IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received November 5, 2020, accepted November 14, 2020, date of publication November 20, 2020,
date of current version December 7, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3039568

Reinforcement Learning Based Beamforming Jammer for Unknown Wireless Networks

GYUNGMIN KIM[1], (Graduate Student Member, IEEE), AND HYUK LIM[2], (Member, IEEE)
[1]School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju 61005, Republic of Korea
[2]Artificial Intelligence Graduate School, Gwangju Institute of Science and Technology, Gwangju 61005, Republic of Korea

Corresponding author: Hyuk Lim (hlim@gist.ac.kr)

This work was supported by the Electronic Warfare Research Center, Gwangju Institute of Science and Technology, funded by Defense Acquisition Program Administration and Agency for Defense Development, Republic of Korea.

ABSTRACT A jamming attack refers to adversarial activities to cause an interruption of communication among legitimate nodes in a wireless network by transmitting a jamming signal. Among various jamming techniques, network jamming is an attack technique that is performed to maximize the jamming impact on the entire network within the capability of the jammer. In this article, we consider a jamming attack to an unknown wireless network where no *a priori* network information is provided to the jammer except the radio frequency signal information acquired by overhearing the shared wireless channel. To increase the impact of a jamming attack on an unknown network, we propose a reinforcement learning based beamforming attack strategy. In the proposed attack, a jammer learns the beam direction and angle width to maximize the impact of jamming using the multi-armed bandit technique. As a reward for reinforcement learning, we develop a new metric that can quantitatively evaluate the impact of a jamming attack by measuring the statistical change of the channel busy times before and after each attack. Through extensive simulations, we evaluate the performance of the proposed jamming strategy in an unknown wireless network.

INDEX TERMS Jamming, beamforming, unknown networks, learning, multi armed bandits.

I. INTRODUCTION

Since the advent of wireless communication, there has been steady research on jamming attacks, which are attacks that interfere with the data transmission of legitimate nodes in a wireless network. Such research is necessary not only to be able to degrade the communication capabilities of hostile networks but also to develop countermeasures by improving the understanding of attack methods performed by the adversary. In modern times, with increasing reliance on wireless networks, research on jamming has become more crucial.

Studies on jamming attacks generally aim to improve jamming performance by increasing the success rate or energy efficiency of attacks [1]–[5]. Some jamming techniques, called network jamming attacks, aim to maximize the impact of jammers from the perspective of the entire network, such as by placing jammers in an optimal location for attack or by attacking nodes that play a major role in the network [6]–[12]. These network jamming attacks attempt to degrade the performance of as many links in the victim network as possible within the jammer's capability. Network jamming is an important problem in many applications, however, it has received relatively little attention in the literature. Specifically, network jamming is a key technique for increasing the impact of jammers, and it is especially important when an attacker must maximize performance with limited resources.

Conventionally, network jamming has been developed by using either optimization or game-theoretic principles [6]–[10]. A major disadvantage of these approaches, however, is that they assume a sufficient amount of *a priori* information about the victim network, such as the network topology, and location of nodes, which may not be available in practice. Recently, studies on jamming have employed machine learning techniques to improve the performance of jamming attacks in networks where sufficient information is not guaranteed for an attacker [12]–[14]. However, these studies assume that the attacker has some *a priori* information about the victim network, such as information on the number of nodes or the medium access control protocol of the victim network. Other studies assume that the jammer can selectively target specific nodes in the victim network [12]. To the best of our knowledge, there have been no studies on

The associate editor coordinating the review of this manuscript and approving it for publication was Ding Xu.

network jamming attacks on an unknown network where no *a priori* information is given to the attacker. Thus, a jamming strategy for improving the impact of jamming using only the radio frequency (RF) signal level information such as received waveform is required when jammers cannot obtain any *a priori* information about the victim network.

In this article, we propose a beamforming-based jamming strategy that can improve the performance of jammers that attack a victim network without any *a priori* information. The proposed jamming strategy uses beamforming transmission that steers the direction and angle width of jamming signals to improve jamming performance in an unknown network, where the jammer is deployed randomly in the network due to the lack of information about the victim network. By concentrating the jamming attack with a specific direction and angle width, the jammer can cause increased performance degradation of the victim network. The jamming strategy proposed in this article utilizes the statistical change in the observed channel busy times before and after an attack to evaluate the jamming impact according to the selected beam direction and angle width. Since the effectiveness of an attack is evaluated based only on the change in the observed channel busy times, the proposed jamming strategy can operate in unknown environments where no *a priori* information is provided. In addition, for the jammer to inflict the most damage on the victim network while measuring jamming performance indirectly through statistical values, we formalize the problem using a multi-armed bandit (MAB) framework, in which the jammer can select the beam direction and beam angle width.

The main contributions of this work are as follows:

- We propose a jamming strategy to improve jamming performance through a beamforming-based attack. With the proposed method, the performance degradation of the victim network can be increased even in the absence of *a priori* information.
- To evaluate the jamming impact on an unknown victim network, we utilize the statistical change in the observed channel busy times before and after each attack. We propose a novel metric that can evaluate the impact of jamming attacks indirectly based on the channel busy times.
- In the proposed jamming strategy, a jammer selects the direction and angle width of the beam using a MAB-based algorithm to maximize the impact of jamming on an unknown network.
- We evaluate the jamming performance of the proposed method using simulations in which a jammer performs a jamming attack on unknown networks with several node distributions.

The remainder of this article is organized as follows. In Section II, we present an overview of related work on network jamming attack. In Section III, we introduce the beamforming-based jamming scenario and the system model. In Section IV, we explain the advantage of the proposed beamforming-based jammer for network jamming, and, in Section V, we propose a jamming strategy for an unknown network where no *a priori* information is provided to the jammer. In Section VI, we present the performance evaluation of the proposed method, and in Section VII, we provide the conclusions of the study.

II. RELATED WORK

Commander *et al.* [6] proposed a strategic jammer placement algorithm to neutralize communication in a victim network. The authors presented a method for determining the location of jammers from a set of available discrete locations while minimizing the number of deployed jammers. Vadlamani *et al.* [7] solved the bi-level min-max jammer placement problem in which an attacker places jammers and determines the optimal channel hopping strategy to minimize the throughput of the victim network, while the defender attempts to maximize the throughput of the network by changing the channels to send data. The problem was modeled as a Nash equilibrium channel hopping game between a defender and an attacker. Using the concept of graph clustering, Feng *et al.* [8] examined how to efficiently place jammers to minimize the number of jammers required to disrupt a network by partitioning the network into a specified number of disconnected subnetworks, each of which was limited in size. They formulated the jammer placement problem as a binary integer linear program and solved this problem via a meta-heuristic algorithm. In [9], Gezici *et al.* proposed an optimal jammer placement method in a wireless localization system.

These jammer placement studies assume that the attacker has a sufficient amount of *a priori* information about the victim network, such as the network topology, the number and location of nodes in the victim network, and the signal-to-interference-plus-noise ratio threshold for the successful transmission of the victim nodes. However, such information may not be available in practice. To address the jammer placement problem in an unknown network, Commander *et al.* [10] derived upper and lower bounds for the optimal number of jammers required to compromise the functionality of a victim network when the jammers were located at the vertices of a uniform grid. Here, the jammers operated in the omnidirectional mode, and it was assumed that the victim network lay in a square range. However, the authors assumed that the attacker had information about the threshold for a successful attack based on the distance between the jammer and victim node. In addition, only the case in which jammers were deployed in some vertices of a uniform grid was considered. However, when an attacker does not have any information about the victim network, it is not guaranteed that the attacker can place a jammer at the optimal location. Therefore, it is necessary to devise a method to increase the impact of jamming on a network without sufficient information about the victim network.

Amuru *et al.* [11] investigated the impact of jammers that are randomly deployed according to a binomial point process.

Without information about the location of the victim nodes, the authors studied jamming against the wireless network from a physical layer perspective by employing tools from stochastic geometry. Sagduyu *et al.* [15] considered stochastic network traffic and evaluated the effects of traffic uncertainty on jamming attacks. In addition, recent studies have shown interest in optimizing the jamming performance using beamforming technique [16]–[19]. In [17], Karlsson *et al.* presented a method for jamming a time-division duplex point-to-point link by utilizing beamforming. In [18], Liu *et al.* designed the optimal jamming signal to achieve maximum degradation in signal-to-interference-plus-noise ratio of legitimate nodes. In [20], Li *et al.* proposed a directional reactive jamming scheme for eavesdropper detection.

