep learning-based precoding optimization for artificial noise scheme," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 3, pp. 3465–3469, 2020.

[166] Y. Zhang, Z. Mou, F. Gao, J. Jiang, R. Ding, and Z. Han, "Uav-enabled secure communications by multi-agent deep reinforcement learning," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 11 599–11 611, 2020.

[167] Y. Lu, P. Cheng, Z. Chen, W. H. Mow, and Y. Li, "A learning approach to cooperative communication system design," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 5240–5244.

[168] Y. Lu, P. Cheng, Z. Chen, Y. Li, W. H. Mow, and B. Vucetic, "Deep autoencoder learning for relay-assisted cooperative communication systems," *IEEE Transactions on Communications*, vol. 68, no. 9, pp. 5471–5488, 2020.

[169] Y. Su, X. Lu, Y. Zhao, L. Huang, and X. Du, "Cooperative communications with relay selection based on deep reinforcement learning in wireless sensor networks," *IEEE Sensors Journal*, vol. 19, no. 20, pp. 9561–9569, 2019.

[170] F. Irrum, M. Ali, M. Naeem, A. Anpalagan, S. Qaisar, and F. Qamar, "D2d-enabled resource management in secrecy-ensured 5g and beyond heterogeneous networks," *Physical Communication*, vol. 45, p. 101275, 2021.

[171] H. Z. Khan, M. Ali, M. Naeem, I. Rashid, S. Mumtaz, A. A. Khan, and A. N. Akhtar, "Secure resource management in beyond 5g heterogeneous networks with decoupled access," *Ad Hoc Networks*, vol. 125, p. 102737, 2022.

[172] H. Z. K. *et al.*, "Secure resource management in beyond 5g heterogeneous networks with decoupled access," *Ad Hoc Networks*, vol. 125, p. 102737, 2022.

[173] L. Wang and X. Liu, "Secure cooperative communication scheme for vehicular heterogeneous networks," *Vehicular Communications*, vol. 11, pp. 46–56, 2018.

[174] D. Marabissi, L. Mucchi, and S. Morosi, "User-cell association for security and energy efficiency in ultra-dense heterogeneous networks," *Sensors*, vol. 21, no. 2, p. 508, 2021.

[175] Z. Dou, G. Si, Y. Lin, and M. Wang, "An adaptive resource allocation model with anti-jamming in IoT network," *IEEE Access*, vol. 7, pp. 93 250–93 258, Mar. 2019.

[176] T. T. Do, E. Björnson, E. G. Larsson, and S. M. Razavizadeh, "Jamming-resistant receivers for the massive MIMO uplink," *IEEE Trans. Inf. Forensics Secur.*, vol. 13, no. 1, pp. 210–223, Aug. 2017.

[177] L. Jia, Y. Xu, Y. Sun, S. Feng, L. Yu, and A. Anpalagan, "A multi-domain anti-jamming defense scheme in heterogeneous wireless networks," *IEEE Access*, vol. 6, pp. 40 177–40 188, Jun. 2018.

[178] W. Wang, K. C. Teh, S. Luo, and K. H. Li, "Physical layer security in heterogeneous networks with pilot attack: A stochastic geometry approach," *IEEE Transactions on Communications*, vol. 66, no. 12, pp. 6437–6449, 2018.

[179] F. Irrum, M. Ali, M. Naeem, A. Anpalagan, S. Qaisar, and F. Qamar, "D2d-enabled resource management in secrecy-ensured 5g and beyond heterogeneous networks," *Physical Communication*, vol. 45, p. 101275, 2021.

[180] X. Hu, B. Li, K. Huang, Z. Fei, and K.-K. Wong, "Secrecy energy efficiency in wireless powered heterogeneous networks: A distributed admm approach," *IEEE Access*, vol. 6, pp. 20 609–20 624, 2018.

[181] Y. Zhang, C. Kang, Y. Teng, S. Li, W. Zheng, and J. Fang, "Deep reinforcement learning framework for joint resource allocation in heterogeneous networks," in *2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall)*, 2019, pp. 1–6.

[182] I. AlQerm and B. Shihada, "Enhanced machine learning scheme for energy efficient resource allocation in 5g heterogeneous cloud radio access networks," in *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2017, pp. 1–7.

