Security in Mobile Edge Caching with Reinforcement Learning

Liang Xiao, Xiaoyue Wan, Canhuang Dai, Xiaojiang Du, Xiang Chen, and Mohsen Guizani

Abstract

Mobile edge computing usually uses caching to support multimedia contents in 5G mobile Internet to reduce the computing overhead and latency. Mobile edge caching (MEC) systems are vulnerable to various attacks such as denial of service attacks and rogue edge attacks. This article investigates the attack models in MEC systems, focusing on both the mobile offloading and the caching procedures. In this article, we propose security solutions that apply reinforcement learning (RL) techniques to provide secure offloading to the edge nodes against jamming attacks. We also present lightweight authentication and secure collaborative caching schemes to protect data privacy. We evaluate the performance of the RL-based security solution for mobile edge caching and discuss the challenges that need to be addressed in the future.

Introduction

Mobile edge computing provides data storage, computing, and application services with edge devices such as access points (APs), laptops, base stations, switches, and IP video cameras at the network edge. Being closer to customers than cloud, mobile edge computing can support the Internet of Things (IoT), cyber-physical systems, vehicular networks, smart grids, and embedded artificial intelligence (AI) with lower latency, and location awareness and mobility support [1–3]. Mobile edge caching (MEC) reduces the duplicated transmissions and backhaul traffic, improves the communication efficiency, and provides quality of service for caching users. Collaborative caching in mobile edge computing shares popular data such as multimedia contents in video games with augmented reality among end users and significantly reduces the traffic load and service latency in the fifth generation (5G) mobile Internet [4].

Security and data privacy are critical and become the bottleneck for the development of MEC, as edge devices are located at the edge of the heterogenous networks and physically closer to attackers. With limited computation, energy, communication, and memory resources, the edge devices are protected by different types of security protocols, which are in general less secure compared to cloud servers and data centers. In addition, MEC systems consist of distributed edge devices that are controlled by selfish and autonomous people. The edge device owners might be curious about the data contents stored in their cache and sometimes even launch insider attacks to analyze and sell the privacy information of the customers. Therefore, MEC systems are more vulnerable to security threats such as wireless jamming, distributed denial of service (DoS) attacks, spoofing attacks including rogue edge and rogue mobile devices, man-in-the-middle attacks, and smart attacks [5].

In this article, we briefly review the security and privacy challenges of MEC and investigate the trade-off between the MEC security performance and the protection overhead in terms of computation complexity and time, communication overhead, and energy consumption. Edge devices and mobile devices have different computing and storage resources, battery levels, communication bandwidths, and locations. Each node has to optimize its defense strategy and choose the key parameters in the security protocols, which are challenging in the heterogenous dynamic network as the dynamic network model and attack model are difficult to estimate. For instance, the test threshold as a key parameter in the physical (PHY) authentication is set based on the known radio propagation and spoofing model, or a large amount of training data. However, neither the network model nor the large volume of training data can be readily obtained in time for an edge node or a mobile device to authenticate each received message [6].

The dilemma in MEC security can be addressed by reinforcement learning (RL) techniques, which enable a learning agent to derive an "optimal" strategy via trial and error. It has been proved that Q-learning, the model-free and widely used RL algorithm, can achieve the highest cumulative reward in the Markov decision process (MDP) [7]. By applying RL techniques, cyber systems such as AlphaGo have beaten human players in various games and have attracted extensive attention from both academia and industry. In recent years, RL techniques have been used to study the dynamic security games, and the proposed RL-based security schemes, such as the anti-jamming channel access scheme, the authentication scheme, and the malware detection scheme, exceed the benchmark deterministic schemes [6, 8, 11]. Therefore, we investigate the repeated

This work was supported in part by the National Natural Science Foundation of China under Grant 61671396, 91638204 and 61501527, the open research fund of National Mobile Communications Research Laboratory, Southeast University (No.2018D08), and the State's Key Project of Research and Development Plan (No.2016YFE0122900-3), the Fundamental Research Funds for the Central Universities, Guangdong Science and Technology Project (No.2016A010106002), and 2016 Major Project of Collaborative Innovation in Guangzhou (No.201604046008).

Digital Object Identifier: 10.1109/MWC.2018.1700291

Xiaoyue Wan and Canhuang Dai are with Xiamen University; Liang Xiao is with Xiamen University and the National Mobile Communications Research Laboratory, Southeast University; Xiaojiang Du is with Temple University; Xiang Chen is with Sun Yat-sen University; Mohsen Guizani is with the University of Idaho.