Machine learning approaches have also recently been studied to improve jamming performance in an unknown network in which jammers cannot obtain sufficient information about the target network. Erpek *et al.* [13] proposed a deep-learning-based jammer that reliably predicts the subsequent successful transmissions without knowledge of the transmitter's algorithm. The jammer collects the channel status and ACKs to train the neural network and only jams if the successful transmission is predicted by the trained neural network. Amuru *et al.* [14] proposed an online learning algorithm to maximize the jamming efficacy using the MAB framework. The jammer learns optimal physical layer jamming strategies without any *a priori* information about the transmitter nodes. Here, the jammer adjusts the transmission power, signal duration, and modulation of the jamming signal according to the selected arm, and employs the presence or absence of an ACK packet as a reward for MAB learning. In addition, Amuru *et al.* [12] proposed a MAB-based blind network attack strategy for the case in which the topology of the victim network is unknown *a priori*. The authors presented an algorithm to learn the node that is most important to attack in a network to minimize the number of messages that are successfully exchanged. To calculate the reward, the jammer measures the fraction of the total number of flows stopped when each node is under attack. It is assumed that the jammer can selectively attack a node and receive ACK packets from the victim network.

In a victim network in which sufficient information is not guaranteed for the attacker, a method to increase the impact of jamming and a metric for evaluating jamming performance are key research topics. To address these topics, existing studies assume that the attacker has minimal information about the victim network [12]–[14]. However, in a hostile network in which an attacker cannot obtain any *a priori* information about the network, a new jamming strategy is required. In this article, we propose a beamforming-based jamming strategy that can improve the performance of a jammer that is deployed at an arbitrary location in a victim network. In addition, we propose a new metric that evaluates the jamming performance based on only the statistical changes in the observed channel busy times before and after an attack.

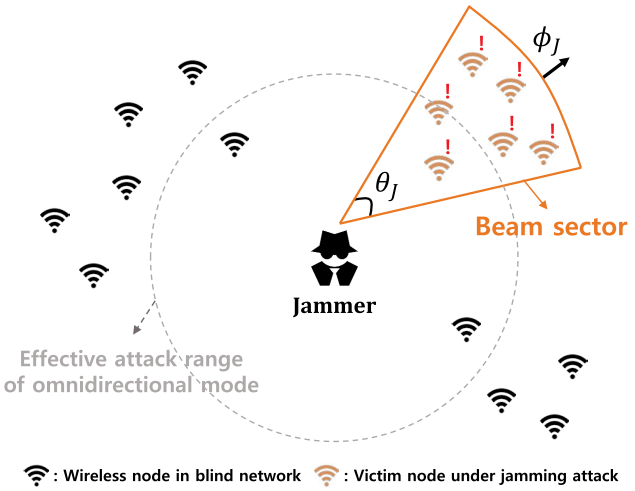

FIGURE 1. Beamforming jammer in an unknown network.

III. SYSTEM MODEL

We consider a jamming attack that exploits antenna-array beamforming technology to effectively attack a victim wireless network. Using the beamforming transmission rather than omnidirectional transmission, the jammer can concentrate jamming signals on the attacks in a specific direction and interfere with distant nodes. Here, it is assumed that nodes are placed at a fixed location and have a stationary transmission probability distribution in the victim network.

For beamforming transmission, we use a beam sector model, which specifies the antenna gain with respect to the antenna array beamforming patterns as depicted in Figure 1, where ϕ_J and θ_J denote the azimuth and angle width of the beam for jamming, respectively. Here, the main lobe gain $g_m^{\theta_J}$ is determined depending on θ_J, and $g_m^{\theta_J}$ gets bigger as θ_J gets narrower. It is assumed that the jammer has a limited power in practice [21]. In the beam sector model, the beam gain between the jammer and node located in the beam sector is constant $g_m^{\theta_J}$ for all angles within θ_J; otherwise, it is the sidelobe gain $g_s^{\theta_J}$. Under the assumption that the side lobe gain is sufficiently small, we ignore the sidelobe gain and let $g_s^{\theta_J}$ be zero. Therefore, the jammer can transmit/receive signals only to/from nodes in the beam sector. Note that the proposed jamming strategy is not limited to the beam sector model. In this article, the beam sector model is adopted to emphasize that an attack is concentrated in a specific area; however, the proposed jamming scheme can be applied to other beamforming models.

A. ATTACK MODEL

A jammer selects the angle width and azimuth of the beam to attack nodes in a specific beam sector using a beam-steering technique. Here, we assume that the jammer can select the angle width θ_i from the set $\mathcal{B}_\theta \in \{\theta_1, \cdots, \theta_{N_\theta}\}$, and the azimuth ϕ_j for the beam direction from the set $\mathcal{B}_\phi \in \{\phi_1, \cdots, \phi_{N_\phi}\}$, where N_θ and N_ϕ are the number of available angle widths and azimuths by the applied beam steering

technique, respectively. Let $\mathcal{S}_{\phi_j}^{\theta_i}$ denote the beam sector that is determined by θ_i and ϕ_j.

We adopt a reactive jamming method to perform a jamming attack that is energy-efficient with a low risk of being detected, which is achieved by transmitting disturbing signals only when communication activity is detected on a target channel [22]. Here, it is assumed that a reactive jammer transmits a jamming signal in a certain duration, which is long enough to cause a reception failure of victim nodes. Because the beamforming-based jammer can only receive signals from nodes located in the beam sector $\mathcal{S}_{\phi_j}^{\theta_i}$ and only perform attacks on the corresponding areas, both the target transmitter and receiver node pair should be located in $\mathcal{S}_{\phi_j}^{\theta_i}$ for a successful attack. If only the transmitter node is located in the beam sector, the attack signal will not reach the receiver node, and if only the receiver node is located in the beam sector, the jammer will not be able to detect the transmitted signal and will not perform the attack.

Let Γ_r denote the signal-to-noise (SNR) threshold for the jammer to detect the transmitted signal from a legitimate node. Then, the jammer can detect communication activity to perform a reactive attack when the SNR of a signal from the victim node is greater than Γ_r. To determine the success of a jamming attack on legitimate receiver nodes, we consider a signal-to-jamming-plus-noise ratio (SJNR) model. The receiver node fails to receive a transmitted signal from its associated transmitter node when the SJNR of the received signal is lower than a certain threshold depending on the transmission rate. We assume that the signal attenuation in the victim network follows a log-distance path loss model as follows:

$$\mathrm{PL}(d)[dB] = \mathrm{PL}(d_0) + 10\gamma \log_{10}(\frac{d}{d_0}) + X_n, \tag{1}$$

where $\mathrm{PL}(d_0) = 20\log_{10}(\frac{4\pi}{\lambda_c})$ denotes the path loss at the reference distance $d_0 = 1$, λ_c denotes the wavelength of light in meters, γ denotes the path loss exponent, and X_n denotes a Gaussian random variable with zero mean and standard deviation σ in the dB scale.

Let n_α and n_β be a pair of legitimate transmitter and receiver nodes in the victim network. Here, the two nodes are separated from each other by Euclidean distance $d_{\alpha,\beta}$, and the signal between the two nodes is denoted as s_α^β. Using the path loss model, the received signal strength (RSS) at n_β for the signal from n_α can be obtained as

$$P_r^{\alpha,\beta} = \kappa P_t^\alpha d_{\alpha,\beta}^{-\gamma}, \tag{2}$$

where P_t^α is the transmission power of n_α, and $\kappa = 10^{-\mathrm{PL}(d_0)/10}$ is a scaling factor. Similarly, the RSS at n_β for the beamforming signal from the jammer can be obtained as

$$P_r^{J,\beta} = \kappa g_m^{\theta_B} P_t^J d_{J,\beta}^{-\gamma} f_{\mathcal{S}}(s_\alpha^\beta), \tag{3}$$

where P_t^J is the transmission power of the jammer, $d_{J,\beta}$ is the Euclidean distance between the jammer and n_β, and $f_{\mathcal{S}}(s_\alpha^\beta)$ is a binary beam sector function for signal s_α^β. Here, $f_{\mathcal{S}}(s_\alpha^\beta) = 1$ if both n_α and n_β are located in $\mathcal{S}_{\phi_j}^{\theta_i}$; otherwise, $f_{\mathcal{S}}(s_\alpha^\beta) = 0$. Consequently, the SJNR of the received signal at n_β, denoted SJNR_β, is given by

$$\mathrm{SJNR}_m(s_\alpha^\beta) = \frac{\kappa P_t^\alpha d_{\alpha,\beta}^{-\gamma}}{\kappa g_m^{\theta_J} P_t^J d_{J,\beta}^{-\gamma} f_{\mathcal{S}}(s_\alpha^\beta) + N_0 W}, \tag{4}$$

where N_0 is the noise power spectral density, and W is the channel bandwidth. Here, for simplicity, we ignore interference signals that are transmitted from other nodes at the same time. Let Γ_J denote the SJNR threshold for a successful jamming attack on n_β. Then, the jammer can successfully attack n_β when $\mathrm{SJNR}_m(s_\alpha^\beta)$ is lower than Γ_J (i.e. $\mathrm{SJNR}_m(s_\alpha^\beta) < \Gamma_J$).

