IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 8, 2018, accepted June 2, 2018, date of publication June 15, 2018, date of current version July 6, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2844548

T2FA: Transparent Two-Factor Authentication

JILIANG ZHANG[1], (Member, IEEE), XIAO TAN[1], XIANGQI WANG[2], AIBIN YAN[3], AND ZHENG QIN[1]

[1]College of Computer Science and Electronic Engineering, Hunan University, Changsha 410082, China
[2]College of Mathematics and Computational Science, Hunan First Normal University, Changsha 410205, China
[3]School of Computer Science and Technology, Anhui University, Hefei 230601, China

Corresponding author: Aibin Yan (abyan@mail.ustc.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61602107, Grant 11726632, and Grant 61604001, in part by the National Natural Science Foundation of Hunan Province, China, under Grant 2018JJ3072, and in part by the Fundamental Research Funds for the Central Universities.

ABSTRACT Traditional username and password-based single-factor authentication is easy to deploy but vulnerable to dictionary attacks, snooping, and brute force attacks. Two-factor authentication (2FA) has been proposed to improve the security, where smart devices are used as the second authentication factor. However, the interaction between human and the smart device is required, which is inconvenient to users. In addition, an attacker is able to get the second authentication factor through fraud, thus invalidating current 2FA mechanisms. In order to solve these problems, we propose a transparent two-factor authentication (T2FA) based on physical unclonable function (PUF) and voiceprint. The second authentication authenticates the user's mobile phone through the PUF. The third one is to determine whether the login terminal and the user's mobile phone are in the same environment with the environment voiceprint. The second and third authentication in the second factor is completely transparent to users. Therefore, T2FA avoids the tedious interaction and provides the same high user experience satisfaction as the single-factor authentication and exhibits high security simultaneously. Moreover, the fraud is eliminated technically due to the transparency of the authentication.

INDEX TERMS Two-factor authentication, physical unclonable function (PUF), voiceprint.

I. INTRODUCTION

A. MOTIVATION

With the rapid development of e-commerce, more and more companies conduct business online. However, due to the openness, versatility and multi-service of the Internet, network security issues become more and more serious. According to the statistics from China Internet Network Information Center (CNNIC) [1], in 2015, 42.7% of Internet users encountered network security problems.

User authentication is to establish the trust between users and devices and has become the most forefront defense for cyber-security. The static single-factor authentication such as ''username + password'' has been used widely because it is easy to deploy without additional devices. However, the security of single-factor authentication depends on the password. Such authentication is effective in the early Internet, where remote access was not used widely and the attack pattern was single. Nowadays, Trojans are able to intercept the user's keyboard record and even decrypt the user's login account and password by collecting the location clicked by the mouse, thus breaking the password protection technology. In 2011, Chinese internet suffers the most serious user data leak in history. China's largest software programmers' web site China Software Developer Network (CSDN) was hacked, and account information for more than 6 million users was leaked and quickly spread via the Internet [2]. In 2015, accounts and passwords disclose became the second serious Internet security incident in computers and mobile phones, accounting for 22.9% [1]. In May 2016, Google announced to completely cancel the password [3]. Therefore, single-factor authentication has become increasingly unsuitable.

In order to enhance the security, Two-Factor Authentication (2FA) was proposed. 2FA is a method of confirming a user's claimed identity by utilizing a combination of two different factors: 1) something they know, 2) something they have, or 3) something they are, i.e., a combination of passwords and physical entities such as smart cards [4], mobile phones [5], tokens, or fingerprints. However, compared with the password-based single-factor authentication, two-factor authentication brings inconvenience to users when a physical entity is used as the second authentication factor, where many additional operation steps are added. For example, the dynamic token method is a one-time password and

provides the high security, but it requires carrying different tokens when the user visits different sites.

In addition, current 2FA approaches are being questioned. For example, McAfee and Guardian Analytics released a joint report titled "Dissecting Operation High Roller" [6]. It mentioned an international criminal group who used an automated operation to attempt to steal large sums of money through unauthorized and fraudulent transfers. Since criminal group implanted malicious software on the victim's computer, the password information could be easily stolen. The theft of passwords was then integrated into the process for automated attacks. Criminals even could manipulate the user's authentication process so that the two-factor authentication tokens used to verify bank account access authorization became useless. There are many other reports that the two-factor authentication is threatened, and even fraud can be achieved without any technology. In 2016, China Central Television (CCTV) reported a piece of news about telecom fraud on the two-factor authentication. Even with the two-factor authentication, criminals can still steal all assets of the victim easily through the unpopular business of China Mobile's online 4G card replacement [7]. The root cause of this event is that the current authentication process of the second factor involves the user interactions, and the criminals can obtain the second authentication factor by fraudulent means so as to complete the authentication process.