[183] F. Tang, Y. Zhou, and N. Kato, "Deep reinforcement learning for dynamic uplink/downlink resource allocation in high mobility 5g hetnet," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 12, pp. 2773–2782, 2020.

[184] H. Ding, F. Zhao, J. Tian, D. Li, and H. Zhang, "A deep reinforcement learning for user association and power control in heterogeneous networks," *Ad Hoc Networks*, vol. 102, p. 102069, 2020.

[185] J. Li, X. Zhang, J. Zhang, J. Wu, Q. Sun, and Y. Xie, "Deep reinforcement learning-based mobility-aware robust proactive resource allocation in heterogeneous networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 1, pp. 408–421, 2020.

[186] L. Zhang and Y.-C. Liang, "Deep reinforcement learning for multi-agent power control in heterogeneous networks," *IEEE Transactions on Wireless Communications*, vol. 20, no. 4, pp. 2551–2564, 2021.

[187] B. M. Roger *et al.*, "Game theory: analysis of conflict," *The President and Fellows of Harvard College, USA*, vol. 66, 1991.

[188] M. Ahmed, Y. Li, Z. Yinxiao, M. Sheraz, D. Xu, and D. Jin, "Secrecy ensured socially aware resource allocation in device-to-device communications underlaying hetnet," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 5, pp. 4933–4948, 2019.

[189] K. C. Lalropuia and V. Gupta, "A bayesian game model and network availability model for small cells under denial of service (dos) attack in 5g wireless communication network," *Wirel. Netw.*, vol. 26, no. 1, p. 557–572, jan 2020.

[190] Y. Gao, Y. Xiao, M. Wu, M. Xiao, and J. Shao, "Game theory-based anti-jamming strategies for frequency hopping wireless communications," *IEEE Trans. Wireless Commun.*, vol. 17, no. 8, pp. 5314–5326, Jun. 2018.

[191] Z. Shen, K. Xu, and X. Xia, "Beam-domain anti-jamming transmission for downlink massive MIMO systems: A Stackelberg game perspective," *IEEE Trans. Inf. Forensics Secur.*, vol. 16, pp. 2727–2742, Mar. 2021.

[192] S. t. Weerasinghe, "Deep learning based game-theoretical approach to evade jamming attacks," in *Decision and Game Theory for Security*. Cham: Springer International Publishing, 2018, pp. 386–397.

[193] D. G. *et al.*, "Game theory based privacy preserving approach for collaborative deep learning in iot," *CoRR*, vol. abs/2103.15245, 2021. [Online]. Available: https://arxiv.org/abs/2103.15245

[194] K. Chung, C. A. Kamhoua, K. A. Kwiat, Z. T. Kalbarczyk, and R. K. Iyer, "Game theory with learning for cyber security monitoring," in *2016 IEEE 17th International Symposium on High Assurance Systems Engineering (HASE)*, 2016, pp. 1–8.

[195] J. Khoury and M. Nassar, "A hybrid game theory and reinforcement learning approach for cyber-physical systems security," in *NOMS 2020 - 2020 IEEE/IFIP Network Operations and Management Symposium*, 2020, pp. 1–9.

[196] L. Feng, Z. Yang, Y. Yang, X. Que, and K. Zhang, "Smart mode selection using online reinforcement learning for vr broadband broadcasting in d2d assisted 5g hetnets," *IEEE Transactions on Broadcasting*, vol. 66, no. 2, pp. 600–611, 2020.

[197] A. E. Thangaraj, "Error-control coding for physical-layer secrecy," 2015.

[198] A. Thangaraj, S. Dihidar, A. R. Calderbank, S. W. McLaughlin, and J.-M. Merolla, "Applications of ldpc codes to the wiretap channel," *IEEE Transactions on Information Theory*, vol. 53, no. 8, pp. 2933–2945, 2007.

[199] V. Rathi, M. Andersson, R. Thobaben, J. Kliewer, and M. Skoglund, "Performance analysis and design of two edge-type ldpc codes for the bec wiretap channel," *IEEE transactions on information theory*, vol. 59, no. 2, pp. 1048–1064, 2012.