B. VICTIM NETWORK MODEL

We consider an unknown victim network in which the jammer is restricted from obtaining *a priori* information about the network, such as the number of nodes, or the pair of transmitter and receiver nodes. Therefore, the assumption about the victim network is not strong except for the following. Generally, wireless communication protocols have the capability to change the transmission parameters such as the transmit power and modulation technique according to the channel condition for stable data transmission and high channel use efficiency. In this article, we assume that nodes in the victim network change the transmission parameters in the event of transmission failure. For example, a node that adopts the adaptive rate control algorithm, such as the automatic rate fallback (ARF) algorithm, can lower the transmission rate after packet transmission failures [23]. Note that if the jammer cannot perceive any change in the victims or receive any explicit feedback after an attack, it is not possible to develop a method to evaluate the jamming impact on the victims. Thus, we have made that a jammer can detect at least RF-level signal changes because a victim changes transmission-related parameters to achieve more reliable communication when its transmission is deteriorated or interfered with. In an unknown network, the RF level signal waveform on the target channel band is the only information available to the jammer, and it is difficult for the jammer to obtain additional information from the waveform.

IV. NETWORK JAMMING ATTACK BASED ON BEAMFORMING TRANSMISSION

A. NONUNIFORMITY OF TRAFFIC DISTRIBUTION IN A WIRELESS NETWORK

In general, nodes in a wireless network are geographically unevenly distributed rather than uniformly distributed. Particularly, if a network has a clustered topology and multi-hop topology, nodes are concentrated around relay or gateway nodes due to the limited transmission range, and the nonuniformity of the node distribution increases. Because the amount of traffic normally increases as the number of nodes increases, nonuniformity of the node distribution leads to nonuniformity of the traffic amount. That is, the amount of traffic varies depending on the location. Moreover, each node

has a different role in the network, and a different amount of data to transmit, thereby further increasing the traffic nonuniformity. Consequently, the amount of traffic affected by the jamming attack depends on the geographical area attacked by the jammer, and this in turn greatly affects the impact of jamming on the victim network. Therefore, to increase the impact of jamming, the jamming attack should be performed to a location such that areas with the highest traffic volume are located within the attack range of the jammer. If such a fine-tuned attack is performed, the jammer can inflict the greatest damage on the victim network when the attack area contains the most traffic.

B. JAMMING ATTACK BASED ON BEAMFORMING TRANSMISSION

In unknown networks without *a priori* information, it is not guaranteed that a jammer can be placed in an optimal geographical location for an attack. Thus, if an omnidirectional transmission-based jammer is placed at an arbitrary location in an unknown network, the jammer will have difficulty increasing its impact on the network by means other than increasing the transmission power. However, due to the limited transmission power of a jammer, increasing the transmission power is not a suitable solution for the efficient jamming attack.

In this article, we exploit beamforming transmission for performing a fine-tuned jamming attack on an unknown victim network. With beamforming transmission, a jammer can inflict greater damage to more victims on the network than when operating in omnidirectional mode with the same transmission power. For example, as depicted in Figure 1, suppose that a jammer is likely to be placed in a random position in an unknown network. Here, the beamforming-based jammer can improve jamming performance by steering the phase array to a target area where the jammer can attack more traffic at a given location. In addition, because the attack signal can be transmitted with a higher intensity to receiver nodes through beamforming transmission, the success rate of jamming can also be increased. Furthermore, the chance that the jamming attack is detected by the victim network is reduced because the nodes located outside of the beam sector cannot receive the jamming signal from the attacker.

V. JAMMING STRATEGY IN AN UNKNOWN NETWORK

A. PROBLEM OF MEASURING JAMMING IMPACT ON AN UNKNOWN NETWORK

It is important for the beamforming-based jammer to select the appropriate beam sector to increase the impact of jamming because the number of victims and their traffic intensity is different depending on which beam sector is attacked by the jammer. To select the location at which to attack, the jammer must be able to compare the jamming impacts when each beam sector is attacked. The jamming impact can be evaluated by the jamming performance, and existing jamming studies mostly use the number of dropped packets or

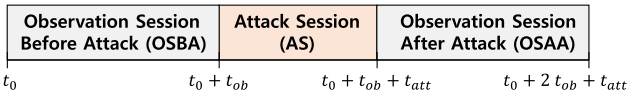

FIGURE 2. Sessions for the jamming strategy.

the transmission failure rate due to the attack to evaluate the jamming performance. To measure the performance, jammers require feedback on whether the receiver node successfully receives data transmitted from the transmitter node. If the victim network adopts a DATA/ACK handshake protocol for reliable data transmission, ACK packets from the transmitter node can be exploited to verify the successful or unsuccessful data transmission. For example, if an ACK packet is not detected after an attack is inflicted on a target data signal, the attack is considered successful, and the evaluated transmission failure rate increases. However, in an unknown network in which *a priori* information, such as the protocol of the victim network, is not provided, the jammer cannot reliably distinguish ACK packets from other signals, and the reception of ACK packets cannot be used as an indicator of success of data transmissions in the victim network.

A jammer that operates in an unknown network has difficulty in evaluating jamming performance in a conventional manner due to a lack of information. In such an unknown network, it is assumed to be impossible to decode captured signals and to obtain an explicit indicator to confirm the impact of jamming, and thus any protocol-dependent quantity cannot be used to define a jamming performance evaluation metric. Even in an unknown network, RF level signal waveform is the available information that the jammer can obtain in the field. Therefore, it is essential for a beamforming-based jammer to be able to measure and compare the performance of an attack on each beam sector using only the acquired RF waveforms to select the most effective beam sector to attack in an unknown network.

B. JAMMING STRATEGY IN AN UNKNOWN NETWORK

In this article, we exploit the change in the observed channel busy times before and after an attack to evaluate the performance of the attack on each beam sector. Here, we expect that if the jammer performs an effective attack on the victim network, there will be a change in the observed channel busy times after the attack, and the discrepancy of busy time patterns before and after the attacks will increase as the jamming affects more traffic transmissions of the victim. To obtain the channel busy times before and after the attack, the jammer divides the performance evaluation process into three sessions, as illustrated in Figure 2. The jammer attacks the beam sector during the attack session (AS), which is in the middle of the evaluation process, and observes the channel without jamming and collects the busy times during the observation session before the attack (OSBA) and observation session after the attack (OSAA). The jammer attacks the target beam sector for the duration of t_{att} in the AS and collects the busy times from the same beam sector for the duration of

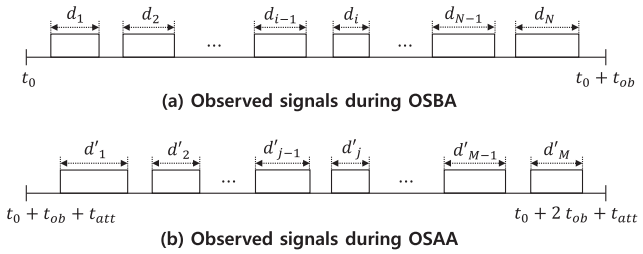

FIGURE 3. Observed channel busy times during observation sessions.

t_{ob} in the OSBA and OSAA. We set t_{ob} and t_{att} to be large enough values to obtain the statistical characteristics of the channel busy times and to cause the victim nodes to change the transmission parameters, respectively. In addition, it is assumed that $2t_{ob} + t_{att}$ is within the coherence time of the channel.

During the OSBA and OSAA, the jammer receives the RF level signal waveforms of the channel and compares the SNR of the waveforms with Γ_r, that is, the SNR threshold, to detect the transmitted signal from legitimate nodes. Here, the observed signals represent the times at which the channel is busy. The channel is considered busy if the SNR of the received waveform is greater than Γ_r at that time. Figure 3 illustrates the observed signals during the OSBA and OSAA. First, in Figure 3(a), d_i denotes the duration of the observed signal (DOS) in the OSBA. Let $\mathcal{D} = \{d_1, d_2, \cdots, d_N\}$ denote the set of the DOS in the OSBA, where N is the number of observed signals in the OSBA. Second, in Figure 3(b), d'_j denotes the DOS in the OSAA. Let $\mathcal{D}' = \{d'_1, d'_2, \cdots, d'_M\}$ denote the set of the DOS in the OSAA, where M is the number of observed signals in the OSAA.