1536-1284/18/$25.00 © 2018 IEEE

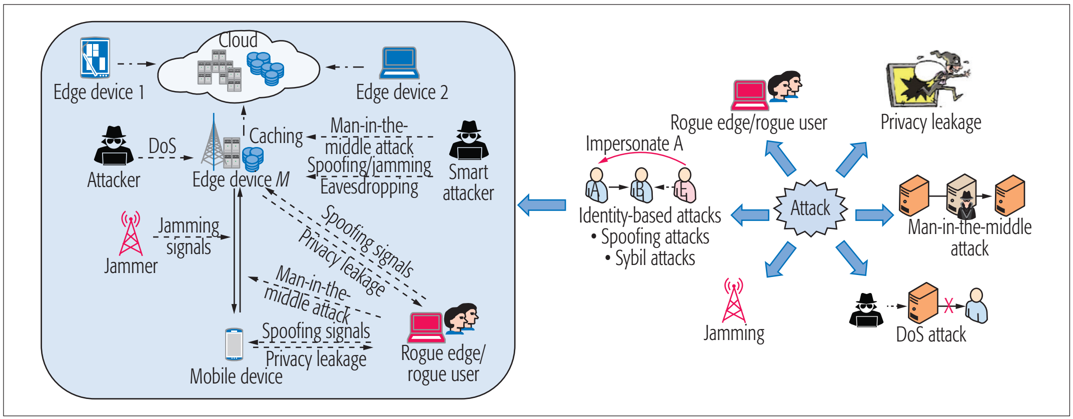

FIGURE 1. Threats in mobile edge caching.

MEC systems are more vulnerable to security threats such as wireless jamming, distributed denial of service attacks (DoS), spoofing attacks including rogue edge and rogue mobile devices, man-in-the-middle attacks, and smart attacks.

game between the MEC systems and attackers and discuss how to build the RL-based security solutions, such as secure mobile offloading against jamming and smart attacks, light-weight authentication with multiple protection levels, and collaborative caching, to resist eavesdropping.

We briefly review the RL-based security techniques and compare their performance via simulations. The challenges to implement the RL-based edge security solutions on practical MEC are discussed. Developed mostly for games such as Go and video games, most RL techniques require the agent to accurately observe the environment state and receive an immediate reward from each action. Unfortunately, these conditions rarely hold in the MEC security game, and MEC systems have to be protected from the security disasters due to the trial errors of the RL algorithms.

This article is organized as follows. In the next section we review the main security issues and present the attack models in MEC systems. We then describe how to build the RL-based MEC security solutions and evaluate their performance. We also identify the challenges ahead and point out several possible directions for future work. Finally, we draw conclusions.

Threat Model in Mobile Edge Caching

In MEC, an adversary can compromise a number of "weak" edge nodes such as video cameras protected by lightweight authentication and encryptions. By using the compromised edge nodes and/or commercial radio devices, an attacker can attack the mobile devices and edge nodes. In addition, selfish customers and curious owners of the edge devices who are hungry for secrets and money also have motivations to attack MEC systems if they know their illegal gains do not incur any punishment. Moreover, by applying advanced machine learning techniques and smart radio transmission devices, a smart attacker learns the ongoing network status and chooses its attack strategy accordingly, which makes it more dangerous for MEC systems.

As illustrated in Fig. 1, MEC has to address a large number of attacks during the mobile offloading procedure from the caching perspective. During mobile offloading, the radio communication channels of an MEC system are vulnerable to attacks launched from the PHY layer or medium access control (MAC) layers, such as jamming, rogue edge nodes/mobile devices, eavesdropping, man-in-the-middle attacks, and smart attacks. The data stored in the cache of the edge devices have to be protected to avoid privacy leakage. We briefly review some important types of attacks as follows.

Jamming: A jammer sends faked signals to interrupt the ongoing radio transmissions of the edge node with cached chunks or caching users and prevent the caching users accessing the cached contents. Another goal of jammers is to deplete the bandwidth, energy, central processing unit (CPU), and memory resources of the victim edge nodes, caching users, and sensors during their failed communication attempts [12].

DoS: DoS attacks are one of the most dangerous security threats, in which attackers aim to break down the victim computer network or cyber systems and interrupt their services. MEC systems are especially vulnerable to distributed DoS attacks, in which some distributed edge devices that are not well protected by security protocols can easily be compromised and then used to attack other edge nodes. Some attackers also aim to prevent the collaborative caching users from accessing the caching data. Jamming can be viewed as a special type of DoS attack.

Spoofing attacks/rogue edge/rogue mobile user/Sybil attacks: An attacker sends spoofing signals to edge nodes with cached chunks or the caching users with the identity of another node such as the MAC address to obtain illegal access to the network resources, and perform further attacks such as DoS and man-in-the-middle attacks [6]. For example, an attacker claims to be an edge node to fool the mobile devices in the area in rogue edge attacks, or sends spoofing messages to the edge node with the identity of another user in rogue user attacks. Faked caching space claimed by a rogue edge can result in significant data loss among the caching users in a collaborative MEC shared with a large number of users. In a Sybil attack, another type of identity-based attack, a caching user claims to be multiple users and requests more network and storage resources.

Man-in-the-middle attacks: A man-in-the-middle attacker sends jamming and spoofing signals to fake an edge node [5] with the goal of hijacking the private communication of the victim edge nodes or mobile devices and even control them.