As discussed above, the traditional single-factor authentication faces a great threat that once the password is leaked, the user will have no security at all. In addition, although current two-factor authentications are able to improve security, the tedious interactions between users and entities are added, which not only reduces the user experience significantly, but also makes it vulnerable to fraud and other threats.

B. OUR CONTRIBUTIONS

In order to mitigate the above issues, we propose the concept of Transparent Two-Factor Authentication (T2FA), and propose the first T2FA technique based on physical unclonable functions [8] [9] and voiceprint. The second factor in the T2FA consists of two authentications: 1) legal verification of smart devices by using PUF; 2) the match of the same environment between the computer and the user's mobile phone when the user logs in. Without changing the user experience, we propose the PUF-based device authentication and the similarity comparison of the environment background sound between the mobile phone and the computer. In T2FA, the second factor is completely transparent to the user, which avoids the tedious interaction between the user and the device. As there is no need for users to participate in the authentication of the second factor, the human-machine interaction and the threat of user-oriented fraud can be eliminated technically. Therefore, T2FA is able to improve user experience and enhance the authentication security significantly.

The rest of this paper is organized as follows. Related work is elaborated in Section II. Section III gives a detailed introduction about our proposed PUF and voiceprint-based transparent two-factor authentication. The detailed experimental results and analysis are reported in Section IV. Finally, we conclude in Section V.

II. RELATED WORK

A. HARDWARE TOKEN

Time-based hardware tokens follow the standard of 2FA. The SecurID [10] is stored as the second factor in the dongle. However, this mechanism requires the interaction between the user and the hardware token. Besides, the server needs to provide each client with a separate hardware token, which makes the deployment of token expensive. It also requires each site having its own authentication devices and needs users to carry them. Therefore, such 2FA mechanism is only used in e-government, online banking and other similar fields.

B. SOFTWARE TOKEN

Google 2-Step authentication [11] is a typical software token mechanism based on user's phone and authentication code. The software authentication code is sent by short message service or an application running on the mobile phone. Such mechanism needs the user to copy the authentication code on the mobile phone to the login interface of the browser. Sound-Proof [12] uses the sound of the environment as an authentication factor, but it does not consider the legality authentication of the mobile phone and suffers security issues in a silent environment.

C. SHORT-RANGE WIRELESS COMMUNICATION

Short-range wireless communication implements the second factor authentication with Bluetooth, NFC or WiFi, which is able to reduce the interaction between the user and mobile phone. Bluetooth [13] is the most widely used technique. The browser and the phone generate a challenge-response pair via Bluetooth to compute the distance between them. However, the Bluetooth API is no longer supported by current mainstream browsers. NFC is a new short-range wireless communication method for smart devices. However, the mainstream browsers do not provide APIs to support NFC devices, and most of personal computers are not integrated with NFC. Moreover, NFC requires the user to hold his mobile phone to complete the two-factor authentication. Therefore, the interaction between the user and the mobile phone is complicated. WiFi communication requires that the computer located in the same network as the mobile phone. The computer needs to use additional software to generate an access node so that the user's mobile phone can be connected to the network where the computer is located. Such method requires that the mobile application continuously monitors the login request sent by the browser, which would degrade the performance of smart devices.

D. SHORT-RANGE ULTRASONIC

The browser and mobile phone can communicate with each other by the short-range ultrasound [14] which can

be recognizable to the computer and not audible to the human ear. Due to the limited performance of the phone, this communication method can only produce highly directional near-field ultra-high frequency sound waves [15], and the signal attenuation is quite fast. When this technique is deployed for authentication, mobile phones cannot use external devices such as earphones, and such high-frequency sound wave will have a health impact on children and animals [16].

E. LOCATION INFORMATION

With the Global Position System (GPS) information, the server can detect whether the computer and phone are located in the same environment. GPS sensors have been deployed in mainstream smartphones but not in most of computers. Many APIs provided by the browser can obtain geolocation information to complete the login [17]. However, the geolocation information is not accurate. For example, when the device is located in a VPN network or an enterprise's large management network, the geolocation is not the real location of the device and can be obtained by attackers, which would result in an authentication failure.

As discussed above, 2FA has attracted much attention in academia and industry. However, due to the tedious interaction, it is inevitable to change the user experience and bring the security issues for authentication. This paper proposes a Transparent Two-Factor Authentication (T2FA) technique based on PUF and voiceprint, where the second authentication is set as the PUF to authenticate the smart devices, and the third authentication is to authenticate the voiceprint between the computer and the mobile phone. The legitimacy of the login request is judged by measuring the matching degree of the environmental sound. Since T2FA is completely transparent to the user, the user experience can be improved greatly and the threat of user-oriented fraud also can be technically eliminated.