To measure the change in the observed signals before and after the attack, we use the two sets $\mathcal{D}$ and $\mathcal{D}'$. We generate two vectors, $V_{\mathcal{D}}$ and $V_{\mathcal{D}'}$ that correspond to $\mathcal{D}$ and $\mathcal{D}'$, respectively, to measure the extent to which the attack has affected the beam sector based on the similarity between the two vectors. We generate the vectors by sorting the elements of $\mathcal{D}$ and $\mathcal{D}'$ in descending order. Because there is randomness in a wireless network, such as transmission order and data size, the order of the signals is not particularly important, as it may differ depending on the observation time instance. Instead of generating $V_{\mathcal{D}}$ and $V_{\mathcal{D}'}$ in order of appearance, we generate them in order of signal duration length. For example, if the longest duration among $\mathcal{D}$ is increased after a jamming attack, this may imply that the victim node starts to perform an anti-jamming method, such as the reduction of the transmission rate.

In this article, we evaluate the impact of a jamming attack on the beam sector using the similarity between $V_{\mathcal{D}}$ and $V_{\mathcal{D}'}$. If the attack is effective, there is a large change in the channel busy time patterns before and after the attack; thus, the similarity value is small. We define the jamming impact as follows:

$$\mathcal{J} \;=\; (1 - S_C(V_{\mathcal{D}}, V_{\mathcal{D}'})) \times \sqrt{\frac{\sum_{i=1}^{N} d_i}{t_{ob}}}. \qquad (5)$$

In (5), $S_C(\cdot)$ denotes the cosine similarity function. We employ this function to measure the change in the received signals before and after the attack and compare the degree of change between attacks on each beam sector. In addition, $\sqrt{\sum_{i=1}^{N} d_i / t_{ob}}$ is used as a weight factor to give higher priority to beam sectors with higher traffic intensity. Using the proposed metric in (5), the jammer can produce high jamming impact values when attacking a beam sector in which high traffic is generated and low similarity is measured from the observed signals before and after the attack. Consequently, the jammer can identify the best beam sector in which it can degrade the network performance the most. The attack impact is evaluated every $2t_{ob} + t_{att}$. When N and M, the number of elements of $\mathcal{D}$ and $\mathcal{D}'$, are not equal, then zero padding is used to fit a vector with a smaller size to a larger vector after sorting in descending order.

It should be noted that if nodes of the victim network perform beamforming transmissions, the jammer may less chance to receive the victim's signals and to interfere with the victim's transmission to its receiver node. As a result, the jamming performance could deteriorate, but it still works to perform the jamming on the victim.

C. ONLINE LEARNING JAMMING ALGORITHM

In this study, the jammer measures the impact of jamming when attacking each beam sector using (5). However, due to the randomness in a wireless network, such as the transmission probability of nodes, order of transmitted signals, and channel condition, the jamming impact is measured randomly from a probability distribution specific to each beam sector, and the variation in measurement may increase because the jamming impact is indirectly measured in an unknown network. In addition, it is difficult for a jammer operating in an unknown network to obtain or define information about the state of the environment. Therefore, it is necessary to utilize an online learning algorithm that can improve the jamming impact gradually without information about the state of the environment.

To address the above problem, we propose an online learning jamming algorithm based on the MAB framework for the beamforming jammer that searches for the optimal direction and angle width of the beam to maximize the jamming impact in an unknown network. In the proposed algorithm, the jammer selects its beam direction and angle width and steers its antenna array phases according to the selection. A joint selection of the beam direction and angle width is referred to as an arm. Thus, the arms of the MAB are defined by the pair $\{\mathcal{B}_\phi, \mathcal{B}_\theta\}$. Let $\mathcal{A}$ denote an arm set consisting of $N_\phi \times N_\theta$ elements that belong to $\mathcal{B}_\phi \times \mathcal{B}_\theta$ where $\times$ represents the Cartesian product. Then, the jammer selects an arm $a_s \in \{1, \ldots, N_\phi \times N_\theta\}$ for the s-th action to attack according to the policy of the MAB algorithm. Here, we utilize the UCB1 algorithm which achieves expected logarithmic regret uniformly over time, for all reward distributions, without requiring prior knowledge of the reward distribution [24].

Algorithm 1 Proposed MAB-Based Beam Jamming Algorithm

// Initialization
$\bar{\mathcal{J}}(k), n(k) \leftarrow 0\ , \forall\ k \in \{1, \ldots, N_\phi \times N_\theta\}$
for $n = 1, \ldots, N_\phi \times N_\theta$ **do**
    Select arm $a_s = n$
    Collect signals d_i from $\mathcal{S}_{a_s}$ during OSBA
    Perform jamming attack on $\mathcal{S}_{a_s}$ during AS
    Collect signals d_j' from $\mathcal{S}_{a_s}$ during OSAA
    Calculate the reward $\mathcal{J}_s(k)$ using (5)
    Update $\bar{\mathcal{J}}(k)$, $n(k)$, $c(k)$ based on policy of UCB1 [24]
end for

// Main Loop
while 1 **do**
    Select arm

$$a_s = \underset{1 \le k \le N_\phi \times N_\theta}{\operatorname{argmax}} \ \bar{\mathcal{J}}(k) + c(k)$$

    Steer antenna array in accordance with a_s
    Collect signals d_i from $\mathcal{S}_{a_s}$ during OSBA
    Perform jamming attack on $\mathcal{S}_{a_s}$ during AS
    Collect signals d_j' from $\mathcal{S}_{a_s}$ during OSAA
    Calculate the reward $\mathcal{J}_s(k)$ using (5)
    Update $\bar{\mathcal{J}}(k)$, $n(k)$, $c(k)$ based on policy of UCB1 [24]
end while

At the start time of the s-th action, the jammer selects an arm $k \in \mathcal{A}$ to attack, and calculates the instantaneous reward $\mathcal{J}_s(k)$ in accordance with arm k using (5). The jammer learns a reward in accordance with each arm over time in order to identify the arm that maximizes the jamming impact on the victim network. Here, if the jammer targets multiple channels, a reward is determined by averaging the values calculated in each channel using (5).

The proposed MAB-based jamming algorithm is presented in Algorithm 1. First, the jammer initializes the UCB1 algorithm with $\mathcal{A}$ by selecting each arm once, and updates $\bar{\mathcal{J}}(k)$, $n(k)$, and $c(k)$ accordingly, where $\bar{\mathcal{J}}(k)$ denotes the empirical mean, $n(k)$ denotes the number of times that arm k is selected, and $c(k)$ is a padding function. A standard choice is $c(k) = \mathcal{B}\sqrt{2 \ln n / n(k)}$, where $\mathcal{B}$ is an upper bound on the rewards, and n is the number of taken actions. Here, we set $c(k)$ to $\sqrt{2 \ln n / n(k)}$ in Algorithm 1. Thereafter, the jammer selects arm a_s based on the policy of UCB1. For each action, the jammer steers the antenna array in accordance with a_s, and the target beam sector $\mathcal{S}_{a_s}$ is then determined. Next, the jammer measures the jamming impact $\mathcal{J}_s$ using the evaluation method described in Section V-B. Namely, the jammer collects signals d_i and d_j' from $\mathcal{S}_{a_s}$ during the OSBA and OSAA, respectively, and performs beam jamming attack on $\mathcal{S}_{a_s}$ during the AS. Then, the jamming impact $\mathcal{J}_s$ is calculated using (5). Lastly, $\bar{\mathcal{J}}(k)$, $n(k)$, and $c(k)$ are updated based on the policy of UCB1. Note that we do not claim optimal performance, as it is not possible to guarantee optimality when no information is provided to the attacker. Because the jammer only has limited observables that provide indirect feedback about its attack performance, it is not guaranteed that the measured reward will be the largest when attacking the beam sector with the optimal beam direction and angle width that maximally degrade the network performance.