[200] A. Subramanian, A. Thangaraj, M. Bloch, and S. W. McLaughlin, "Strong secrecy on the binary erasure wiretap channel using large-girth ldpc codes," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 585–594, 2011.

[201] H. Mahdavifar and A. Vardy, "Achieving the secrecy capacity of wiretap channels using polar codes," *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 6428–6443, 2011.

[202] E. Hof and S. Shamai, "Secrecy-achieving polar-coding," in *2010 IEEE Information Theory Workshop*. IEEE, 2010, pp. 1–5.

[203] M. Andersson, V. Rathi, R. Thobaben, J. Kliewer, and M. Skoglund, "Nested polar codes for wiretap and relay channels," *IEEE Communications Letters*, vol. 14, no. 8, pp. 752–754, 2010.

[204] Y. Liang, H. V. Poor, S. Shamai *et al.*, "Information theoretic security," *Foundations and Trends® in Communications and Information Theory*, vol. 5, no. 4–5, pp. 355–580, 2009.

[205] J. M. Renes, R. Renner, and D. Sutter, "Efficient one-way secret-key agreement and private channel coding via polarization," in *International Conference on the Theory and Application of Cryptology and Information Security*. Springer, 2013, pp. 194–213.

[206] Y. Zhang, A. Liu, C. Gong, G. Yang, and S. Yang, "Polar-ldpc concatenated coding for the awgn wiretap channel," *IEEE Communications Letters*, vol. 18, no. 10, pp. 1683–1686, 2014.

[207] J.-C. Belfiore and F. Oggier, "Secrecy gain: A wiretap lattice code design," in *2010 International Symposium On Information Theory & Its Applications*. IEEE, 2010, pp. 174–178.

[208] F. Oggier, P. Solé, and J.-C. Belfiore, "Lattice codes for the wiretap gaussian channel: Construction and analysis," *IEEE Transactions on Information Theory*, vol. 62, no. 10, pp. 5690–5708, 2015.

[209] A.-M. Ernvall-Hytonen, "On a conjecture by belfiore and solé on some lattices," *IEEE Transactions on Information Theory*, vol. 58, no. 9, pp. 5950–5955, 2012.

[210] J.-C. Belfiore and F. Oggier, "Lattice code design for the rayleigh fading wiretap channel," in *2011 IEEE International Conference on Communications Workshops (ICC)*. IEEE, 2011, pp. 1–5.

[211] L.-C. Choo and C. Ling, "Superposition lattice coding for gaussian broadcast channel with confidential message," in *2014 IEEE Information Theory Workshop (ITW 2014)*. IEEE, 2014, pp. 311–315.

[212] X. He and A. Yener, "Providing secrecy with structured codes: Two-user gaussian channels," *IEEE Transactions on Information Theory*, vol. 60, no. 4, pp. 2121–2138, 2014.

[213] ——, "The gaussian many-to-one interference channel with confidential messages," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 2730–2745, 2011.

[214] ——, "Strong secrecy and reliable byzantine detection in the presence of an untrusted relay," *IEEE transactions on information theory*, vol. 59, no. 1, pp. 177–192, 2012.

[215] L. Lai, H. El Gamal, and H. V. Poor, "The wiretap channel with feedback: Encryption over the channel," *IEEE Transactions on Information Theory*, vol. 54, no. 11, pp. 5059–5067, 2008.

[216] L. H. Ozarow and A. D. Wyner, "Wire-tap channel ii," *AT&T Bell Laboratories technical journal*, vol. 63, no. 10, pp. 2135–2157, 1984.

[217] H. Jeong and P. Grover, "Energy-adaptive error correcting for dynamic and heterogeneous networks," *Proceedings of the IEEE*, vol. 107, no. 4, pp. 765–777, 2019.

[218] E. Arikan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Transactions on Information Theory*, vol. 55, no. 7, p. 3051–3073, Jul 2009. [Online]. Available: http://dx.doi.org/10.1109/TIT.2009.2021379

[219] Q. Du, Y. Xu, W. Li, and H. Song, "Security enhancement for multicast over internet of things by dynamically constructed fountain codes," *Wireless Communications and Mobile Computing*, vol. 2018, 2018.

[220] K. M. Rege, K. Balachandran, J. H. Kang, and K. Karakayali, "Interference mitigation in heterogeneous networks with simple dirty paper coding," *Wireless Networks*, vol. 26, no. 4, pp. 2755–2767, 2020.

