SCIENCE CHINA
Information Sciences

July 2022, Vol. 65 170304:1–170304:16
https://doi.org/10.1007/s11432-021-3284-y

• RESEARCH PAPER •
Special Focus on Cyber Security in the Era of Artificial Intelligence

Reliable resource allocation with RF fingerprinting authentication in secure IoT networks

Weiwei WU, Su HU*, Di LIN & Gang WU

National Key Laboratory of Science and Technology on Communications, University of Electronic Science and Technology of China, Chengdu 610054, China

Received 4 February 2021/Revised 23 April 2021/Accepted 9 June 2021/Published online 20 June 2022

Abstract The unprecedented growth of the Internet of Things (IoT) has led to a huge amount of wireless resource consumption in a network. Due to limited wireless resources, a network can only guarantee the quality of service (QoS) of authenticated users rather than that of all users. By acknowledging this limitation, we realise that user authentication would be a big issue in IoT networks. Although traditional authentication methods can enhance network security to a certain extent, their vulnerability to malicious attacks and the relevant complicated computations restrict IoT deployments. In this paper, a radio frequency (RF) fingerprinting based authentication scheme is proposed under the architecture of convolutional neural network (CNN). It can effectively prevent unauthenticated users from consuming valuable wireless resources and significantly improve QoS performance for legitimate users. By solving an NP-hard optimization problem with the objective of minimizing efficient energy density, we demonstrate an approximate optimal resource allocation scheme in consideration of an RF-fingerprinting based authentication process. The analytic results show that our proposed scheme can dramatically reduce the efficient energy density compared with traditional cryptography based authentication schemes.

Keywords user authentication, Internet of things, convolutional neural network, RF fingerprinting, NP-hard optimization problem

Citation Wu W W, Hu S, Lin D, et al. Reliable resource allocation with RF fingerprinting authentication in secure IoT networks. Sci China Inf Sci, 2022, 65(7): 170304, https://doi.org/10.1007/s11432-021-3284-y

1 Introduction

With the explosive growth of the Internet of Things (IoT), the requirements of high-speed connectivity, low-latency transmission, as well as high-quality video services have facilitated the development of 5th-generation (5G) mobile communications [1]. As a promising technology in 5G communications, mobile edge computing (MEC) can significantly reduce latency and improve bandwidth efficiency by offloading tasks from the cloud to the edge [2–5]. However, a large number of devices are normally employed in most industrial IoT applications, which leads to resource and energy scarcity in wireless networks. While a few authenticated users who are willing to pay for services can access a particular network, other unauthenticated users cannot be allowed to access the network [6]. Aiming at experiencing the high quality of services (QoS) without payment, an unauthenticated user may disguise as an authenticated user by cracking legitimate account passwords.

To guarantee QoS level, it is critical to validate whether users are authenticated. Varieties of traditional authentication techniques are applied to ensure that the services at a high QoS level can merely be accessed by authenticated users in wireless networks [7]. The traditional schemes are mostly built on classical cryptography-based authentication techniques that rely on complicated mathematical operations and protocols from upper communication layers, e.g., the medium access control (MAC) layer and network layer. These schemes could mitigate security problems to a particular extent by validating users' identities and maintaining information confidentiality. In a typical industrial IoT scenario, a star network topology is normally employed [8]. A star network regularly includes a few MECs with a large number

* Corresponding author (email: husu@uestc.edu.cn)

of user devices, which usually have limited computation capacity. If a classical cryptography-based authentication scheme is implemented in this case, it is a big challenge for user devices to complete the entire process of encryption and decryption by using complicated computations [9]. Additionally, the classical methods are most likely to be susceptible to malicious attacks, e.g., duplicating and changing IP or MAC addresses [10–12]. Moreover, most of the techniques are user-dependent. They would be inefficient when users select weak and gullible passwords that can be cracked within several minutes or hours. According to Avast's report in 2018, more than 83% of Americans are using weak passwords since they set their passwords less than 10 characters. These passwords do not spontaneously include numbers, special symbols, as well as uppercase and lowercase letters [13]. Particularly, in an IoT network with increasing numbers of devices, it would be a troublesome burden for a user to set and memorize strong but complicated passwords of various devices. In summary, the existing security methods at the upper communication layers would be vulnerable to malicious attacks, require a large amount of computing resource consumption, and mostly depend on the passwords selected by users. So it is necessary to investigate an appropriate scheme that is tailored to the needs of IoT applications.

As a light-weighted and user-independent authentication solution, radio-frequency (RF) fingerprinting has recently received increasing research attention [14]. It can recognize a mobile device by analyzing and identifying its unique and inherent features at the physical layer, thereby preventing the impersonation of devices for security credentials. Specifically, RF fingerprinting does not require both communication parties to have sufficient resources to go through an energy-consuming process of encryption and decryption. Instead, RF fingerprint features, similar to the DNA of IoT devices, are extracted from the signals from transmitters. These features could then be analyzed by classification models for validating the identities of user devices. Moreover, RF fingerprinting does not rely on the IP addresses and password selections. Thus, RF fingerprinting could be a suitable authentication solution in IoT scenarios. Typically, RF fingerprinting has the following characteristics.

(1) Versatility: In the process of wireless device manufacturing, each individual device has its own unique inherent features in hardware components, even when they are manufactured in the same batch. RF fingerprints are designed to reflect the unique features of an individual device. The features are attributed to manufacturing tolerances, component aging, and workplace changes. The typical examples of the RF fingerprints include frequency deviation of oscillators, phase noise, non-linear distortion of power amplifiers, filter distortion, I/Q imbalance, phase imbalance, frequency error, etc. [15]. We can identify each wireless device by properly selecting fingerprint features.

(2) Short-term invariance: Unlike human beings who remain their fingerprint features invariant throughout their lives, the components of wireless devices would inevitably face aging. This characteristic leads to the differences between the actual fingerprint features and those registered in the fingerprint feature database. However, the aging process of components might take a long time, so the impact on RF fingerprint features during a short period is relatively low [16]. The study in [16] shows that RF fingerprint features extracted by a wireless network can remain unchanged within 5 months, so the device recognition rate is quite stable.

(3) Uniqueness: As human beings have different fingerprint features, the RF fingerprints of wireless devices are also unique due to the uniqueness of their hardware components. The state-of-the-art technologies facilitate integrated circuit chips that result in subtle differences between hardware components. Thus, how to identify the wireless devices by their unique fingerprint features is a primary issue.

In our paper, we propose a novel algorithm that not only optimizes wireless resources but also takes an RF-based authentication scheme into account. Despite a rich body of literature on the optimization of wireless resources by deploying an MEC architecture, most existing algorithms may not work well for the IoT applications, which faces the challenges of authorizing numerous wireless devices. The main contributions of our work are summarized as follows.

(1) We present the topic of resource allocation in combination with user authentication for IoT applications. This research topic not only can enhance network security for IoT applications, but also can avoid the waste of wireless resources that are allocated to unauthenticated users.

(2) We compare the performance of a physical-layer authentication method with that of the traditional authentication methods, in terms of resource allocation. To our knowledge, there are very few quantitative studies on the comparison between the performance of authentication methods at the physical layer and that of the traditional methods at high layers, e.g., network layer and application layer.