Privacy leakage: Some owners of edge devices are curious about the data stored in their caching, and apply machine learning techniques and

Attack	RL techniques	Action	Performance	Ref
Spoofing	Q-learning Dyna-Q DQN	Test threshold Offloading rate Auth. level	False alarm rate Misdetection rate Utility of the receiver	[6, 13]
Jamming	Q-learning PDS Hotbooting Q DQN Fast DQN	Channel selection Power control Offloading rate	SINR BER Energy consumption	[8, 10–14]
Eavesdropping	Q-learning DQN Fast DQN	Defense mode Offloading rate	Secrecy data rate	[10]
Malware	Q-learning Dyna-Q PDS	Offloading rate	Detection accuracy Detection delay	[10]

TABLE 1. Summary of the RL-based security methods in wireless networks.

data analysis software to scan the caching data. In addition, the lightweight authentication protocols cannot always prevent rogue caching users from accessing the caching data. Therefore, MEC systems have to protect the caching user privacy information such as the preferences and travel histories of a specific user during mobile offloading and the caching process.

Smart Attacks: By using smart radio devices such as universal software radio peripherals (USRPs), an attacker can observe the network state such as the traffic pattern in the area, compromise some edge nodes with insufficient security protections, and wiretap the public control channels of the edge network. The attacker can also use machine learning techniques to investigate the network pattern and attack the MEC systems accordingly, possibly with multiple steps. For example, a proactive eavesdropper may first send jamming or spoofing signals to the victim edge node to receive more information from it. In [13], a smart attacker can choose the type of attacks according to its distances to the edge nodes, which has been proved to be more dangerous to MEC systems than traditional attackers that can launch a single type of attack.

RL-Based MEC Security Solutions

Each edge device or mobile device in MEC systems has to make a number of decisions to address the security threats mentioned in the previous section. For instance, a mobile device has to choose the data, the transmit power, channel and time, and the edge node in the mobile offloading against smart attackers who launch jamming, eavesdropping, rogue edge, and man-in-the-middle attacks according to the ongoing offloading policies and the network states. Most existing edge security solutions are either fixed strategies based on a certain fixed network and attack model or the optimization results based on the accurate knowledge on a number of parameters that are challenging to be obtained by an edge node in a practical edge system, because many of the network and attack parameters change significantly over time and are difficult to be estimated. Therefore, an MEC system has to find a proper security strategy without heavily depending on a specific network and attack model, which cannot be formulated as an optimization problem that is easy to address by an edge node or a mobile device.

This dilemma is promising to be addressed by applying reinforcement learning techniques such as deep Q-network (DQN), and RL-based security solutions enable a wireless device to optimize its policy in the repeated security game via trial and error. In the RL-based security scheme, a learning agent such as an edge node or a mobile device observes the current state and a quality function or Q-function to choose its action such as the security complexity and defense levels. The state corresponds to the status of the other nodes in the MEC system and the attack characters that can be observed by the node. If the future reward to the node is independent of the previous state for the given current state and strategy, the node can achieve the optimal strategy after sufficient interactions with the attackers in the dynamic edge system.

One of the first wireless security issues to apply RL techniques is anti-jamming communications [8, 10–14], showing that a transmitter can use RL algorithms such as Q-learning to optimize its transmit power and channel selection in some simplified communication scenarios, such as very few feasible actions and possible states, without being aware of the network model and the jamming model. As summarized in Table 1, RL techniques have also been used in detection of spoofing [6, 13], smart attacks [13], and malware [9]. Therefore, RL is promising to improve MEC security, although the RL-based MEC security solutions are complicated with more challenges to address. For concrete examples, we show how to apply RL techniques in the anti-jamming offloading, authentication and anti-eavesdropping transmission issues as follows.

RL-Based Anti-Jamming Mobile Offloading: In the anti-jamming MEC offloading system, a mobile device has to choose its offloading policy, such as the part of the data to offload, the transmit power, channel and time, and to which edge nodes to connect, each from a given finite feasible action set. The goal is to improve the offloading quality, such as the signal-to-noise-plus-interference ratio (SINR) and bit error rate (BER) of the signals received by the edge nodes against jamming and interference, and save the computation and communication energy consumption.

As the future state observed by a mobile device is independent of the previous states and actions for a given state and offloading strategy in the current time slot, the mobile offloading strategy chosen by the mobile device in the repeated game with jammers and interference sources can be viewed as an MDP with finite states [13]. Therefore, a mobile device can apply reinforcement learning techniques to achieve the optimal offloading policy without being aware of the jamming model and the MEC model.