[221] X. Yang, L. Zhang, and Z. Wu, "A unified convolutional neural network classifier aided intelligent channel decoder for coexistent heterogeneous networks," *IEEE Systems Journal*, 2021.

[222] F. Liang, C. Shen, and F. Wu, "An iterative bp-cnn architecture for channel decoding," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 1, pp. 144–159, 2018.

[223] M. Zhang, Q. Huang, S. Wang, and Z. Wang, "Construction of ldpc codes based on deep reinforcement learning," in *2018 10th International Conference on Wireless Communications and Signal Processing (WCSP)*, 2018, pp. 1–4.

[224] Y. Liao, S. A. Hashemi, J. Cioffi, and A. Goldsmith, "Construction of polar codes with reinforcement learning," in *GLOBECOM 2020-2020 IEEE Global Communications Conference*. IEEE, 2020, pp. 1–6.

[225] T. Wadayama and S. Takabe, "Deep learning-aided trainable projected gradient decoding for ldpc codes," in *2019 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2019, pp. 2444–2448.

[226] Y. Liao, S. A. Hashemi, J. M. Cioffi, and A. Goldsmith, "Construction of polar codes with reinforcement learning," *IEEE Transactions on Communications*, vol. 70, no. 1, pp. 185–198, 2021.

[227] S. Cammerer, T. Gruber, J. Hoydis, and S. ten Brink, "Scaling deep learning-based decoding of polar codes via partitioning," in *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, 2017, pp. 1–6.

[228] Y. Zhang, R. H. Deng, E. Bertino, and D. Zheng, "Robust and universal seamless handover authentication in 5g hetnets," *IEEE Transactions on Dependable and Secure Computing*, vol. 18, no. 2, pp. 858–874, 2021.

[229] A. K. *et al.*, "Design of a usim and ecc based handover authentication scheme for 5g-wlan heterogeneous networks," *Digital Communications and Networks*, vol. 6, no. 3, pp. 341–353, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352864817302821

[230] A. Kumar and H. Om, "Handover authentication scheme for device-to-device outband communication in 5g-wlan next generation heterogeneous networks," *Arabian Journal for Science and Engineering*, vol. 43, 04 2018.

[231] Z. Han, T. Lei, Z. Lu, X. Wen, W. Zheng, and L. Guo, "Artificial intelligence-based handoff management for dense wlans: A deep reinforcement learning approach," *IEEE Access*, vol. 7, pp. 31 688–31 701, 2019.

[232] M.-A.-F. Rihani, M. Mroue, J.-C. Prevotct, F. Nouvel, and Y. Mohanna, "A neural network based handover for multi-rat heterogeneous networks with learning agent," in *2018 13th International Symposium on Reconfigurable Communication-centric Systems-on-Chip (ReCoSoC)*, 2018, pp. 1–6.

[233] A. G. Mahira and M. S. Subhedar, "Handover decision in wireless heterogeneous networks based on feedforward artificial neural network," in *Computational Intelligence in Data Mining*, H. S. Behera and D. P. Mohapatra, Eds. Singapore: Springer Singapore, 2017, pp. 663–669.

[234] Y. Koda, K. Yamamoto, T. Nishio, and M. Morikura, "Reinforcement learning based predictive handover for pedestrian-aware mmwave networks," in *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 2018, pp. 692–697.

[235] T. t. Ma, "Securing 5g hetnets using mutual physical layer authentication," in *Proceedings of the 2019 7th International Conference on Information Technology: IoT and Smart City*, 2019, pp. 275–278.

[236] P. Zhang, "Physical layer authentication for wireless communications," 2020.

[237] J. Liu, X. Wang, and H. Tang, "Physical layer authentication enhancement using maximum snr ratio based cooperative af relaying," *Wireless Communications and Mobile Computing*, vol. 2017, 2017.

[238] S. Tomasin, "Analysis of channel-based user authentication by key-less and key-based approaches," *IEEE Transactions on Wireless Communications*, vol. 17, no. 9, pp. 5700–5712, 2018.

[239] S. Han, S. Xu, W. Meng, and C. Li, "Dense-device-enabled cooperative networks for efficient and secure transmission," *IEEE Network*, vol. 32, no. 2, pp. 100–106, 2018.