(3) We address an optimization problem with the objective of minimizing efficient energy density, and this problem is proven to be non-polynomial (NP) hard. We demonstrate a fully polynomial-time approx-

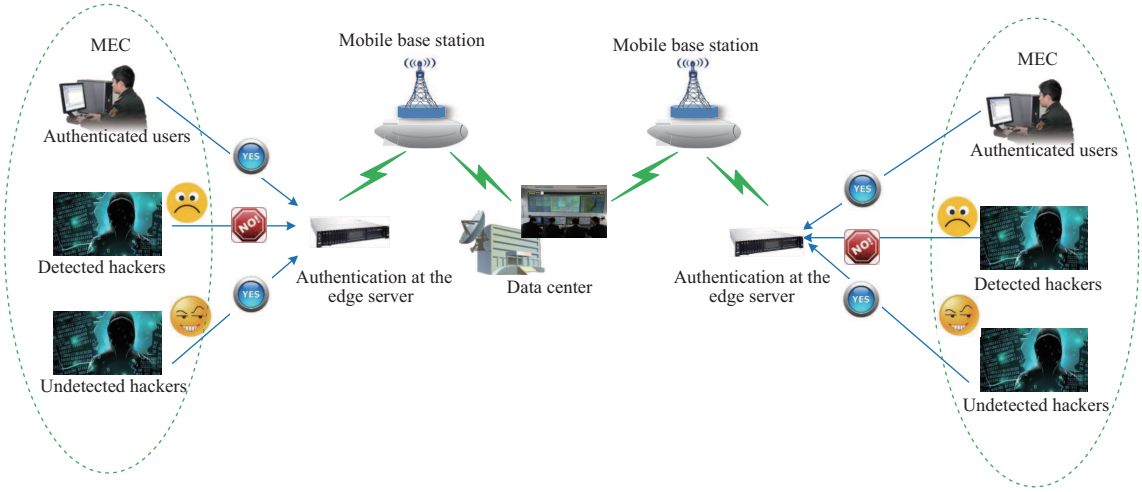

Figure 1 (Color online) Architecture of edge computing network for IoT.

imation scheme (FPTAS) based approximation scheme for optimal resource allocation in consideration of user authentication level and provide insights into our proposed method's computation complexity.

The rest of this paper is organized as follows. Section 2 demonstrates system models, including both the authentication model and the communication model. Then, an authentication based resource allocation problem and its approximation solutions are described in Sections 3 and 4, respectively. Section 5 shows the experiment results. The conclusion is discussed in Section 6.

2 System model

In typical IoT networks, a huge number of connected devices result in consuming a huge amount of wireless resource consumption. In order to reasonably allocate limited resources, a validation process that distinguishes authenticated users from unauthenticated ones would be a big issue. An appropriate authentication method has a positive impact on wireless resource optimization, but it still consumes extra energy. As the accuracy of authentication algorithms increases, the computational complexity would also correspondingly increase. In other words, a huge amount of computing resource consumption would necessarily be required for achieving the goal of the high accuracy of learning algorithms. By balancing the accuracy of the authentication process and the optimization of resource allocation, our system model should consider three perspectives: an authentication model, a communication model, and a computing model.

Against the background, a system model of an edge computing network for IoT is shown in Figure 1. Under this architecture, the network is primarily composed of three ties. The first tie is composed of several core networks and base stations, and they could be either located on earth or suspended on an airship. The second tie is composed of edge computing networks, and they are mobile servers suspended on a small-scaled vehicle or an unmanned aerial vehicle. The third tie is composed of devices at the user ends, including mobiles and laptops. A few users, including authenticated users and hackers or illegitimate users, attempt to access the network. However, the hackers may sneak into the focal network when an authentication algorithm fails to detect abnormal user devices. The traditional authentication methods, including wired equivalent privacy (WEP), Wi-Fi protected access (WPA) and WPA2, can mitigate network security problems to a certain degree, but they are vulnerable to malicious attacks, require a large amount of computing resource consumption, and mostly depend on the passwords selected by users. In the following, we first investigate an RF fingerprinting based authentication method, which is tailored to the needs of IoT applications. We then demonstrate a communication model and a computing model.

2.1 Authentication with RF fingerprinting recognition

The research on RF fingerprint recognition can primarily be categorized into two streams: machine-learning based RF fingerprint recognition and deep-learning based RF fingerprint recognition. In the first stream, most relevant fingerprint features have to be delicately selected and then employed to

train a machine-learning based classification model. The second stream uses deep-learning algorithms to automatically learn the internal characteristics of wireless signals. In the following, we demonstrate the related work of machine-learning and deep-learning algorithms for RF fingerprint recognition.

On one hand, particular studies focus on machine learning algorithms for RF fingerprint recognition [17–19]. A few studies employ various support-vector-machines (SVM) algorithms to identify mobile devices, including the PolyKernel algorithms [17] and the Pearson VII universal Kernel (PuK) algorithms [18]. Of these SVM-related algorithms, the PuK algorithms are fairly effective in RF fingerprinting recognition, dramatically improving recognition performance and reducing computation time. Besides SVM algorithms, Jian et al. [19] proposed a method of extracting spectral features with k-nearest neighbor (kNN). To improve the accuracy of kNN algorithm, a combination of support vector machine (SVM) and kNN algorithms was presented to classify 802.11 devices by analyzing the RF fingerprint, and this method could achieve an accuracy of 95%. In summary, by deploying the traditional machine learning recognition schemes, researchers first standardize original signals through power normalization, noise reduction, and label setting. Primary features are then extracted from normalized data and stored into a fingerprint library. Finally, the features are used to identify the types of mobile devices. As for traditional recognition schemes, the key is to select appropriate signal characteristics from RF fingerprints. The selection process heavily depends on expert experiences. Also, the selected features might only be effective as RF fingerprints in certain scenarios, but they cannot be generalized to other communication scenarios.

On the other hand, relevant studies have shifted their attention from machine learning algorithms to deep learning ones in recent years [16,20–23]. Sankhe et al. [20] used 2-layered convolution neural networks (CNN) to train RF fingerprinting data from 16 X310 USRP SDRs. Wu et al. [16] employed a deep neural net (DNN) with the activation functions of rectified linear units (ReLU) to run a training model for the RF fingerprinting recognition of 12 Ettus USRP N210. They further propose an incremental learning based neural network to accelerate the learning process. However, these algorithms are designed for the classification of devices, and they cannot recognize unknown devices. In addition, the above-mentioned studies do not consider negative effects of unauthenticated users on the performance of resource allocation. A lack of user authentication under the architecture of resource allocation motivates us to propose the work.

Compared with traditional machine learning algorithms, deep learning-based recognition schemes do not mandatorily require experts' previous experiences and a long-consuming process of feature extraction [21–23]. Therefore, in this paper, we use a deep learning model instead of a machine learning model to differentiate authenticated devices from unauthenticated ones. Specifically, we present a computing model (shown in Figure 2) to authenticate IoT devices. Each mobile device transmits its RF signal data to the edge server for authentication in such a model. We also consider the various detailed CNN models, including ResNet, LeNet, DenseNet, WideResNet.

In view of the CNN models, we select the following structure for authentication by balancing its accuracy and model complexity. The input data can be represented as 2×128 I/Q samples of RF signals. The model is composed of 5 layers, including 2 convolution layers, 2 pooling layers, and a fully-connected layer. Specifically, the first convolution layer is characterized as a $50 \times 1 \times 3$ network with the kernel of ReLU, and the first pooling layer is characterized as a max-pooling layer. The second convolution layer and pooling layer have the same structure as the first ones. The last layer is a fully-connected layer with the kernel of Softmax. As expected, the final output is the identity classification of a mobile device.

In summary, the edge server needs to complete service tasks and user authentication tasks, sharing the bandwidth resources and the edge computing resources of servers (detailed in Subsection 3.2).