The complexity of the proposed beam jamming scheme is given by $\mathcal{O}(nT)$ for T iterations because the number of arms is n so the complexity of updating $\bar{\mathcal{J}}(k)$, $n(k)$, and $c(k)$ per iteration is given by $\mathcal{O}(n)$ where n is $N_\phi \times N_\theta$, and the computational complexity of calculating the reward is $\mathcal{O}(1)$ per arm per iteration. Here, as N_ϕ or N_θ increases, the computational complexity increases linearly with the memory requirement of $\mathcal{O}(n)$.

D. COUNTERMEASURES

In a wireless network, since transmitted signals are received not only from a receiver node but also from an adversary, even when the communication protocol or encryption scheme of the network is not revealed, the impact of jamming can be improved by using only received waveform at the jammer as in the proposed method. Therefore, it is necessary for the nodes of the network to take into account an adversary attempting to exploit the waveform including the transmitted signals. One possible approach is to use artificial noise to make it difficult for an adversary to evaluate jamming performance. By continuously propagating the artificial noise, which is mutually agreed among the nodes of the network, on the channel, it is possible to disguise the channel to keep appearing busy and make it hard to judge whether jamming is successful. In addition, it is also possible to reduce network performance degradation by creating a fake high-traffic zone to drive attacks elsewhere. An alternative short-term solution is to use beamforming communication to reduce the probability that the network traffic information is exposed to adversaries through the directional transmission.

VI. PERFORMANCE EVALUATION

In this section, we present the results of a simulation designed to evaluate the performance of the proposed jamming strategy in an unknown wireless network. In the simulation, the beamforming-based jammer learned the beam direction and angle width in order to determine the critical area in which the jamming signal could cause the greatest degradation of network performance with limited transmission power. Because the jammer has no *a priori* information about the victim network, we considered the situation in which an attacker deploys the jammer at a random location within the victim network. To validate the learning performance, we compared the results of the proposed method with optimal results that were obtained through a method of minimizing the throughput of the victim network by beamforming-based jamming when the jammer had complete information about the victim network. For comparison, we consider two naive

TABLE 1. Simulation parameters.

Parameter	Value
Transmission power of nodes	12 dBm
Size of network	$300 \times 300\ m^2$
Carrier frequency	2.4 GHz
Channel bandwidth	20 MHz
Path loss exponent	2.5
Noise power spectral density	-90 dBm/MHz
Duration of an iteration	30 ms
Attack session ratio in one action	0.3
Number of Monte Carlo simulations	$10,000$

attack methods using beamforming transmissions. A random beamforming attack is to select one of the beam sectors randomly. A traffic intensity-based attack is to attack the beam sector where the channel utilization rate is the highest. The traffic intensity-based method is a greedy method that can be taken using acquired RF waveforms without any *a priori* information about the network. We have compared the results of the proposed method with those of the naive methods. We set the number of beam directions to 36 (10° for the beam direction interval). In addition, the set of angle widths for the beam was $\{\pi/6, \pi/3, 2\pi/3, 2\pi\}$ in radians, and the set of beam gains for the corresponding angle width was $\{10, 6, 2, 0\}$ in dBi. Here, the jammer performed MAB learning using only the RF level signal waveform from the beam sector and performed reactive jamming on the beam sector determined according to the selected arm. To determine whether the channel is busy, we adopt the threshold-based signal detection method used in the clear channel assessment mechanism. For example, the signal detect threshold is around 4 dB SNR for IEEE 802.11 radios [25]. In our simulation, the SNR threshold Γ_r for signal detection was set to 5 dB, and the received waveform was compared with Γ_r every 1μs.

We constructed victim network with nonuniform node distribution. Here, it is assumed that nodes are arranged in multiple clusters within the network. In the simulation, we set the nodes in the victim network to adopt the ARF algorithm for signal transmission. Using this algorithm, the transmitter nodes adjusted the transmission rate according to the channel condition that was obtained based on the success or failure of the packet transmission, or the estimated SNR [23]. ARF is an algorithm for adjusting the signal transmission rate according to the channel conditions in a wireless network, and most wireless networks adopt adaptive rate control algorithms for dynamically adjusting the transmission rate to increase channel usage efficiency. In the victim network, a packet was dropped if the SJNR of the received signal at the receiving node was less than the SNR threshold for successful reception of the transmitted data rate. The set of data rates of the nodes in the victim network was $\{6, 12, 24, 48\}$ in Mbps and the set of SNR thresholds for the corresponding data rates was $\{2, 5, 11, 18\}$ in dB. The simulation was implemented in MATLAB, and the simulation parameters are listed in Table 1.

A. INSIDE JAMMER CASE

First, we evaluated the jamming performance in the case where the jammer is placed inside a victim network and clusters of nodes were arranged in a victim network based on a uniform distribution. We considered the scenario in which one jammer and multiple clusters of nodes were placed at random locations over an area of $300 \times 300\ m^2$, and nodes in each cluster were uniformly distributed within a cluster with a radius of 20 m.

Figure 4 illustrates the jamming performance against the number of iterations when the number of clusters was 5 and the transmission power of the jammer was 20 dBm. Here, each cluster contained 10 nodes. Figure 4(a), (b), and (c) present the average success rates of the jamming attack, the average number of dropped packets, and the throughput per cluster, respectively. Figure 4(a) indicates that the average success rate of the proposed method increased as the online learning progressed, and the proposed method achieved a higher success rate than the traffic intensity-based method and omnidirectional method. This is because the nodes of the victim network adjusted the transmission method before and after the attack when the attack was successful. In the ARF-based network, the average duration of the collected signals during the OSAA was higher than that of the OSBA when a transmission failure occurred due to a successful attack. Here, the more successful the attack on the beam sector was, the greater the rate of change was. Using the proposed metric, the jammer obtains a larger reward from the beam sector in which the jammer can perform a more successful attack, and converges to perform an attack on the corresponding beam sector. Through the proposed jamming strategy, the jammer can improve the energy efficiency of jamming and reduce the risk of being detected due to attacks with a lower success rate.

Figure 4(b) demonstrates that the average number of dropped packets due to the proposed jamming attack gradually increased and converged. Here, as the jammer learned the beam sector to attack using not only the rate of change before and after the attack but also the channel usage rate of the beam sector, the number of dropped packets increased. In Figure 4(c), the average throughput per cluster in the victim network of the proposed method decreased as online learning progressed, and the proposed method had a larger impact on the victim network than other methods. Figure 5 compares the learning performance of the proposed method with respect to the optimal strategy in the inside jammer case. It should be noted that the proposed method quickly outperformed the conventional omnidirectional method even though it required time to converge. Furthermore, Figure 4 demonstrates that the performance of the proposed method was close to the optimal method although the reward was calculated through indirect performance evaluation with limited observables.

Figure 6 presents the jamming performance against the number of clusters in the victim network when the transmission power of the jammer was 20 dBm.

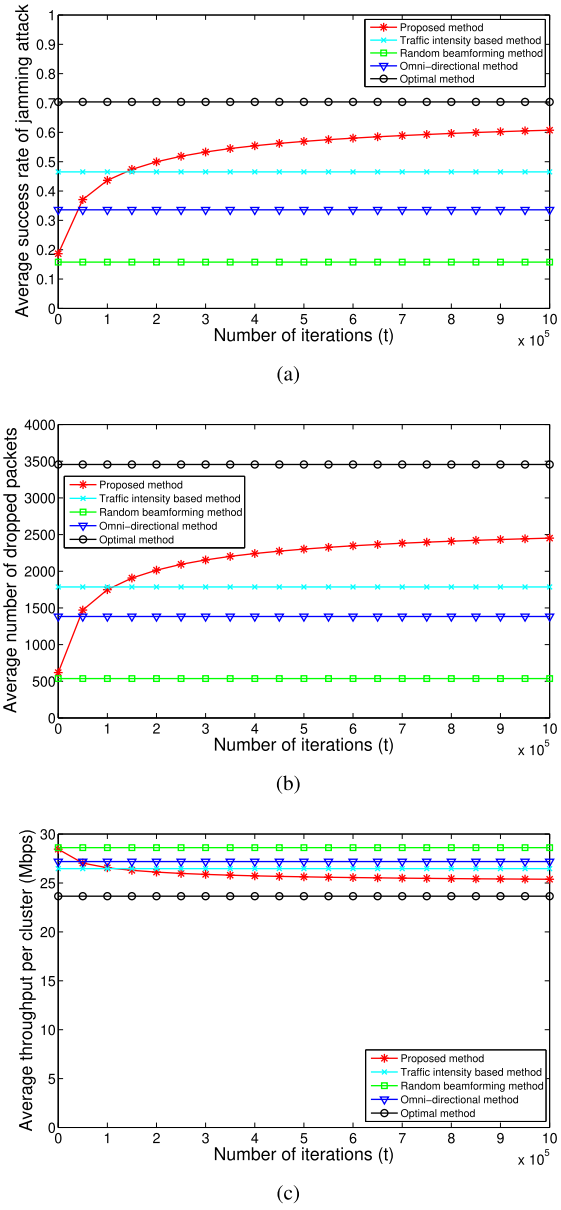

FIGURE 4. The jamming performance in the inside jammer case against number of iterations (the number of clusters is 5 and transmission power of jammer is 20 dBm). (a) Average success rate of jamming attack. (b) Average dropped packets in AS. (c) Average throughput of victim network in AS.