In the RL-based offloading scheme as presented in [9], the mobile device observes the received jamming power, the radio channel bandwidth, the battery levels, and the user density to formulate the state. As illustrated in Fig. 2, the mobile device chooses the offloading policy such as the edge selection and offloading rate based on the current state and the Q-function, which is the expected

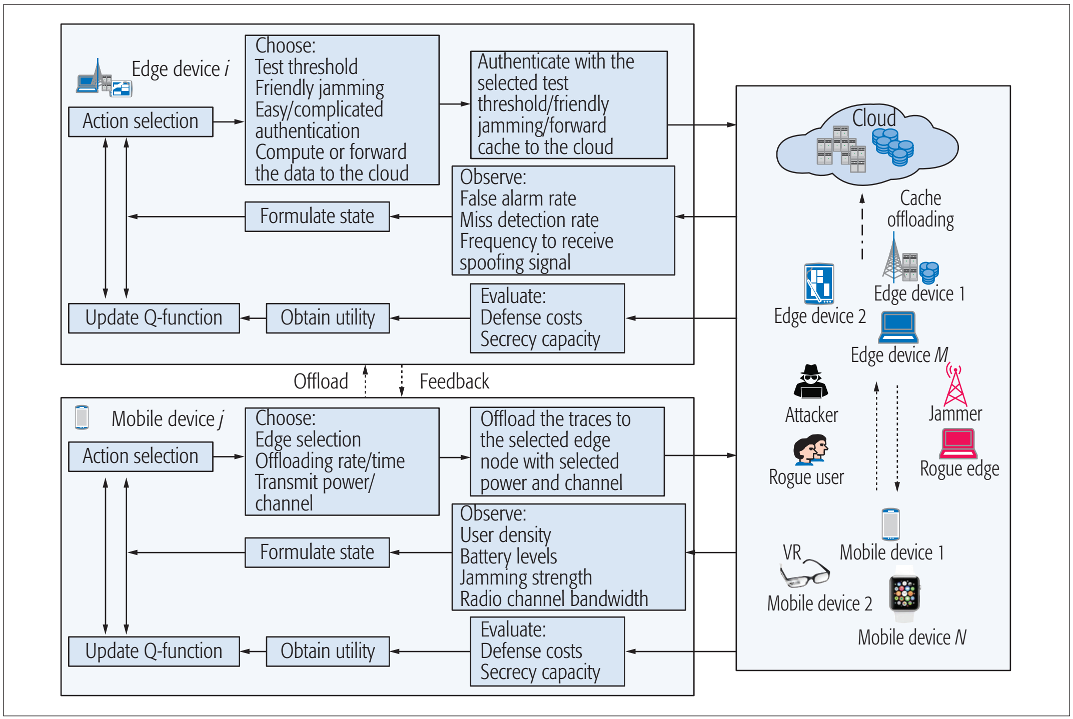

FIGURE 2. RL-based security solutions for MEC systems.

Each edge device or mobile device in MEC systems has to make a number of decisions to address the security threats mentioned in the previous section. For instance, a mobile device has to choose the data, the transmit power, the channel and time, and the edge node in the mobile offloading against smart attackers who launch jamming, eavesdropping, rogue edge, and man-in-the-middle attacks according to the ongoing offloading policies and the network states.

discounted long-term reward for each action-state pair and represents the knowledge obtained from the jamming defense history. The values of the Q-function are updated via the iterative Bellman equation in each time slot according to the current offloading policy, the network state, and the utility received by the mobile device against jamming.

The utility of the mobile device received in a time slot is evaluated according to the anti-jamming communication efficiency such as the SINR of the signals, the BER of the received messages, and the defense costs such as the offloading energy consumption. In the DQN-based offloading scheme, convolutional neural networks (CNNs) and the strategy sequence pool as shown in Fig. 3 are used to estimate the Q-values and provide a faster learning speed. The CNN consists of two convolutional (Conv) layers and two fully connected (FC) layers, and the weights of the CNN are updated based on the stochastic gradient descent (SGD) algorithm according to the previous anti-jamming communication experience in the memory pool [12]. The output of the CNN is used for estimating the values of the Q-function for each offloading policy. By applying the ε-greedy algorithm, the mobile device chooses the offloading policy that maximizes its current Q-function with a high probability $1 - \varepsilon$ and the other policies with a small probability. This scheme can make a trade-off between the exploration (i.e., to avoid being trapped in the local optimal strategy) and the exploitation (i.e., to improve the utility).

RL-Based Authentication: Due to the limited memory, energy, and computational resources, a mobile device usually has difficulty estimating the ongoing spoofing model and prefers lightweight authentication protocols to detect identity-based attacks such as spoofing attacks, Sybil attacks, and rogue edge attacks. Each edge node also needs the fast detection of a large number of spoofing messages and rogue users. To this end, PHY-authentication techniques that reuse the existing channel estimates of the source node and/or the ambient radio signals provide lightweight protection against identity-based attacks without leaking user privacy such as their locations [6].

However, most existing PHY-authentication builds hypothesis tests to compare the radio channel with the channel record of the claimed node. Therefore, the receiver has to determine the test threshold in the authentication for each incoming message, which is challenging in an MEC system with a time-variant radio channel model and a spoofing model. This issue can be addressed by RL-based authentication schemes, in which the key authentication parameters such as the test threshold are obtained via RL techniques. For example, according to the RL-based authentication scheme developed in [6], an edge node observes the recent spoofing detection accuracy and the spoofing frequency and chooses the test threshold according to the Q-function, which is updated similar to the anti-jamming offloading mentioned above. In another example, similar to [13], RL techniques can be used for an edge node to determine its authentication methods, that is, the edge node automatically applies more authentication protocols if finding itself in a risky network with smart attackers.