2.2 Models in an MEC network

2.2.1 *Communication model of an MEC network*

This subsection presents the communication model of an MEC based cellular network. An MEC based cellular network contains a cellular mobile station and M mobile users. A mobile user can access the Internet via the core of a cellular network, and they share the computation resources of an MEC server that is connected with a base station in a cell. In our study, we assume that these mobile users employ orthogonal frequency division multiplexing (OFDM) schemes for communications. We also assume that the OFDM scheme is well designed with guard period and cyclic prex schemes, ensuring that both inter-symbol interference and inter-channel interference can be ignored.

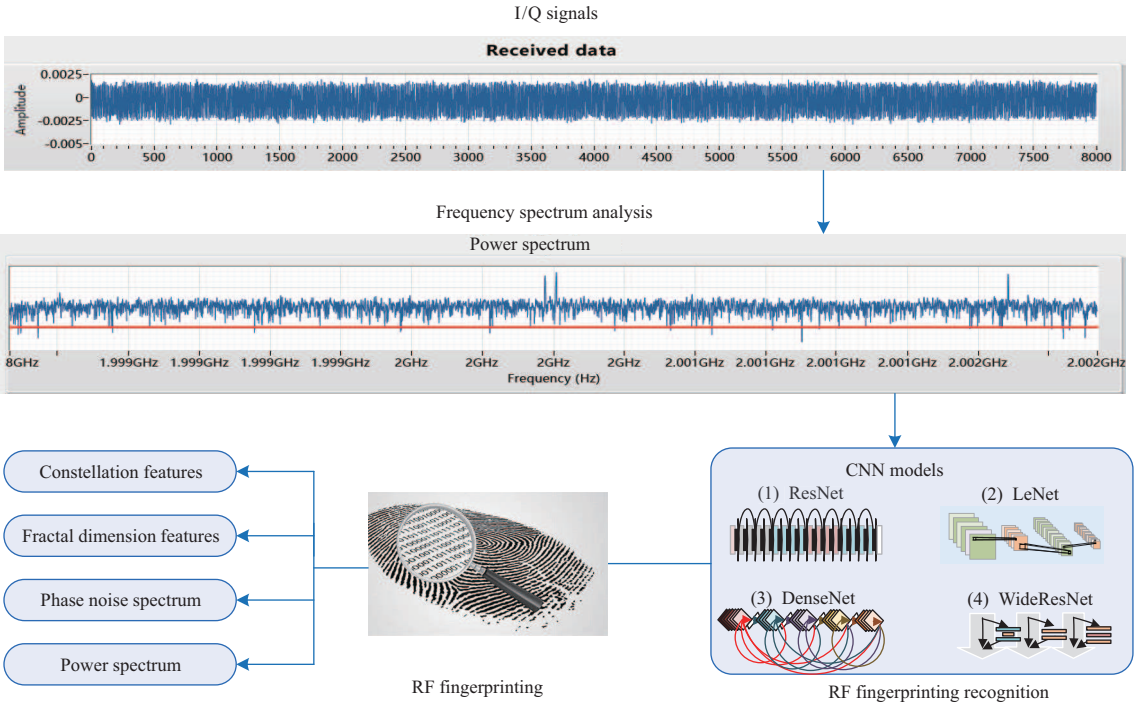

Figure 2 (Color online) Model structure of a CNN network for RF fingerprinting recognition.

Let k denote the kth mobile user, W_k^s Hz denote the amount of wireless bandwidth per OFDM subcarrier, N_k^R denote the number of resource blocks, N_k^s denote the number of subcarriers per resource block, δ_k denote the signal to noise ratio (SNR) of device k. By Shannon theorem, we can define the data rate of a mobile device k as

$$C_k = N_k^R N_k^s W_k^s \log_2(1+\delta_k), \tag{1}$$

where $N_k^R = N_k^A + N_k^S$, given N_k^A and N_k^S to be the number of resource blocks for data transmission in the process of user authentication and for service, respectively.

2.2.2 *Local computation model*

In the following model for local service computation, we compute energy consumption E_k^l and running time T_k^l as (2) and (3).

$$E_k^l = \eta_k U_k^l, \tag{2}$$

where η_k denotes the consumption of energy per computation unit, and U_k^l denotes the number of required computation units for service tasks.

$$T_k^l = \frac{U_k^l}{f_k^l}, \tag{3}$$

where f_k^l denotes the computation frequency of each local computation unit for service tasks.

2.2.3 *Edge computation model*

In the following model for edge service computation, we compute energy consumption $E_k^S = S_k^S P_k / C_k$ and running time $T_k^S = S_k^S / C_k + U_k^S / f_k^e$, where S_k^S denotes the amount of service data uploaded to edge servers, C_k denotes data transmission rate, P_k denotes average transmission power per user, S_k^S denotes the amount of service data uploaded to edge servers, f_k^e denotes the computation frequency of each edge computation unit, and U_k^S denotes the number of required computation units for data service.

In the edge computation for user authentication, we compute energy consumption $E_k^A = S_k^A P_k / C_k$ and running time $T_k^A = S_k^A / C_k + U_k^A / f_k^e$, where S_k^A denotes the amount of authentication data uploaded to edge servers, S_k^A denotes the amount of authentication data uploaded to edge servers, and U_k^A denotes the number of required computation units.

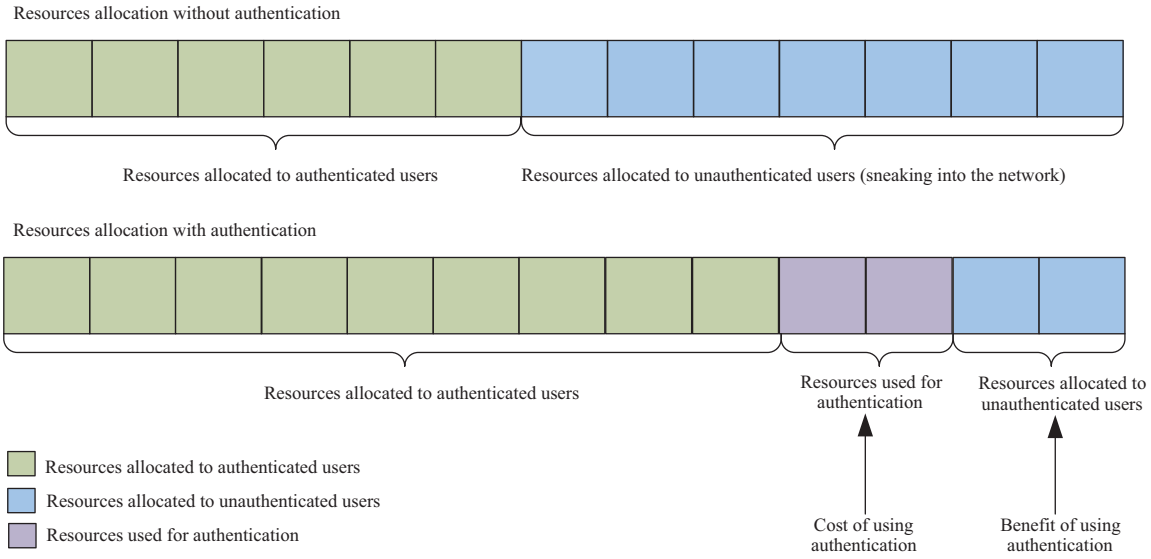

Figure 3 (Color online) Illustration of resource allocation with authentication.