Figure 6(a), (b), and (c) present the average success rates of a jamming attack, the average number of dropped packets, and the throughput per cluster, respectively. Figure 6(a) reveals that the performance of the proposed method was higher than that of the omnidirectional method and was similar to that of the optimal method. Here, the performance gap between the two methods was evident when the number of clusters was small. However, as the number of clusters gradually increased, the gap between the two methods decreased as the performance of the omnidirectional method gradually increased. As the number of clusters increased, the probability of a jammer being placed around the cluster

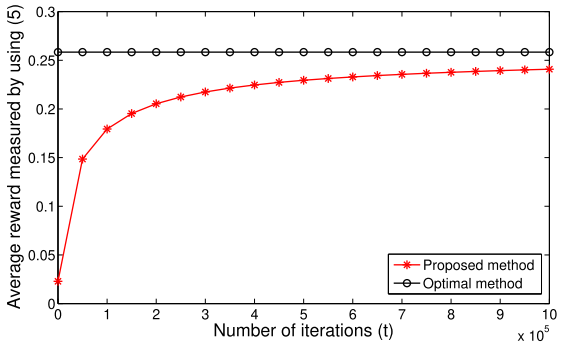

FIGURE 5. Average reward in the inside jammer case against number of iterations.

increased as well. Thus, the average distance between the jammer and nodes decreased, and the success rate of the omnidirectional method increased. In addition, as illustrated in Figure 6(b), the omnidirectional method outperformed the proposed method as the number of clusters increased even if the jamming success rate was lower. This is because the attack range of a jamming signal with the same intensity is wider in the omnidirectional method than in the beamforming method.

The beamforming method can propagate the signal more strongly and farther away; however, the total geographical area in which a signal of the same intensity reaches is reduced as the angle width of the beam decreases. Accordingly, when the node density of the network increases, the attack is performed on more transmissions through the omnidirectional method. As a result, more victim nodes are affected by the jamming attack even when the success rate is low, resulting in a situation in which the total number of dropped packets is similar to or greater than that of the proposed method in a dense network. Likewise, when the node density in the network is increased, the network throughput performance by the omnidirectional method deteriorated, as illustrated in Figure 6(c). However, it should be noted that the situation in which a jammer can be placed inside a dense hostile network rarely happens in practice. Therefore, we focused on developing a method to increase the jamming impact even when the jammer is placed outside the center of the network.

B. OUTSIDE JAMMER CASE

In the second part of the simulation, we evaluated the jamming performance when clusters of nodes were intensively deployed in part of the network, creating a hot spot, and a jammer was placed outside the hot spot area. We considered the scenario in which multiple clusters were placed within one quadrant of the network and a jammer was randomly placed outside the quadrant. Nodes in each cluster were uniformly distributed within a cluster with a radius of 20 m.

Figure 7 presents the jamming performance against the number of iterations when the number of clusters was 5 and the transmission power of the jammer was 20 dBm. Here, each cluster contained 10 nodes. Figure 7(a), (b), and (c) present the average success rates of the jamming attack,

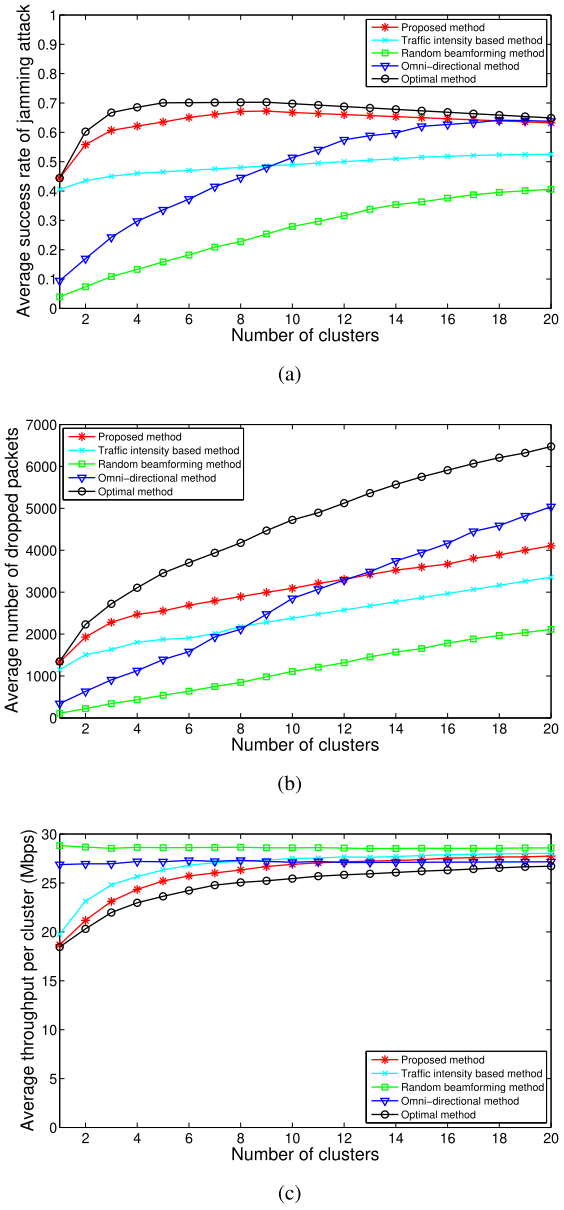

FIGURE 6. The jamming performance in the inside jammer case against number of clusters (transmission power of jammer is 20 dBm). (a) Average success rate of jamming attack. (b) Average dropped packets in AS. (c) Average throughput of victim network in AS.

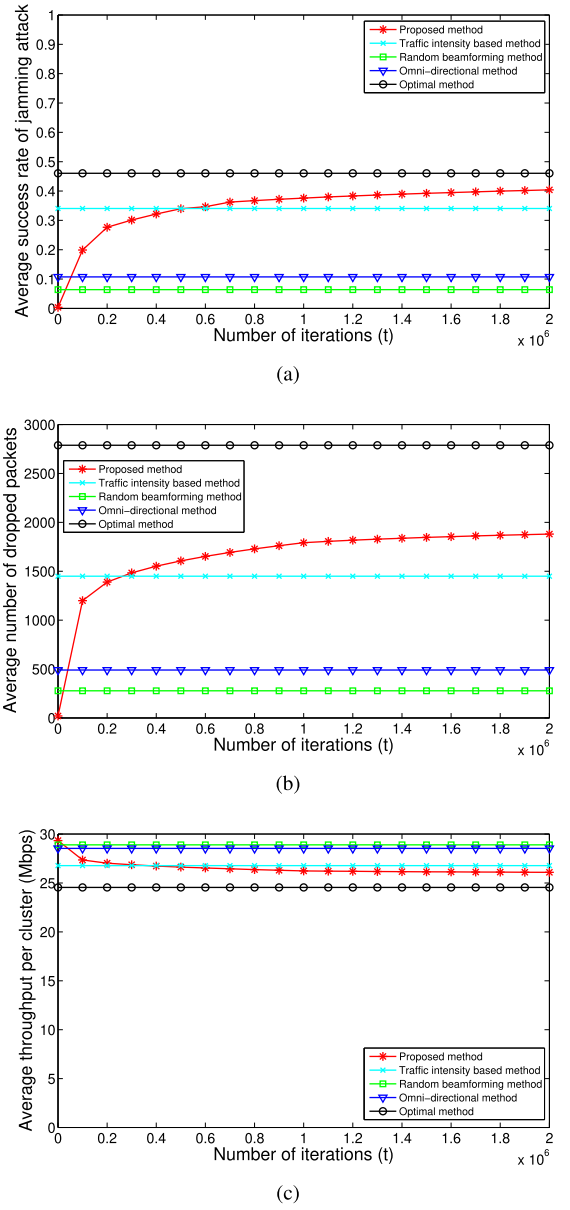

FIGURE 7. The jamming performance in the outside jammer case against number of iterations (the number of clusters is 5 and transmission power of jammer is 20 dBm). (a) Average success rate of jamming attack. (b) Average dropped packets in AS. (c) Average throughput of victim network in AS.

the average number of dropped packets, and the throughput per cluster, respectively. In Figure 7(a), as in Figure 4(a), the average success rate of the proposed method gradually increased as the online learning progressed, and the proposed method achieved better performance than the other methods. Likewise, as in Figure 4(b) and (c), the average number of dropped packets due to the proposed jamming attack gradually increased in Figure 7(b), and the average throughput per cluster in the victim network of the proposed method gradually decreased, as illustrated in Figure 7(c). Figure 8 compares the learning performance of the proposed method with respect to the optimal strategy in the outside jammer case. It should be noted that the overall performance was lower than in the inside jammer case described in Section VI-A. This is because, in the inside jammer case, the jammer was placed inside the area in which nodes were distributed, whereas in the outside jammer case, the jammer was placed outside this area. As a result, the average distance between the jammer and victim nodes was higher in the outside jammer case than in the inside jammer case.