RL-Based Friendly Jamming: Secure collaborative caching in MEC has to protect data privacy and resist eavesdropping. For example, an edge node can send friendly jamming signals according to the data stored in the caching system to prevent an eavesdropping attacker understanding the information sent from a mobile node or another edge node. In this way, each edge node has to determine whether to attend the friendly jamming according to the network topology, the channel models, and the presence of the attackers. An

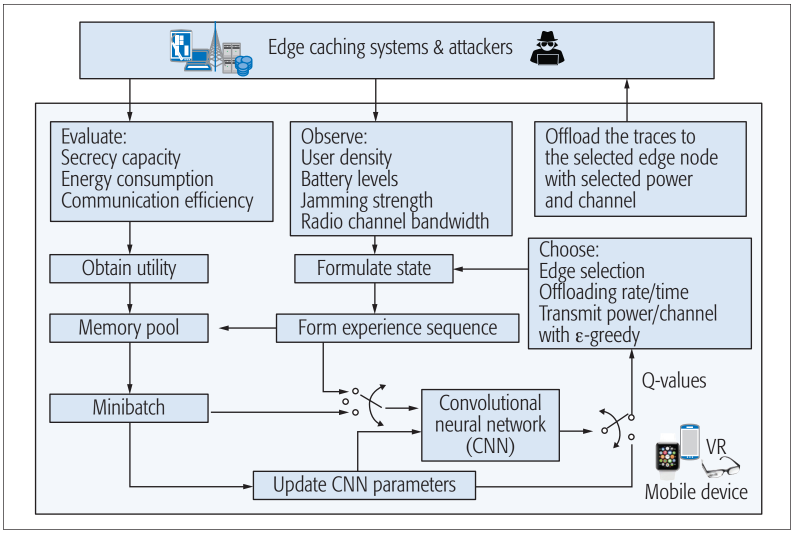

FIGURE 3. Illustration of the DQN-based secure offloading in mobile edge caching.

edge node has to decide whether to compute the data or to forward the data received from the mobile device to the cloud, and whether to store the "popular" data in the edge against privacy leakage and DoS attacks.

Performance Analysis

As a widely used RL algorithm, Q-learning has been applied to resist spoofing, malware, jamming, and eavesdropping in wireless networks, as summarized in Table 1. Without requiring any knowledge of the network and attack model, the Q-learning-based security schemes apply the iterative Bellman equation to update the Q-values, and have two parameters (i.e., the learning rate and discount factor) to control their learning performance. More specifically, the learning rate is set to weight the current experience in the learning process, and the discount factor represents the uncertainty on the feature rewards. In the Q-learning-based authentication scheme presented in [6], the learning rate is set as 0.7, and the discount factor is at 0.1 to achieve accurate spoofing detection. These schemes can easily be implemented in mobile or edge nodes with low computational and storage overhead, and enable it to achieve the optimal strategy with probability one after a sufficiently large number of interactions with the attackers in an MDP even with randomness.

However, Q-learning-based edge security suffers from the "high-dimensional disaster," as the mobile or edge nodes have to explore all the feasible state and action pairs to understand the network and attacks before the network state significantly changes or the attackers change their policies. It has been found that the learning speed of a Q-learning-based scheme is usually slower than the network variation speed, which seriously degrades the edge security performance.

Therefore, the Dyna-Q-based security methods such as the authentication scheme developed in [6] use both real defense experience and virtual experience generated by the Dyna architecture to find the optimal strategy. The Dyna-Q-based authentication scheme utilizes hypothetical experience to accelerate the learning process and thus improve the spoofing detection accuracy. However, the virtual experience is not always true, especially at the beginning of the security game, which decreases the learning rate of the security methods.

To address this issue, the edge security schemes based on the post decision state (PDS) [15] apply the known information regarding the network, attack, and channel models to accelerate the exploration and use Q-learning to study the unknown state space. On the other hand, the edge node, without being aware of any network model, can resort to the DQN technique that compresses the state space with deep learning. The DQN-based security schemes converge to the optimal strategies faster compared to the RL techniques mentioned above, especially when the edge node witnesses a large network state space. However, the implementation of the CNN in these schemes requires high computational complexity and memory, which exceeds the capability of many edge and mobile devices. To this end, a hotbooting method as a special case of transfer learning exploits the learning experience in similar scenarios to initialize the weights of the CNN and reduce the random explorations at the beginning of the learning process. A fast DQN-based anti-jamming communication method presented in [12] applies both DQN and the hotbooting technique to improve the communication efficiency against jamming.