Thus, we can compute total energy consumption as

$$E_k^e = E_k^S + E_k^A. \tag{4}$$

Assuming that the processes of authentication and service tasks are completed in a sequence, we can compute total delay as

$$T_k^e = T_k^S + T_k^A. \tag{5}$$

3 Optimization problem

As shown in Figure 3, when hackers or illegitimate users sneak into the focal network, a huge amount of wireless resource consumption (blue blocks in Figure 3) would result in bad or even frustrating experiences for authenticated users (green blocks in Figure 3). In achieving the goal of optimizing wireless resources, user authentication should be considered at the MEC end, and the benefit of using authentication methods is exactly saving the resources allocated to unauthenticated users (blue blocks in Figure 3). However, authentication process would consume extra energy of MECs (purple blocks in Figure 3), which is cost of using authentication methods. To minimize the energy consumption of MECs, we should further consider the factor of energy consumption from authentication.

Thus, an optimization problem to minimize energy density in consideration of user authentication is presented. Additionally, while the authentication that can detect hackers has a positive impact on energy saving, extra energy consumption of authentication should also be taken into account. In the following, we would consider a tradeoff between the energy consumption from authentication and the energy saving from authentication.

Let $\boldsymbol{K}_A = \{1, 2, \ldots, K_A\}$ denote the set of authenticated mobile users, and $\boldsymbol{K}_U = \{1, 2, \ldots, K_U\}$ denote the set of unauthenticated mobile users. The optimization of energy density can be transformed into searching for beneficial offloading strategy, which is defined as follows.

Definition 1. The beneficial offloading strategy refers to a task offloading process which can achieve the goal of lower energy consumption by edge servers than that by local computation, while guaranteeing the running time at the edge server below a delay constraint.

In the following, we mathematically define the beneficial offloading strategy. Specifically, we define the optimization problem with an authentication process.

3.1 Establishment of optimization problem

Even though using an authentication method, a few unauthenticated users can be detected and are denied to access the network, and let $\bar{\boldsymbol{K}}_U = \{1, 2, \ldots, \bar{K}_U\}$ denote the set of undetected mobile users in $\boldsymbol{K}_U$.

When no authentication method is used, $\bar{\boldsymbol{K}}_U = \boldsymbol{K}_U$ since all the unauthenticated users can sneak into the network. To simplify the presentation, we consider the case of not using any authentication method as a special case of using an authentication method, which is invalid and cannot detect any unauthenticated users.

Mathematically, the optimization problem using an authentication method can be represented as

$$\begin{aligned}
\text{(P1)} \quad & \min_{\{\phi_k, N_k^R\}} \quad D\left(\phi_k, N_k^R\right) = \frac{E\left(\boldsymbol{K}_A\right) + E\left(\bar{\boldsymbol{K}}_U\right)}{N^R\left(\boldsymbol{K}_A\right) + N^R\left(\bar{\boldsymbol{K}}_U\right)} \\
\text{s.t.} \quad & \text{C1}: \sum_{k=1}^{K} \phi_k N_k^R \leqslant N_T^R, \\
& \text{C2}: \phi_k \in [0,1], \\
& \text{C3}: E_k^e \leqslant E_k^l, \\
& \text{C4}: T_k^e \leqslant T_k^M, \\
& \text{C5}: N_k^R \in \mathbb{Z}^+, \\
& \text{C6}: K = K_A\left(1 + \mu(1-\kappa)\right),
\end{aligned} \tag{6}$$

given

$$\begin{aligned}
E\left(\boldsymbol{K}_A\right) &= \sum_{m=1}^{K_A} \phi_m E_m^e + (1-\phi_m) E_m^l, \\
E\left(\bar{\boldsymbol{K}}_U\right) &= \sum_{n=1}^{\bar{K}_U} \phi_n E_n^e + (1-\phi_n) E_n^l, \\
N^R\left(\boldsymbol{K}_A\right) &= \sum_{m=1}^{K_A} \phi_m N_m^R, \\
N^R\left(\bar{\boldsymbol{K}}_U\right) &= \sum_{n=1}^{\bar{K}_U} \phi_n N_n^R, \\
K &= K_A + \bar{K}_U,
\end{aligned}$$

where m represents the mth authenticated mobile user, n represents the nth unauthenticated mobile user, K represents the number of users in the network, K_A represents the number of authenticated mobile users, $\bar{K}_U$ represents the number of unauthenticated mobile users who are not recognized, ϕ_k denotes the proportion of data offloaded to edge servers, N_T^R denotes the total number of resource blocks, T_k^M denotes the maximal allowed delay of device k.

Constraint C1 indicates that the number of actually used resource blocks cannot be greater than that of available blocks. Constraint C2 indicates that the proportion of data that is offloaded to edge servers should be in the range of $[0,1]$. Constraint C3 indicates that energy consumption should satisfy the beneficial offloading strategy in Definition 1. Constraint C4 indicates that the delay of a device should be below the acceptable delay. Constraint C5 indicates that the number of resource blocks must be an integer. Constraint C6 indicates the relationship between the number of users in a network user and the recognition accuracy of an authentication algorithm.

In problem (P1), we define the accuracy of detecting unauthenticated users by an authentication method κ as

$$\kappa = 1 - \frac{\bar{K}_U}{K_U}. \tag{7}$$

Also we define the ratio between the number of unauthenticated users and the number of authenticated users μ as

$$\mu = \frac{K_U}{K_A}. \tag{8}$$

Thus, the total number of users in the network K can be denoted as

$$K = K_A\left(1 + \mu(1-\kappa)\right). \tag{9}$$

3.2 Representation of optimization problem

In this subsection, we represent equation (6) by replacing constraints C3 and C4 with a single constraint on N_k^R. To achieve this, we first present the following theorem based on the beneficial offloading strategy in Definition 1.

Theorem 1. To meet the requirements of beneficial offloading strategy in Definition 1, the number of resource blocks for an arbitrary device k (i.e., N_k^R) should satisfy

$$N_k^R \geqslant \max(N_k^E, N_k^T), \tag{10}$$

where

$$\begin{aligned}
&N_k^E = \left\lceil \frac{S_k P_k}{\eta_k U_k^l N_k^s W_k^s \log_2(1+\delta_k)} \right\rceil, \\
&N_k^T = \left\lceil \frac{S_k}{(T_k^M - U_k/f_k^e) N_k^s W_k^s \log_2(1+\delta_k)} \right\rceil, \\
&S_k = S_k^S + S_k^A, \\
&U_k = U_k^S + U_k^A.
\end{aligned}$$

Proof. The beneficial offloading strategy in Definition 1 is composed of two requirements, including $E_k^e \leqslant E_k^l$ and $T_k^e \leqslant T_k^M$.

As for $E_k^e \leqslant E_k^l$, we have

$$\begin{aligned}
&\frac{S_k P_k}{C_k} \leqslant \eta_k U_k^l, \\
&C_k = N_k^R N_k^s W_k^s \log_2(1+\delta_k) \\
&\Longrightarrow\ N_k^R \geqslant \frac{S_k P_k}{\eta_k U_k^l N_k^s W_k^s \log_2(1+\delta_k)}.
\end{aligned}$$

Given that N_k^E is an integer, let

$$N_k^E = \left\lceil \frac{S_k P_k}{\eta_k U_k^l N_k^s W_k^s \log_2(1+\delta_k)} \right\rceil.$$

We have

$$N_k^R \geqslant N_k^E.$$

As for $T_k^e \leqslant T_k^M$, we have

$$\begin{aligned}
&\frac{S_k}{C_k} + \frac{U_k}{f_k^e} \leqslant T_k^M, \\
&C_k = N_k^R N_k^s W_k^s \log_2(1+\delta_k) \\
&\Longrightarrow\ N_k^R \geqslant \frac{S_k}{(T_k^M - U_k/f_k^e) N_k^s W_k^s \log_2(1+\delta_k)}.
\end{aligned}$$

Given that N_k^T is an integer, let

$$N_k^T = \left\lceil \frac{S_k}{(T_k^M - U_k/f_k^e) N_k^s W_k^s log_2(1+\delta_k)} \right\rceil,$$

and we have $N_k^R \geqslant N_k^T$.