For a successful attack, the received jamming signal strength at the victim node must be sufficiently strong; therefore, as the distance between the victim node and the jammer increases, the attack success rate decreases. Therefore, the jamming performance was lower in the outside jammer

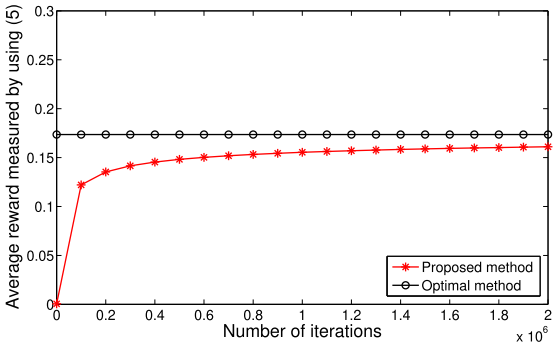

FIGURE 8. Average reward in the outside jammer case against number of iterations.

case, in which the distance between the victim node and jammer was relatively large. Figure 7 also indicates that the gap between the jamming impact of the proposed method and the omnidirectional method widened. For example, the performance of the proposed method in Figure 4(a) was approximately two times higher than that of the omnidirectional method; however, Figure 7(a) indicates that it was approximately four times higher. This is because the beamforming method reduces the jamming performance degradation by reducing the angle width of the beam and increasing the received jamming signal strength at the victim node for a successful attack when the distance between the jammer and victim node increases.

Figure 9 presents the jamming performance against the number of clusters in the victim network when the transmission power of the jammer was 20 dBm. Figure 9(a), (b), and (c) present the average success rates of the jamming attack, the average number of dropped packets, and the throughput per cluster, respectively. Unlike Figure 6(a), Figure 9(a) indicates that the performance of the omnidirectional method slightly increased when the number of clusters increased, and the proposed method exhibited much higher performance regardless of the number of clusters. In Figure 9(b) and (c), the proposed method also exhibited higher performance than the omnidirectional method regardless of the number of clusters, unlike in Figure 6(b) and (c). This is because the average distance between the jammer and victim node did not decrease even when the number of clusters increased, as the jammer was not placed inside the hot spot. Because the beamforming method can concentrate the jamming signal on the hot spot, the proposed method had better performance than the omnidirectional method.

It should be noted that situations in which a jammer is placed at the center of the hot spot are rare when a jammer is deployed in a large hostile network, and situations in which jamming attacks must be performed outside the center of the network are more common in the real world. Therefore, when a jammer is placed in an unknown network environment in the absence of *a priori* information about the victim network, jamming performance can be improved in most common situations through the proposed beamforming-based network jamming strategy.

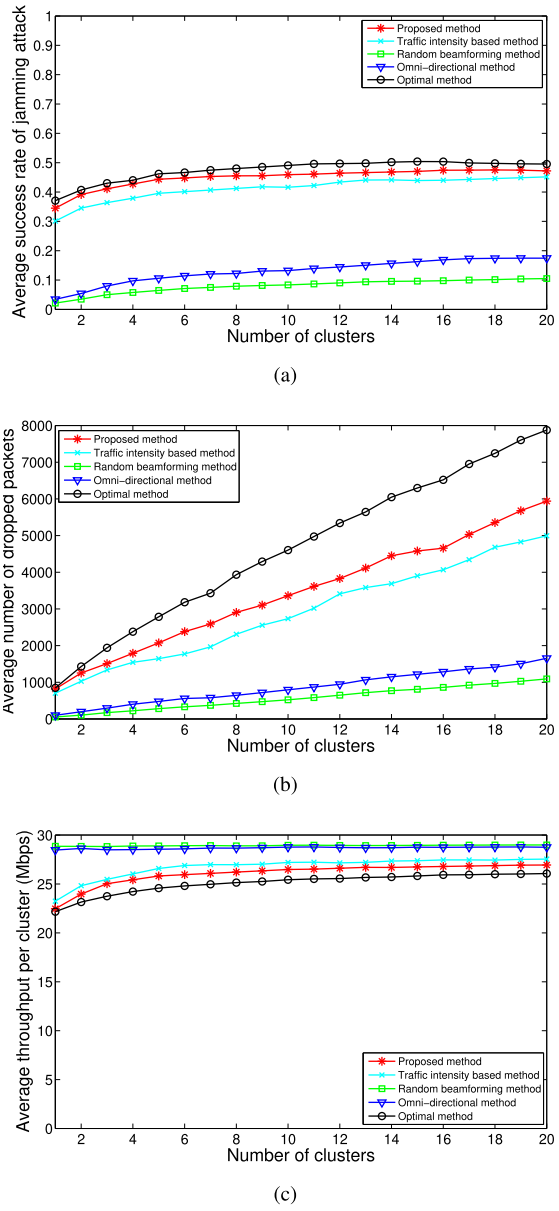

FIGURE 9. The jamming performance in the outside jammer case against number of clusters (transmission power of jammer is 20 dBm). (a) Average success rate of jamming attack. (b) Average dropped packets in AS. (c) Average throughput of victim network in AS.

Table 2 and 3 are numerical summary tables of simulation results for the inside jammer case and the outside jammer case to quantitatively display the results of the proposed jamming method. In both cases, it is shown that the proposed method has better performance than that of the traffic intensity-based method. This is because the traffic intensity-based method cannot perform an attack to increase the jamming success rate even though the traffic intensity-based method can target the area with the largest number of nodes through channel usage observation. On the other hand, the proposed method can further increase the jamming impact by taking into account whether the attack is successful or not together with traffic intensity. In addition, it should be noted that although

TABLE 2. Simulation results in the inside jammer case.

Attack method	Number of clusters in victim network											
	Success rate (%)				Dropped packets				Network throughput (Mbps)			
	5	10	15	20	5	10	15	20	5	10	15	20
Proposed method	63.5	66.8	65.0	63.2	2552	3092	3600	4105	25.2	26.9	27.4	27.8
Traffic intensity-based method	46.5	49.0	51.5	52.5	1876	2379	2869	3359	26.3	27.5	27.8	28.0
Random beamforming method	15.8	27.9	36.3	40.6	540	1109	1657	2114	28.6	28.6	28.5	28.6
Omni-directional method	33.6	51.4	62.0	63.8	1392	2850	3945	5039	27.2	27.1	27.1	27.2
Optimal method	70.0	69.8	67.3	64.9	3458	4723	5751	6475	23.6	25.4	26.2	26.7

TABLE 3. Simulation results in the outside jammer case.

Attack method	Number of clusters in victim network											
	Success rate (%)				Dropped packets				Network throughput (Mbps)			
	5	10	15	20	5	10	15	20	5	10	15	20
Proposed method	44.4	45.9	47.0	47.2	2073	3361	4579	5938	25.8	26.5	26.7	26.9
Traffic intensity-based method	39.6	41.6	43.9	45.2	1644	2733	3902	4993	26.6	27.2	27.4	27.5
Random beamforming method	6.5	8.3	9.6	10.5	279	522	808	1091	28.9	29.0	29.0	29.0
Omni-directional method	10.6	13.2	16.3	17.5	477	793	1216	1654	28.6	28.7	28.8	28.8
Optimal method	46.2	49.1	50.4	50.0	2783	4603	6296	7877	24.6	25.4	25.8	26.0

the success rate of the proposed method is higher than that of the omnidirectional method, the dropped packets of the omnidirectional method is higher than that of the proposed method when the number of clusters is 15 in Table 2. This is because the omnidirectional method can attack relatively more packets in a high node density network, causing more dropped packets despite a lower success rate. Thus, it is essential to take into account not only the attack success rate but also the traffic intensity of the target area. From the above two results, we can confirm the importance of combining traffic intensity and attack success rate as in the proposed method.