Simulations and preliminary experiments built on laptops and USRPs show that the RL-based security solutions are promising to improve edge security. In the simulations, a user device can connect to three edge nodes against a mobile sweeping jammer that chooses its jamming power and location according to the ε-greedy algorithm. The mobile device chooses its offloading policy, including the serving edge node, the offloading rate, and the transmit power, according to the observed environment states such as the previous jamming strength. Both the discount factor and ε are set to be 0.1, the learning rate is 0.7, and the CNN parameters are set according to [12]. As presented in Fig. 4, the DQN-based offloading scheme can significantly reduce the offloading energy consumption and the delay, and increase the SINR of the signals received by the edge nodes compared to the benchmark schemes. All these schemes converge to the optimal strategy that can be validated via the Nash equilibrium of the repeated edge security game after a sufficiently long time, although DQN requires the shortest learning time.

Challenges and Future Work

Most existing RL techniques are first developed for various games, in which a learning agent accurately knows its state and immediate reward from each action (e.g., the change of the scores in a video game). In addition, an agent can tolerate most results of the feasible strategies, especially at the beginning of the game, which is the basis of the trial-and-error methods. Unfortunately, these assumptions do not hold in network security. For instance, a non-optimal network defense decision sometimes leads to forbidden results such as national safety risks. Although the RL-based security techniques are promising to improve edge security and privacy, they have to address the following challenges.

Inaccurate and Delayed State Information: An edge device usually has difficulty estimating the current network and attack state accurately and fast enough to choose the next defense policy. Therefore, the impacts of inaccurate and delayed state information on the MEC security performance have to be investigated. We have to improve the MEC security solutions with advanced RL techniques that require less state information and tolerate the inaccurate and delayed state observation for 5G communication systems. A promising solution is to incorporate the known network and attack information extracted with data mining to accelerate the learning process.

Evaluation of the Utility for Each Security Strategy: An agent has to observe the security gain and the protection cost to evaluate its reward from each action. Each in turn consists of a large number of factors. For example, in a secure mobile offloading, a mobile device has to accurately evaluate the data privacy, the transmission and computation delay, the energy cost, and the rogue edge risks from its last offloading policy, and incorporate them properly to evaluate the utility, which is challenging for most practical MEC systems. The 5G communication systems have to investigate these factors in the utility evaluation instead of using the heuristic model used in most existing RL-based security schemes. It is critical to replace the heuristic RL methods such as Q-learning in the MEC security solutions with the newly developed RL techniques that work well with delayed and inaccurate utility information.

Makeup Protocol for a Bad RL Decision: Existing RL techniques require an agent to try some bad policies to learn the optimal strategy. This exploration, which is dangerous for edge security, indicates a large number of failed defenses against attackers. To this end, transfer learning techniques that use data mining to explore existing defense experiences can be designed to help the RL reduce the random exploration and thus decrease the risks of trying bad defense policies at the beginning of the learning process. Backup protocols have to be designed for the 5G system to avoid a security disaster from a bad decision made in the learning process such as connecting with a rogue edge.

Conclusion

In this article, we study several security challenges in MEC systems and propose a security solution based on reinforcement learning. The solution consists of a secure mobile offloading solution against smart attacks, lightweight authentication, and a caching collaboration scheme to resist wiretaps. We apply RL to choose the defense levels and/or key parameters in the process. RL-based secure mobile edge caching can enhance the security and user privacy of mobile edge caching systems. As shown in the simulation results, the RL-based security solution is effective in protecting MEC systems against various types of smart attacks with low overhead.

References

[1] Y. Mao *et al.*, "A Survey on Mobile Edge Computing: The Communication Perspective," *IEEE Commun. Surveys & Tutorials*, vol. 19, no. 4, Aug. 2017, pp. 2322–58.

[2] X. Du *et al.*, "An Effective Key Management Scheme for Heterogeneous Sensor Networks," *Ad Hoc Networks*, vol. 5, no. 1, Jan. 2007, pp. 24–34.