Thus, $N_k^R \geqslant \max(N_k^E, N_k^T)$.

Based on Theorem 1, we can represent (6) as the following equivalent problem:

$$\begin{aligned}
\text{(P2)} \quad & \min_{\{\phi_k, N_k^R\}} \quad \frac{\sum_{k=1}^{K} \phi_k E_k^e + (1-\phi_k) E_k^l}{\sum_{k=1}^{K} \phi_k N_k^R} \\
\text{s.t.} \quad & \text{C1}: \sum_{k=1}^{K} \phi_k N_k^R \leqslant N_T^R, \\
& \text{C2}: \phi_k \in [0,1], \\
& \text{C3}: N_k^R \geqslant \max(N_k^E, N_k^T), \\
& \text{C4}: N_k^R \in \mathbb{Z}^+, \\
& \text{C5}: K = K_A \left(1 + \mu(1-\kappa)\right).
\end{aligned} \tag{11}$$

Since $\sum_{k=1}^{K} \phi_k E_k^e + (1-\phi_k) E_k^l = \sum_{k=1}^{K} E_k^l - \sum_{k=1}^{K} \phi_k (E_k^l - E_k^e)$, we can represent (11) as

$$\begin{aligned}
\text{(P3)} \quad & \max_{\{\phi_k, N_k^R\}} \quad \frac{\sum_{k=1}^{K} \phi_k (E_k^l - E_k^e)}{\sum_{k=1}^{K} \phi_k N_k^R} \\
\text{s.t.} \quad & \text{C1}: \sum_{k=1}^{K} \phi_k N_k^R \leqslant N_T^R, \\
& \text{C2}: \phi_k \in [0,1], \\
& \text{C3}: N_k^R \geqslant \max(N_k^E, N_k^T), \\
& \text{C4}: N_k^R \in \mathbb{Z}^+, \\
& \text{C5}: K = K_A \left(1 + \mu(1-\kappa)\right).
\end{aligned} \tag{12}$$

3.3 NP-hard proof of optimization problem

In the following, we proof that the optimization problem (P3) is a NP-hard problem.

Theorem 2. The optimization problem (P3) is a NP-hard problem.

Proof. It is easy to show that problem (P3) is a transformation of maximum-density knapsack problem, which is NP-complete. In a maximum-density knapsack problem, the ratio between knapsack value and weight is designed to be maximized by allocating items into the knapsack, with the capacity of knapsack as a constraint.

In problem (P3), the part of $(E_k^l - E_k^e)$ can be assumed as the value of item k, N_k^R can be assumed as the weight of items, ϕ_k can be assumed as the decision whether an item k should be placed into the knapsack, N_T^R can be assumed as the capacity of knapsack.

Thus, problem (P3) is polynomial-time reducible to a maximum-density knapsack problem, and thus a NP-hard problem.

4 Approximate solution

4.1 Upper-bound solution

In this subsection, we investigate the upper-bound approximate solution to problem (P3) by transforming it into the problem of optimizing the summation of density since

$$\frac{\sum_{k=1}^{K} \phi_k (E_k^l - E_k^e)}{\sum_{k=1}^{K} \phi_k N_k^R} \leqslant \sum_{k=1}^{K} \frac{\phi_k (E_k^l - E_k^e)}{\phi_k N_k^R} = \sum_{k=1}^{K} \frac{(E_k^l - E_k^e)}{N_k^R}.$$

Mathematically, the problem of optimizing the summation of density is shown in (13), which can be

viewed as the upper-bound approximate solution.

$$\begin{aligned}
\text{(P4)} \quad & \max_{\{\phi_k, N_k^R\}} \quad \sum_{k=1}^{K} \frac{(E_k^l - E_k^e)}{N_k^R} \\
\text{s.t.} \quad & \text{C1}: \sum_{k=1}^{K} \phi_k N_k^R \leqslant N_T^R, \\
& \text{C2}: \phi_k \in [0,1], \\
& \text{C3}: N_k^R = \max(N_k^E, N_k^T), \\
& \text{C4}: N_k^R \in \mathbb{Z}^+, \\
& \text{C5}: K = K_A \left(1 + \mu(1-\kappa)\right).
\end{aligned} \tag{13}$$

Theorem 3. The solution to the highest density-summation problem can be achieved when the energy consumption at the edge server equals to half of energy consumption at the local end, i.e., $E_k^l = 2E_k^e$.

Proof. Let

$$D(N_k^R) = \frac{(E_k^l - E_k^e)}{N_k^R}, \quad S_k = S_k^S + S_k^A. \tag{14}$$

We have

$$\begin{aligned}
\frac{\partial D(N_k^R)}{\partial N_k^R} &= \frac{\partial(\frac{E_k^l - E_l^e}{N_k^R})}{\partial N_k^R} \\
&= \frac{\partial(\frac{\eta_k U_k^l}{N_k^R} - \frac{S_k p_k}{(N_k^R)^2 N_k^S W_k^S \log_2(1+\delta_k)})}{\partial N_k^R} \\
&= -\frac{\eta_k U_k^l}{(N_k^R)^2} + 2\frac{S_k p_k}{N_k^S W_k^S \log_2(1+\delta_k)(N_k^R)^3} \\
&= -\frac{1}{N_k^R}(E_k^l - 2E_k^e).
\end{aligned}$$

When $\frac{\partial D(N_k^R)}{\partial N_k^R} = 0$, we have $E_k^l - 2E_k^e = 0$, i.e., $E_k^l = 2E_k^e$.

Theorem 4. The upper bound solution to problem (P3) is $N_k^R = N_k^U$, given

$$S_k = S_k^S + S_k^A, \quad N_k^U = \frac{2S_k p_k}{\eta_k U_k^l N_k^S W_k^S \log_2(1+\delta_k)}.$$

Proof. By Theorem 5, we have $E_k^l = 2E_k^e$, i.e.,

$$\eta_k U_k^l = \frac{2S_k p_k}{N_k^R N_k^S W_k^S \log_2(1+\delta_k)}. \tag{15}$$

Thus, we can achieve $N_k^R = N_k^U$, given

$$N_k^U = \frac{2S_k p_k}{\eta_k U_k^l N_k^S W_k^S \log_2(1+\delta_k)}.$$

4.2 FPTAS based approximate solution

In this subsection, we investigate the approximate solution to problem (P3), which is a maximum-density knapsack problem, and it is shown that an FPTAS exists.

Theorem 5. A $(1+\epsilon)$-approximation solution to problem (P3) exists, given ϵ to be an allowed error of solution.

Proof. Let

$$P(\phi_k) = \sum_{k=1}^{K} \phi_k (E_k^l - E_k^e),$$
$$W(\phi_k) = \sum_{k=1}^{K} \phi_k N_k^R,$$
$$D(\phi_k) = \frac{P(\phi_k)}{W(\phi_k)}.$$

When the objective is $\max_{\{\phi_k, N_k^R\}} P(\phi_k)$, we can transform it into a 0-1 knapsack problem. Also it is shown that a FPTAS exists by using a dynamic programming algorithm for a 0-1 knapsack problem [24]. Thus, we have $P(\phi_k) \geqslant \frac{P(\phi_k^{\text{opt}})}{1+\epsilon}$, given ϕ_k^{opt} to be the optimal solution and ϵ to be an allowed error of solution.