VII. CONCLUSION

In this article, we considered a beamforming-based jammer that operates in an unknown network in the absence of *a priori* information about the victim network. The proposed jamming strategy seeks the optimal beam direction and angle width to increase the impact of jamming on the victim network. To measure the jamming performance without precise feedback on the attack or *a priori* information about the victim network, we proposed a new metric for evaluating the impact of a jamming attack using changes in the received waveforms. Using the proposed metric, a jammer can obtain and compare the jamming performance of attacks on each beam sector. In addition, we proposed an online learning-based jamming strategy to find a beam sector with which the jammer can improve jamming performance through beamforming transmission in an unknown network. The results of simulations demonstrated that the proposed jamming strategy can improve jamming performance using only the received waveforms in an unknown victim network.

Our future research will investigate the development of a jamming strategy that learns the beamforming parameters even for an unknown network where moving nodes exist and the network characteristics change over time. Here, we plan to develop an algorithm for mobile network cases using MAB algorithms such as discounted UCB or sliding window UCB while adjusting the weight factor and observation time to take into account a dynamic movement of nodes in the network. In addition, we will study a method that can adaptively adjust the durations of sessions for jamming strategy in accordance with the response of the victim network.

REFERENCES

[1] K. Pelechrinis, M. Iliofotou, and S. V. Krishnamurthy, "Denial of service attacks in wireless networks: The case of jammers," *IEEE Commun. Surveys Tuts.*, vol. 13, no. 2, pp. 245–257, 2nd Quart., 2011.
[2] M. Lichtman, J. D. Poston, S. Amuru, C. Shahriar, T. C. Clancy, R. M. Buehrer, and J. H. Reed, "A communications jamming taxonomy," *IEEE Secur. Privacy*, vol. 14, no. 1, pp. 47–54, Jan. 2016.
[3] V.-L. Nguyen, P.-C. Lin, and R.-H. Hwang, "Energy depletion attacks in low power wireless networks," *IEEE Access*, vol. 7, pp. 51915–51932, 2019.
[4] Y. Zou, J. Zhu, X. Wang, and L. Hanzo, "A survey on wireless security: Technical challenges, recent advances, and future trends," *Proc. IEEE*, vol. 104, no. 9, pp. 1727–1765, Sep. 2016.
[5] K. Grover, A. Lim, and Q. Yang, "Jamming and anti-jamming techniques in wireless networks: A survey," *Int. J. Ad Hoc Ubiquitous Comput.*, vol. 17, no. 4, pp. 197–215, Dec. 2014.
[6] C. W. Commander, P. M. Pardalos, V. Ryabchenko, S. Uryasev, and G. Zrazhevsky, "The wireless network jamming problem," *J. Combinat. Optim.*, vol. 14, no. 4, pp. 481–498, Sep. 2007.
[7] S. Vadlamani, H. Medal, B. Eksioglu, and P. Li, "A bi-level programming model for the wireless network jamming placement problem," in *Proc. Inst. Ind. Syst. Eng. (IISE)*, 2014, p. 1003.
[8] J. Feng, E. L. Pasiliao, W. E. Dixon, and J. M. Shea, "An optimal jamming strategy to partition a wireless network," in *Proc. IEEE Mil. Commun. Conf. (MILCOM)*, Oct. 2015, pp. 978–984.
[9] S. Gezici, S. Bayram, M. N. Kurt, and M. R. Gholami, "Optimal jammer placement in wireless localization systems," *IEEE Trans. Signal Process.*, vol. 64, no. 17, pp. 4534–4549, Sep. 2016.
[10] C. W. Commander, P. M. Pardalos, V. Ryabchenko, O. Shylo, S. Uryasev, and G. Zrazhevsky, "Jamming communication networks under complete uncertainty," *Optim. Lett.*, vol. 2, no. 1, pp. 53–70, Oct. 2007.
[11] S. Amuru, H. S. Dhillon, and R. M. Buehrer, "On jamming against wireless networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 412–428, Jan. 2017.
[12] S. Amuru, R. M. Buehrer, and M. van der Schaar, "Bandit strategies for blindly attacking networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–6.

[13] T. Erpek, Y. E. Sagduyu, and Y. Shi, "Deep learning for launching and mitigating wireless jamming attacks," *IEEE Trans. Cognit. Commun. Netw.*, vol. 5, no. 1, pp. 2–14, Mar. 2019.
[14] S. Amuru, C. Tekin, M. van der Schaar, and R. M. Buehrer, "Jamming bandits—A novel learning method for optimal jamming," *IEEE Trans. Wireless Commun.*, vol. 15, no. 4, pp. 2792–2808, Apr. 2016.
[15] Y. E. Sagduyu, R. A. Berryt, and A. Ephremidesi, "Wireless jamming attacks under dynamic traffic uncertainty," in *Proc. Int. Symp. Modeling Optim. Mobile, Ad Hoc, Wireless Netw.*, 2010, pp. 303–312.
[16] M. A. Maleki Sadr, M. Ahmadian-Attari, R. Amiri, and V. V. Sabegh, "Worst-case jamming attack and optimum defense strategy in cooperative relay networks," *IEEE Control Syst. Lett.*, vol. 3, no. 1, pp. 7–12, Jan. 2019.
[17] M. Karlsson, E. Bjornson, and E. G. Larsson, "Jamming a TDD Point-to-Point link using reciprocity-based MIMO," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 12, pp. 2957–2970, Dec. 2017.
[18] Q. Liu, M. Li, X. Kong, and N. Zhao, "Disrupting MIMO communications with optimal jamming signal design," *IEEE Trans. Wireless Commun.*, vol. 14, no. 10, pp. 5313–5325, Oct. 2015.
[19] S. Sodagari and T. C. Clancy, "Efficient jamming attacks on MIMO channels," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2012, pp. 852–856.
[20] R. Li, Q. Duan, J. Xue, S. Zhang, and C. He, "A directional reactive jamming scheme based on machine learning," in *Proc. 11th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Oct. 2019, pp. 1–5.
[21] K. Cumanan, H. Xing, P. Xu, G. Zheng, X. Dai, A. Nallanathan, Z. Ding, and G. K. Karagiannidis, "Physical layer security jamming: Theoretical limits and practical designs in wireless networks," *IEEE Access*, vol. 5, pp. 3603–3611, 2017.
[22] M. Wilhelm, I. Martinovic, J. B. Schmitt, and V. Lenders, "Short paper: Reactive jamming in wireless networks: How realistic is the threat?" in *Proc. 4th ACM Conf. Wireless Netw. Secur. (WiSec)*, 2011, pp. 47–52.
[23] Ad. Kamerman, and L. Monteban, "WaveLAN-II: A high-performance wireless LAN for the unlicensed band," *Bell Labs Tech. J.*, vol. 2, no. 3, pp. 118–133, 1997.
[24] P. Auer, N. Cesa-Bianchi, and P. Fischer, "Finite-time analysis of the multiarmed bandit problem," *Mach. Learn.*, vol. 47, no. 2, pp. 235–256, 2002.
[25] S. Grimaldi, A. Mahmood, and M. Gidlund, "Real-time interference identification via supervised learning: Embedding coexistence awareness in IoT devices," *IEEE Access*, vol. 7, pp. 835–850, 2019.

GYUNGMIN KIM (Graduate Student Member, IEEE) received the B.S. degree from the School of Electronic Engineering, Ajou University, Suwon, South Korea, in 2016. He is currently pursuing the Ph.D. degree with the School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju, South Korea. His research interests include resource management and security for wireless networks and artificial intelligence.

HYUK LIM (Member, IEEE) received the B.S., M.S., and Ph.D. degrees from the School of Electrical Engineering and Computer Science, Seoul National University, Seoul, South Korea, in 1996, 1998, and 2003, respectively. From 2003 to 2006, he was a Postdoctoral Research Associate with the Department of Computer Science, University of Illinois at Urbana–Champaign, Champaign, IL, USA. He is currently a Full Professor with the Artificial Intelligence Graduate School Gwangju Institute of Science and Technology (GIST), Gwangju, South Korea, and the School of Electrical Engineering and Computer Science, GIST. His research interests include network protocol design, optimization, and the performance evaluation of computer and communication networking systems.