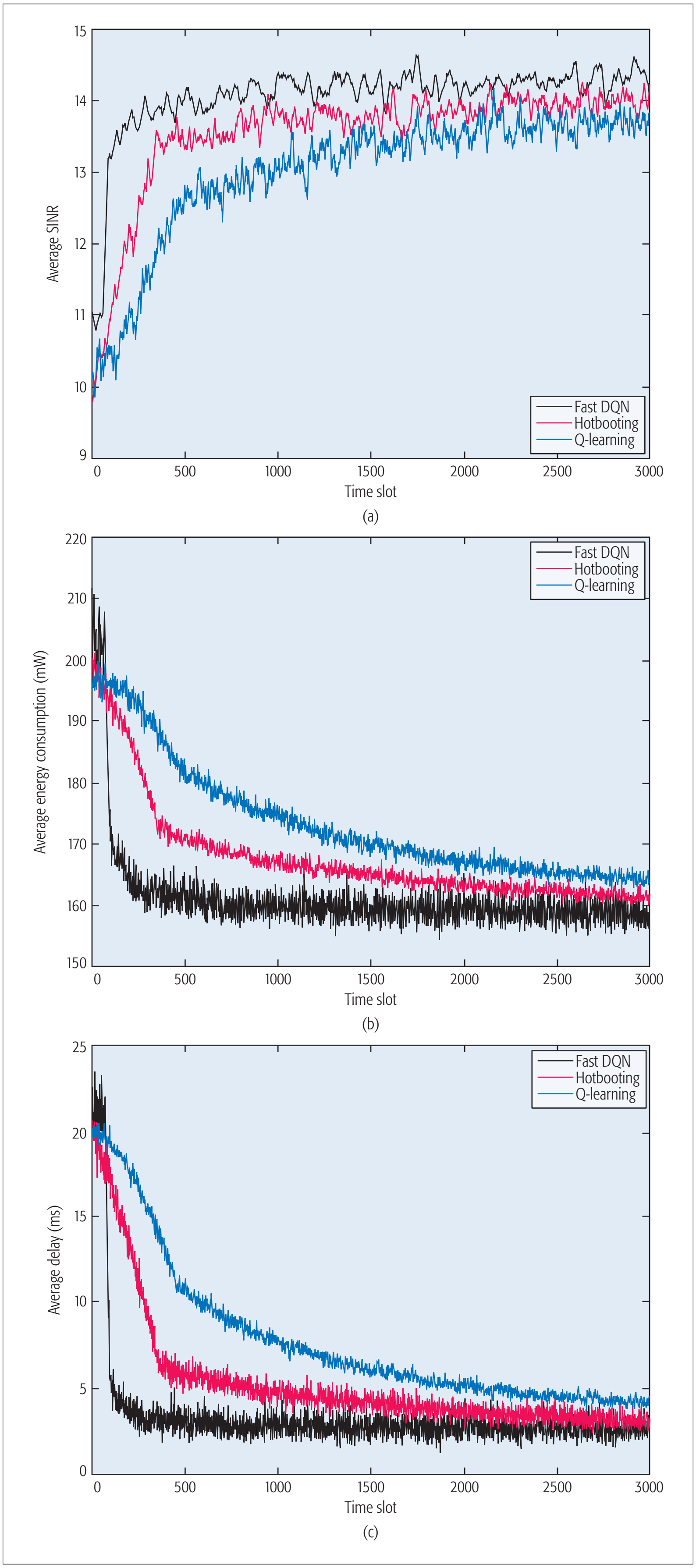

FIGURE 4. Performance of the RL-based offloading for a mobile device that is close to three edge devices against jamming: a) SINR; b) local energy consumption; c) computing delay.

RL-based secure mobile edge caching can enhance the security and user privacy of mobile edge caching systems. As shown in the simulation results, the RL-based security solution is effective in protecting MEC systems against various types of smart attacks with low overhead.

[3] Y. Xiao *et al.*, "A Survey of Key Management Schemes in Wireless Sensor Networks," *Computer Commun.*, vol. 30, no. 11–12, Sept. 2007, pp. 2314–41.

[4] X. Wang *et al.*, "D2D Big Data: Content Deliveries over Wireless Device-to-Device Sharing in Realistic Large-Scale Mobile Networks," *IEEE Wireless Commun.*, vol. 25, no. 1, Feb. 2018, pp. 1–10.

[5] R. Roman, J. Lopez, and M. Mambo, "Mobile Edge Computing, Fog *et al.*: A Survey and Analysis of Security Threats and Challenges," *Future Generation Computer Systems*, vol. 78, no. 2, Jan. 2018, pp. 680–98.

[6] L. Xiao *et al.*, "PHY-Layer Spoofing Detection with Reinforcement Learning in Wireless Networks," *IEEE Trans. Vehic. Tech.*, vol. 65, no. 12, Dec. 2016, pp. 10,037–47.

[7] C. Jiang *et al.*, "Machine Learning Paradigms for Next-Generation Wireless Networks," *IEEE Wireless Commun.*, vol. 24, no. 2, Apr. 2016, pp. 98–105.

[8] M. A. Aref, S. K. Jayaweera, and S. Machuzak, "Multi-Agent Reinforcement Learning Based Cognitive Anti-Jamming," *Proc. IEEE WCNC*, San Francisco, CA, Mar. 2017, pp. 1–6.

[9] L. Xiao *et al.*, "Cloud-Based Malware Detection Game for Mobile Device with Offloading," *IEEE Trans. Mobile Computing*, vol. 16, no. 10, Mar. 2017, pp. 2742–50.

[10] C. Xie and L. Xiao, "User-Centric View of Smart Attacks in Wireless Networks," *IEEE Int'l. Conf. Ubiquitous Wireless Broadband*, Nanjing, China, Oct. 2016, pp. 1–6.

[11] Y. Wu *et al.*, "Anti-Jamming Games in Multi-Channel Cognitive Radio Networks," *IEEE JSAC*, vol. 30, no. 1, Jan. 2012, pp. 4–15.

[12] G. Han, L. Xiao, and H. V. Poor, "Two-Dimensional Anti-Jamming Communication Based on Deep Reinforcement Learning," *IEEE Int'l. Conf. Acoustics, Speech and Signal Processing*, New Orleans, LA, Mar. 2017, pp. 2087–91.

[13] L. Xiao *et al.*, "A Mobile Offloading Game Against Smart Attacks," *IEEE Access*, vol. 4, May 2016, pp. 2281–91.