Also $W(\phi_k)$ increases with ϕ_k since the weight increases when more items are placed into the knapsack. Thus, we have

$$D(\phi_k) = \frac{P(\phi_k)}{W(\phi_k)} \geqslant \frac{P(\phi_k)}{W(\phi_k^{\text{opt}})} \geqslant \frac{P(\phi_k^{\text{opt}})}{W(\phi_k^{\text{opt}})(1+\epsilon)}. \tag{16}$$

So a $(1+\epsilon)$-approximation solution to problem (P3) exists.

The FPTAS algorithm can be characterized as Algorithm 1.

Algorithm 1 FPTAS algorithm for problem (P3)

Step 1: Initialize $\epsilon > 0$, N_u to be the number of devices, the set of allocated devices $S_d(0) = \emptyset$, given $S_d(k)$ represents the kth iteration.
Step 2: Rank the density value of each device in a decreasing order $k = 1, 2, \ldots, N_u$.
Step 3: If $D(\phi_k) > D(\phi_{k-1})$ and all the constraints in problem (P3) are satisfied, we allocate resources to device k, so $S_d(k) = S_d(k-1) \cup \{k\}$. Otherwise, we do not allocate resources to device k, so $S_d(k) = S_d(k-1)$. Based on the above principles, we use a dynamic programming algorithm to compute the set of devices $S_d(N_u)$ [24].
Step 4: If $S_d(N_u) = \emptyset$, return 'No solution'; Otherwise, return the solution.

5 Simulation results

We experimented with a hardware platform consisting of a NI-PXIe 1085 device and three USRP-RIO-2943 (universal software radio peripheral-radio reconfigurable input/output). The hardware set is shown in Figure 4. NI-PXIe 1085 is a computer-based platform for data transmission and graphic display. The two USRP RIO-2943 (RIO1 and RIO2) are used to simulate four different transmitters that need to be identified. One USRP RIO-2943 (RIO3) is used to be a receiver that is responsible for receiving signals from the four transmitters. The data from the hardware platform is collected, and then imported to Matlab software for subsequent simulation and analysis. We compare the analytic results in the three scenarios, including traditional authentication algorithm, authentication algorithm with RF fingerprinting, and without any authentication algorithms. As a low-cost software-defined radio support device, USRP-RIO enables a wide range of applications, including broadcasting, mobile, GPS, WiFi, ISM FM, TV signals, and more. So we can use USRP-RIO to simulate diverse IoT devices. The user interface of USRP-RIO is shown in Figure 5, and users can set the parameters in the simulation through this interface. Specifically, this interface is composed of hardware configuration part, waveform display part and spectrum display part. The primary parameters of USRP-RIO are shown in Table 1. The parameters include central frequency, FFT points, sampling rate, sampling points, bandwidth, and modulation method.

In the simulation, we set the parameters as follows [25]: the amount of wireless bandwidth per OFDM subcarrier $W_k^s = 15$ kHz, the number of subcarriers per resource block $N_k^s = 12$, the consumption of energy per computation unit $\eta_k = 2$ mW, the amount of authentication data uploaded to edge servers $S_k^A = 1000$ kB, the amount of service data $S_k^S = 100$ MB, the number of required edge computation units for service tasks $U_k^S = 0.05$, the number of required edge computation units for authentication $U_k^A = 0.05$, the number of required local computation units for service tasks $U_k^l = 2$, the computation frequency of

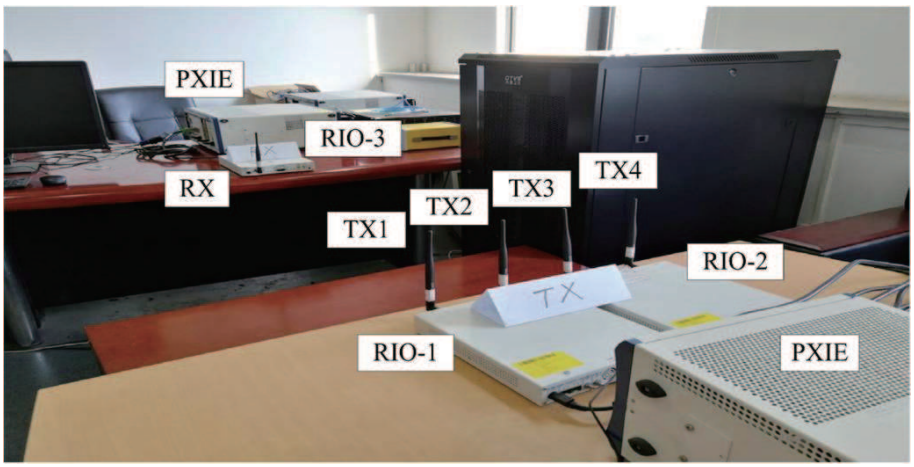

Figure 4 (Color online) Experimental settings for data collection.

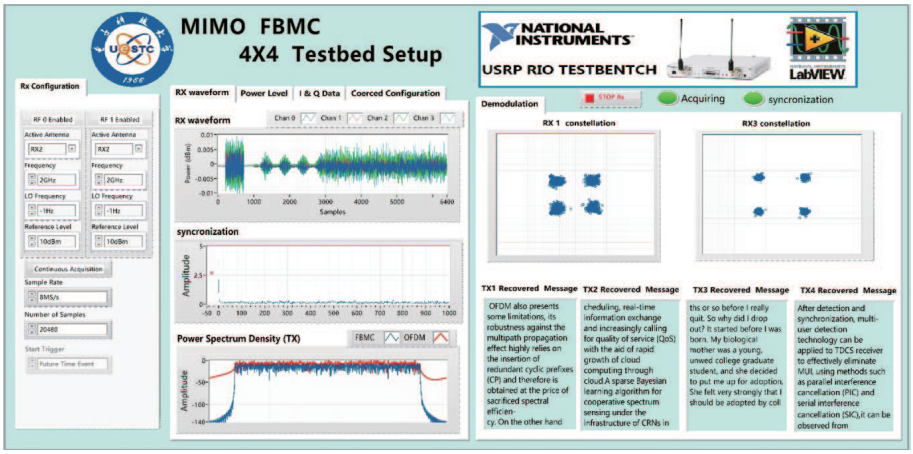

Figure 5 (Color online) User interface of USRP-RIO.

Table 1 Parameters of USRP-RIO interface

Parameter	Value
Central frequency	2 GHz
FFT points	256
Sampling rate	1 MHz
Sampling points	10000
Bandwidth	1 MHz
Modulation method	QPSK

each local computation unit $f_k^l = 500$ M cycles/s, the computation frequency of each edge computation unit $f_k^e = 20000$ M cycles/s. We also assume that the amount of authentication data and the number of required edge computation units for authentication are the same for both traditional algorithm and RF fingerprinting method.

Additionally, we consider various traditional authentication methods in the network layer and application layer, including WEP, WPA, WPA2. All these traditional methods need users to set passwords. If any of the passwords are weak and gullible, a hacker can easily crack them no matter how secure the authentication scheme is. Weak passwords are considered as one of common weaknesses for a traditional authentication algorithm. So we set $\kappa = 17\%$ for a traditional authentication algorithm, since 83% of passwords are simple to be cracked according to Avast's report [13]. On the other hand, for the RF fingerprinting algorithm, the values of κ depend on channel characteristics. In the simulation, we consider the Rayleigh channels under the environments of home and office, and the key parameters of channels refer to Table 2.