[240] H. Fang, X. Wang, and L. Hanzo, "Learning-aided physical layer authentication as an intelligent process," *IEEE Transactions on Communications*, vol. 67, no. 3, pp. 2260–2273, 2018.

[241] H. Fang, X. Wang, and S. Tomasin, "Machine learning for intelligent authentication in 5g and beyond wireless networks," *IEEE Wireless Communications*, vol. 26, no. 5, pp. 55–61, 2019.

[242] L. Senigagliesi, M. Baldi, and E. Gambi, "Physical layer authentication techniques based on machine learning with data compression," in *2020 IEEE Conference on Communications and Network Security (CNS)*, 2020, pp. 1–6.

[243] C. Shi, J. Liu, H. Liu, and Y. Chen, "Smart user authentication through actuation of daily activities leveraging wifi-enabled iot," in *Proceedings of the 18th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, ser. Mobihoc '17. New York, NY, USA: Association for Computing Machinery, 2017. [Online]. Available: https://doi.org/10.1145/3084041.3084061

[244] F. Pan, X. Li, H. Pu, Y. Guo, and J. Liu, "Physical layer authentication based on residual network for industrial wireless cpss," in *IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society*, 2020, pp. 4368–4373.

[245] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5g be?" *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, 2014.

[246] W. Tang, S. Feng, Y. Ding, and Y. Liu, "Physical layer security in heterogeneous networks with jammer selection and full-duplex users," *IEEE Transactions on Wireless Communications*, vol. 16, no. 12, pp. 7982–7995, 2017.

[247] Y. Xu, H. Xie, C. Liang, and F. R. Yu, "Robust secure energy efficiency optimization in swipt-aided heterogeneous networks with a non-linear energy harvesting model," *IEEE Internet of Things Journal*, 2021.

[248] Y. Jiang, Y. Zou, J. Ouyang, and J. Zhu, "Beamforming aided secrecy energy efficiency maximization in heterogeneous cellular networks," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 3, pp. 2576–2589, 2021.

[249] Fierce Wireless, *Qualcomm: 5G is all about sub 6 GHz and mmWave*, 2019. Accessed: Feb. 14, 2022. [Online]. Available: $https://www.fiercewireless.com/wireless/qualcomm-5g-all-about-sub-6-ghz-and-mmwave$

[250] A. Alkhateeb, Y.-H. Nam, M. S. Rahman, J. Zhang, and R. W. Heath, "Initial beam association in millimeter wave cellular systems: Analysis and design insights," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2807–2821, Mar. 2017.

[251] C. Liu, M. Li, S. V. Hanly, I. B. Collings, and P. Whiting, "Millimeter wave beam alignment: Large deviations analysis and design insights," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1619–1631, Apr. 2017.

[252] I. Budhiraja, N. Kumar, and S. Tyagi, "Deep-reinforcement-learning-based proportional fair scheduling control scheme for underlay D2D communication," *IEEE Internet Things J.*, vol. 8, no. 5, pp. 3143–3156, Mar. 2021.

[253] Z. Ding, R. Schober, and H. V. Poor, "No-pain no-gain: Drl assisted optimization in energy-constrained cr-noma networks," *IEEE Transactions on Communications*, vol. 69, no. 9, pp. 5917–5932, 2021.

[254] L. Zhang and Y.-C. Liang, "Deep reinforcement learning for multi-agent power control in heterogeneous networks," *IEEE Transactions on Wireless Communications*, vol. 20, no. 4, pp. 2551–2564, 2020.

[255] D. Lee, N. He, P. Kamalaruban, and V. Cevher, "Optimization for reinforcement learning: From a single agent to cooperative agents," *IEEE Signal Processing Magazine*, vol. 37, no. 3, pp. 123–135, 2020.

[256] Y. Du, L. Han, M. Fang, J. Liu, T. Dai, and D. Tao, "Liir: Learning individual intrinsic reward in multi-agent reinforcement learning," *Adv. Neural Inf. Process. Syst.*, vol. 32, 2019.