[14] Y. Gwon *et al.*, "Competing Mobile Network Game: Embracing Anti-Jamming and Jamming Strategies with Reinforcement Learning," *IEEE Conf. Commun. and Network Security*, National Harbor, MD, Oct. 2013, pp. 28–36.

[15] X. He, H. Dai, and P. Ning, "Improving Learning and Adaptation in Security Games by Exploiting Information Asymmetry," *IEEE INFOCOM*, Hong Kong, China, May 2015, pp. 1787–95.

Biographies

Liang Xiao [M'09, SM'13] (lxiao@xmu.edu.cn) is currently a professor in the Department of Communication Engineering, Xiamen University, Fujian, China. She has served as an Associate Editor of *IEEE Transactions on Information Forensics and Security* and a Guest Editor of the *IEEE Journal of Selected Topics in Signal Processing*. She is the recipient of best paper awards for the INFOCOM 2016 Big Security Workshop and ICC 2017. She received her B.S. degree in communication engineering from Nanjing University of Posts and Telecommunications, China, in 2000, her M.S. degree in electrical engineering from Tsinghua University, China, in 2003, and her Ph.D. degree in electrical engineering from Rutgers University, New Jersey, in 2009. She has been a visiting professor with Princeton University, Virginia Tech, and the University of Maryland, College Park.

Xiaoyue Wan received her B.S. degree in communication engineering from Xiamen University, China, in 2016. She is currently pursuing an M.S. degree with the Department of Communication Engineering, Xiamen University. Her research interests include network security and wireless communications.

Canhuang Dai received his B.S. degree in communication engineering from Xiamen University in 2017. He is currently pursuing an M.S. degree with the Department of Communication Engineering, Xiamen University. His research interests include smart grid and wireless communications.

Xiaojiang Du [SM'09] (dxj@ieee.org) received his B.S. and first M.S. degrees in electrical engineering from Tsinghua University, Beijing, China, in 1996 and 1998, respectively. He received his second M.S. and Ph.D. degrees in electrical engineering from the University of Maryland, College Park in 2002 and 2003, respectively. He is a professor in the Department of Computer and Information Sciences at Temple University, Philadelphia, Pennsylvania. His research interests are security, wireless networks, and systems. He has authored over 250 journal and conference papers in these areas, as well as a book published by Springer. He has been awarded more than US$5 million in research grants from the U.S. National Science Foundation (NSF), Army Research Office, Air Force, NASA, the State of Pennsylvania, and Amazon. He won the best paper award at IEEE GLOBECOM 2014 and the best poster runner-up award at ACM MobiHoc 2014. He served as the lead Chair of the Communication and Information Security Symposium of IEEE ICC 2015 and a Co-Chair of the Mobile and Wireless Networks Track of IEEE WCNC 2015. He is a Life Member of ACM.

Xiang Chen [M'08] (chenxiang@mail.sysu.edu.cn) received his B.E. and Ph.D. degrees from the Department of Electronic Engineering, Tsinghua University in 2002 and 2008, respectively. From July 2008 to July 2012, he was with the Wireless and Mobile Communication Technology R&D Center, Research Institute of Information Technology, Tsinghua University. From August 2012 to December 2014, was with the Tsinghua Space Center, School of Aerospace, Tsinghua University. Since January 2015, he has served as an associate professor at the School of Electronics and Information Technology, Sun Yat-sen University, Guangzhou, China. He is also with the Research Institute of Tsinghua University in Shenzhen as a chief researcher (part-time). His research interests mainly focus on 5G wireless communications, the Internet of Things, and software radio.

Mohsen Guizani [S'85, M'89, SM'99, F'09] (mguizani@ieee.org) received his B.S. (with distinction) and M.S. degrees in electrical engineering, and M.S. and Ph.D. degrees in computer engineering from Syracuse University, New York, in 1984, 1986, 1987, and 1990, respectively. He is currently a professor and the ECE Department Chair at the University of Idaho. Previously, he served as the Associate Vice President of Graduate Studies, Qatar University, Chair of the Computer Science Department, Western Michigan University, and Chair of the Computer Science Department, University of West Florida. He also served in academic positions at the University of Missouri-Kansas City, the University of Colorado-Boulder, Syracuse University, and Kuwait University. His research interests include wireless communications and mobile computing, computer networks, mobile cloud computing, security, and smart grid. He currently serves on the Editorial Boards of several international technical journals, and is the founder and Editor-in-Chief of *Wireless Communications and Mobile Computing* (Wiley). He is the author of nine books and more than 500 publications in refereed journals and conferences. He has guest edited a number of Special Issues in IEEE journals and magazines. He has also served as a member, Chair, and General Chair of a number of international conferences. He has received multiple teaching awards from different institutions as well as the best research award from three institutions. He received the Wireless Technical Committee's Recognition Award in 2017. He was the Chair of the IEEE Communications Society Wireless Technical Committee and the TAOS Technical Committee. He served as an IEEE Computer Society Distinguished Speaker from 2003 to 2005. He is a Senior Member of ACM.