Table 2 Parameters of Rayleigh channels

Environment	RDS (ns)	Number of paths	Sampling rate (Mbps)	Guard interval length (μs)
Home	50	16	20	0.8
Office	100	16	20	0.8

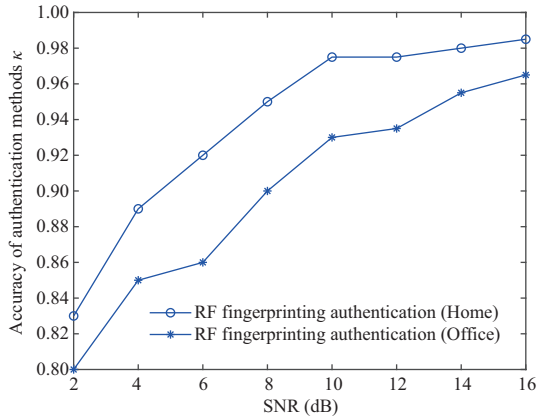

Figure 6 (Color online) Accuracy of authentication methods κ by RF fingerprinting authentication (Home vs. Office).

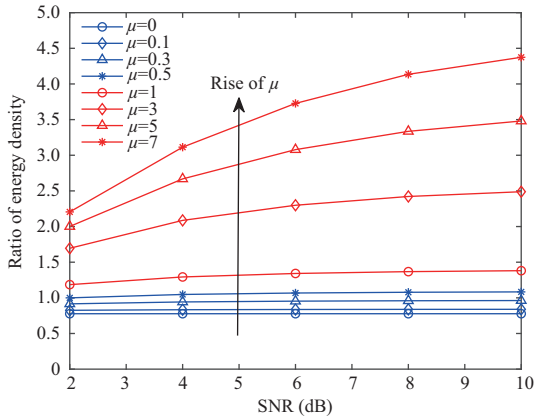

Figure 7 (Color online) Energy density by RF fingerprinting authentication vs. traditional authentication.

5.1 Benefit of the authentication in reducing energy density

In this subsection, we compare the efficient energy density of our proposed RF-fingerprinting scheme to that of traditional ones. In Figure 6, we show that the authentication accuracy κ of our proposed scheme depends on the values of SNRs as well as channel environment. With the rise of SNR, the values of κ under the environments of both home and office also increase. Additionally, the accuracy of authentication κ in the office environment is lower than that at home, since the former has a larger value of root-mean squared delay spread (RDS) than the latter, lowering the accuracy of RF fingerprint recognition.

Because any authentication scheme cannot fully guarantee that all unauthenticated users are rejected to sneak into the network, particular unauthenticated users may consume extra energy by transmitting and processing non-compliant data. For fairness, we define efficient energy density as the energy density that is used only to transmit the data of authenticated users. The efficient energy density can be transformed from (10) into

$$\bar{D}\left(\phi_k, N_k^R\right) = \frac{E\left(\boldsymbol{K}_A\right) + E\left(\bar{\boldsymbol{K}}_U\right)}{N^R\left(\boldsymbol{K}_A\right)}. \tag{17}$$

As shown in Figure 4, our experiment is conducted in the office, and thus we consider the accuracy of authentication κ under such an environment. In Figure 7, we compare the ratio of efficient energy density consumption by using the scheme with traditional authentication to that by using RF-fingerprinting authentication in various settings of μ and SNR. It is interesting to find that the ratio of efficient energy density increases with the rise of SNR since the performance of an RF-fingerprinting authentication algorithm relies on the level of noise. A high level of noise may degrade the accuracy of the RF-fingerprinting recognition model. It should be noted that the ratio of efficient energy density consumption increases with the rise of μ, indicating that our proposed authentication outperforms the traditional authentication when detecting unauthenticated users.

5.2 Analysis of efficient energy density with the authentication based algorithm

In this subsection, we investigate the efficient energy density with our proposed RF-fingerprinting authentication based algorithm in various settings of SNR, μ as well as N_T^R.

Figure 8 shows the change of efficient energy density with different SNRs. It illustrates that the efficient energy density keeps increasing with the rise of SNR when $\mu \leqslant 1$, while the efficient energy density fluctuates with the rise of SNR when $\mu > 1$. The reason is that the efficient energy density relies on both SNR and authentication process's consumption. When $\mu \leqslant 1$, the number of required resource blocks decreases with the rise of SNR. The data rate for the transmission of a resource block

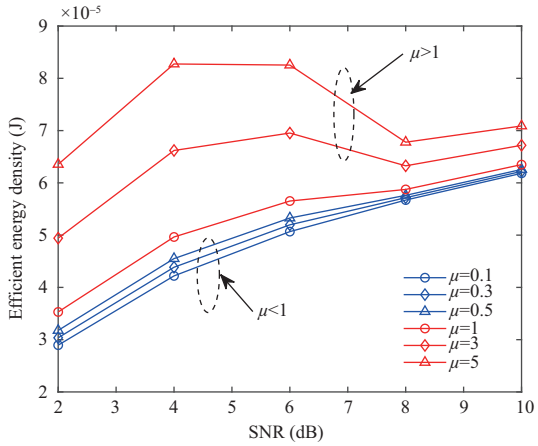

Figure 8 (Color online) Efficient energy density with various SNRs.

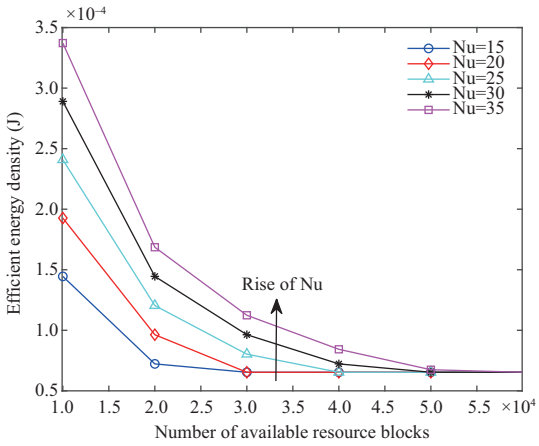

Figure 9 (Color online) Efficient energy density with various number of resource blocks N_T^R.

increases at high SNR, but the total data rate keeps constant as the required number of resource blocks is reduced. Thus, the energy consumption that is determined by data rate is constant at high SNR. The efficient energy density increases with SNR when $\mu \leqslant 1$ as the required number of resource blocks decreases. When $\mu > 1$, the energy consumption decreases because the authentication algorithm can detect unauthenticated users and terminate the data service for them. It should be noted that both the numerator of energy consumption and the denominator of the number of resource blocks decrease with the rise of SNR, so the ratio of efficient energy density fluctuates with the various SNR settings.

Figure 9 shows that the efficient energy density decreases with the rise of N_T^R. At a low N_T^R, the constraint of resource blocks becomes a primary factor. The efficient energy density is relatively high as a few users have no choice to process their data locally, which attributes to limited resource blocks at the edge server. When N_T^R increases to a high value, the constraint of resource blocks is not a primary factor any longer, and thus the efficient energy density converges to a constant.

5.3 Complexity of the authentication based algorithm

In this subsection, we discuss the computation complexity in the process of RF-fingerprinting authentication. CNN is considered as the primary algorithm of detecting unauthenticated users. It is proved that the CNN model would increase time complexity due to the forward and backpropagation process. The process can mathematically be shown as $O(\sum_{m=1}^{M} c_m s_m^2 f_m \rho_m^2)$, where M represents the number of convolutional layers, m represents the index of a convolutional layer, c_m represents the size of input feature map at the mth layer, s_m represents the size of filters at the mth layer, f_m represents the number of filters at the mth layer, and ρ_m represents the size of output feature map at the mth layer [26].