[257] N. Naderializadeh, J. J. Sydir, M. Simsek, and H. Nikopour, "Resource management in wireless networks via multi-agent deep reinforcement learning," *IEEE Transactions on Wireless Communications*, vol. 20, no. 6, pp. 3507–3523, 2021.

[258] V. François-Lavet, P. Henderson, R. Islam, M. G. Bellemare, and J. Pineau, "An introduction to deep reinforcement learning," *arXiv preprint arXiv:1811.12560*, 2018.

[259] Q. Wei, F. L. Lewis, Q. Sun, P. Yan, and R. Song, "Discrete-time deterministic Q-learning: A novel convergence analysis," *IEEE Trans. Cybern.*, vol. 47, no. 5, pp. 1224–1237, Apr. 2016.

[260] A. Kumar, A. Zhou, G. Tucker, and S. Levine, "Conservative Q-learning for offline reinforcement learning," *Adv. Neural Inf. Process. Syst.*, vol. 33, pp. 1179–1191, 2020.

[261] Z. Li, M. Chen, K. Wang, C. Pan, N. Huang, and Y. Hu, "Parallel deep reinforcement learning based online user association optimization in heterogeneous networks," in *Proc. IEEE Int. Conf. Commun. Wrkshps (ICC Workshops)*, Dublin, Ireland, 2020, pp. 1–6.

[262] Z. Wang, T. Schaul, M. Hessel, H. Hasselt, M. Lanctot, and N. Freitas, "Dueling network architectures for deep reinforcement learning," in *Proc. Int. Conf. Machine Learning*. PMLR, New York, NY, USA, 2016, pp. 1995–2003.

[263] H. Yang, A. Alphones, W.-D. Zhong, C. Chen, and X. Xie, "Learning-based energy-efficient resource management by heterogeneous rf/vlc for ultra-reliable low-latency industrial iot networks," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 8, pp. 5565–5576, 2020.

[264] J. W. Wallace and R. K. Sharma, "Automatic secret keys from reciprocal mimo wireless channels: Measurement and analysis," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 3, pp. 381–392, 2010.

[265] Z. Yu, Z. Ji, and P. L. Yeoh, "Secrecy rate comparison of key-based encrypted data transmission and keyless secure transmission," in *2021 IEEE Globecom Workshops (GC Wkshps)*, 2021, pp. 1–6.

[266] M. Lin, J. Ouyang, and W.-P. Zhu, "Joint beamforming and power control for device-to-device communications underlaying cellular networks," *IEEE J. Sel. Areas Commun*, vol. 34, no. 1, pp. 138–150, Jul. 2015.

[267] H. Yang, Z. Xiong, J. Zhao, D. Niyato, Q. Wu, H. V. Poor, and M. Tornatore, "Intelligent reflecting surface assisted anti-jamming communications: A fast reinforcement learning approach," *IEEE Trans. Wireless Commun.*, vol. 20, no. 3, pp. 1963–1974, Nov. 2020.

[268] L. Xiao, X. Lu, D. Xu, Y. Tang, L. Wang, and W. Zhuang, "UAV relay in VANETs against smart jamming with reinforcement learning," *IEEE Trans. Veh. Technol.*, vol. 67, no. 5, pp. 4087–4097, Jan. 2018.

[269] Z. Xiao, B. Gao, S. Liu, and L. Xiao, "Learning based power control for mmwave massive mimo against jamming," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Abu Dhabi, UAE, 2018, pp. 1–6.

[270] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5394–5409, Aug. 2019.

[271] Y. S. Nasir and D. Guo, "Multi-agent deep reinforcement learning for dynamic power allocation in wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 10, pp. 2239–2250, Aug. 2019.

[272] X. Wang, C. Wang, X. Li, V. C. M. Leung, and T. Taleb, "Federated deep reinforcement learning for internet of things with decentralized cooperative edge caching," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 9441–9455, Apr. 2020.

[273] N. Van Huynh, D. N. Nguyen, D. T. Hoang, and E. Dutkiewicz, ""Jam Me If You Can:" Defeating jammer with deep dueling neural network architecture and ambient backscattering augmented communications," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 11, pp. 2603–2620, Aug. 2019.

[274] D. W. K. Ng, E. S. Lo, and R. Schober, "Energy-efficient resource allocation in OFDMA systems with large numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3292–3304, Jul. 2012.