In the CNN model, the time complexity is linearly increasing with the number of convolutional layers M, and this complexity can be shown as

$$TC = O\left(\sum_{m=1}^{M} c_m s_m^2 f_m \rho_m^2\right). \tag{18}$$

As shown in Figure 10, the total delay U_k^A is linearly increasing with the number of layers M, which is in accordance with (19). A CNN model with many layers requires more computing units than that with a few layers. Figure 10 indicates that the total delay increases with the amount of service data S_k^A because a large amount of service data takes long time for transmission and processing. Thus, the total delay is dependent on authentication scheme, data transmission, as well as data processing.

6 Conclusion

This article has proposed an RF fingerprinting based authentication method, which can effectively prevent unauthorized users from occupying wireless resources. Based on the authentication scheme, we

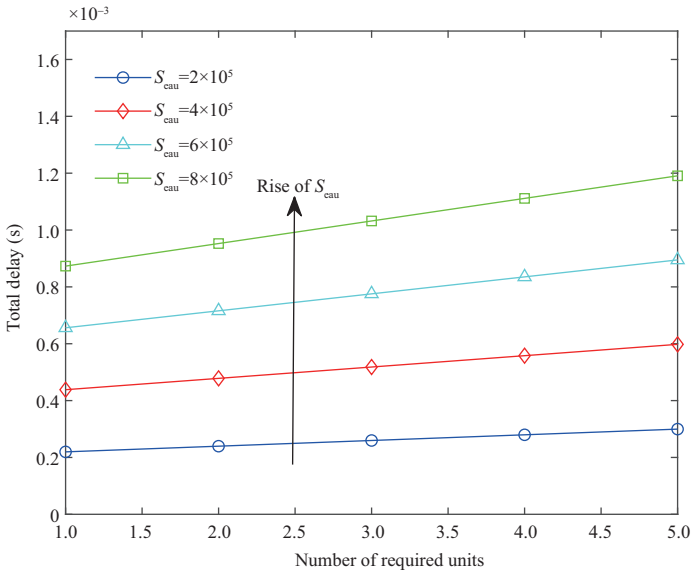

Figure 10 (Color online) Total delay of authentication based algorithm.

address resource allocation algorithms by solving an NP-hard optimization problem to minimize efficient energy density. We also present an FPTAS based approximation scheme for optimal resource allocation in consideration of user authentication. The detailed analysis and extensive experiments show that our proposed scheme can dramatically reduce the efficient energy density compared to traditional authentication methods. The results also demonstrate the robustness of the proposed scheme in various settings of SNRs and N_T^R. Additionally, the time complexity of total delay is also analysed. The result illustrates that the delay linearly increases with the number of layers in the CNN-based authentication model.

Acknowledgements This work was partially supported by Science and Technology Program of Sichuan Province (Grant No. 2021YFG0330), Intelligent Terminal Key Laboratory of SiChuan Province (Grant No. SCITLAB-0001), Fundamental Research Funds for the Central Universities (Grant No. ZYGX2019J076), National Natural Science Foundation of China (Grant No. 61971092), and Province Sichuan Foundation for Distinguished Young Scholars (Grant No. 2020JDJQ0023).

References

1 Liu X, Jia M, Zhang X, et al. A novel multichannel Internet of Things based on dynamic spectrum sharing in 5G communication. IEEE Internet Things J, 2019, 6: 5962–5970
2 Shen S Q, Zhang K, Zhou Y, et al. Security in edge-assisted Internet of Things: challenges and solutions. Sci China Inf Sci, 2020, 63: 220302
3 Li S L, Zhai D, Du P F, et al. Energy-efficient task offloading, load balancing, and resource allocation in mobile edge computing enabled IoT networks. Sci China Inf Sci, 2019, 62: 029307
4 You X H, Wang C-X, Huang J, et al. Towards 6G wireless communication networks: vision, enabling technologies, and new paradigm shifts. Sci China Inf Sci, 2021, 64: 110301
5 Liang B, Fan R, Hu H, et al. Nonlinear pricing based distributed offloading in multi-user mobile edge computing. IEEE Trans Veh Technol, 2021, 70: 1077–1082
6 Kakkavas G, Tsitseklis K, Karyotis V, et al. A software defined radio cross-layer resource allocation approach for cognitive radio networks: from theory to practice. IEEE Trans Cogn Commun Netw, 2020, 6: 740–755
7 Cui J, Wei L, Zhang J, et al. An efficient message-authentication scheme based on edge computing for vehicular ad hoc networks. IEEE Trans Intell Transp Syst, 2019, 20: 1621–1632
8 Peng L, Hu A, Zhang J, et al. Design of a hybrid RF fingerprint extraction and device classification scheme. IEEE Internet Things J, 2019, 6: 349–360
9 Das R, Gadre A, Zhang S, et al. A deep learning approach to IoT authentication. In: Proceedings of IEEE International Conference on Communications (ICC), Kansas City, 2018. 1–6
10 Hao P, Wang X. Integrating PHY security into NDN-IoT networks by exploiting MEC: authentication efficiency, robustness, and accuracy enhancement. IEEE Trans Signal Inf Process over Networks, 2019, 5: 792–806
11 Chen D, Zhang N, Qin Z, et al. S2M: a lightweight acoustic fingerprints-based wireless device authentication protocol. IEEE Internet Things J, 2017, 4: 88–100
12 Xu D Y, Ren P Y, Ritcey J A. Independence-checking coding for OFDM channel training authentication: protocol design, security, stability, and tradeoff analysis. IEEE Trans Inform Forensic Secur, 2019, 14: 387–402
13 Tobi W. Weak password policies: a lack of corporate social responsibility. J Colloq Inf Syst Sec Educ, 2020, 8: 7–8
14 Radhakrishnan S V, Uluagac A S, Beyah R. GTID: a technique for physical device and device type fingerprinting. IEEE Trans Dependable Secure Comput, 2015, 12: 519–532
15 Fan X, Wang F, Wang F, et al. When RFID meets deep learning: exploring cognitive intelligence for activity identification. IEEE Wireless Commun, 2019, 26: 19–25
16 Wu Q, Feres C, Kuzmenko D, et al. Deep learning based RF fingerprinting for device identification and wireless security. Electron lett, 2018, 54: 1405–1407

17 Wang X, Wang X, Mao S. RF sensing in the Internet of Things: a general deep learning framework. IEEE Commun Mag, 2018, 56: 62–67
18 Youssef K, Bouchard L, Haigh K, et al. Machine learning approach to RF transmitter identification. IEEE J Radio Freq Identif, 2018, 2: 197–205
19 Jian T, Rendon B C, Ojuba E, et al. Deep learning for RF fingerprinting: a massive experimental study. IEEE Internet Things M, 2020, 3: 50–57
20 Sankhe K, Belgiovine M, Zhou F, et al. Oracle: optimized radio classification through convolutional neural networks. In: Proceedings of IEEE Conference on Computer Communications, 2019. 370–378
21 Chatterjee B, Das D, Maity S, et al. RF-PUF: enhancing IoT security through authentication of wireless nodes using in-situ machine learning. IEEE Internet Things J, 2019, 6: 388–398
22 Riyaz S, Sankhe K, Ioannidis S, et al. Deep learning convolutional neural networks for radio identification. IEEE Commun Mag, 2018, 56: 146–152
23 Roy D, Mukherjee T, Chatterjee M, et al. RFAL: adversarial learning for RF transmitter identification and classification. IEEE Trans Cogn Commun Netw, 2020, 6: 783–801
24 Cohen R, Katzir L. The generalized maximum coverage problem. Inf Processing Lett, 2008, 108: 15–22
25 Li X, Dang Y, Aazam M, et al. Energy-efficient computation offloading in vehicular edge cloud computing. IEEE Access, 2020, 8: 37632–37644
26 He K, Sun J. Convolutional neural networks at constrained time cost. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015. 5353–5360