III. PUF AND VOICEPRINT-BASED T2FA

The traditional single-factor authentication is based on username and password, which is easy to deploy but vulnerable to dictionary attacks, snooping and brute force attacks. Although using different passwords on different websites can improve the security, it brings a lot of trouble to users in terms of memory. In the urgent need to design a more secure authentication mechanism, two-factor authentication (2FA) comes out, combining the password with the entities such as credit card, mobile phone, token or fingerprint. However, as visiting different sites often require different tokens, when you need to visit many sites at the same time, carrying a long list of tokens will be very troublesome. Therefore, compared with the single-factor authentication, two-factor authentication will bring more inconvenience to users when using different physical entities as authentication factors. Besides, the interactive between human and entities may allow an attacker to obtain the second factor through fraud.

PUF and voiceprint-based T2FA aims to add the authentication of mobile phones and the feature comparison of environmental background sounds without changing the user experience. The authentication is completely transparent to the user, which means that the user only needs to enter the user name and password. This not only enhances the security of the single-factor authentication, but also addresses the security issue that the traditional two-factor authentication requires the user interaction. Therefore, T2FA owns the anti-fraud ability that previous 2FA techniques do not have.

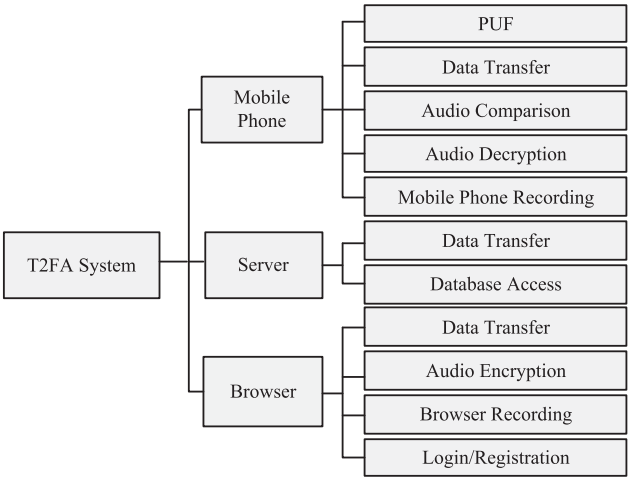


FIGURE 1. Functional modules in our proposed T2FA system.

A. DESIGN OF T2FA

T2FA is composed of login/registration module, browser recording module, mobile phone recording module, PUF module, audio encryption and decryption module, data transmission module, database access module and audio comparison module. The functional modules in the T2FA system are shown in Figure 1.

1) Login/Registration Module. This module is responsible for the authentication and transmission of input information when logging in or registering.
2) Browser Recording Module. This module is responsible for collecting the environment sound around the user's computer after the server passes the user name and password authentication.
3) Mobile Phone Recording Module. This module is responsible for collecting the environment sound around the user's mobile phone after the server passes the user name and password authentication.
4) PUF Module. This module is responsible for the authentication of the user's mobile phone after the server passes the user name and password authentication.
5) Audio Encryption and Decryption Module. This module is responsible for encrypting the audio files collected by the browser and transmitting the encrypted file to the server.
6) Data Transfer Module. This module is responsible for transferring data and instructions between browser, server and mobile phone.

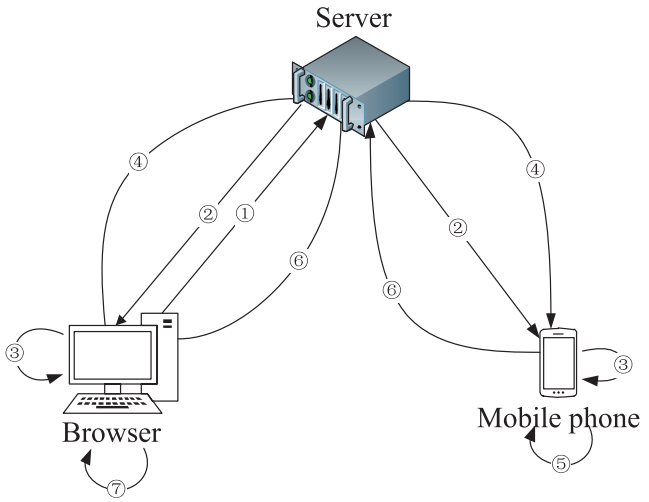


FIGURE 2. Authentication flow of T2FA.

7) Database Access Module. This module is responsible for querying or inserting data in the database during user login or registration.
8) Audio Comparison Module. This module is responsible for extracting the mark information in the two audio time domains and comparing the similarity of the two pieces.

Current two-factor authentication techniques are non-transparent. This paper propose the first transparent two-factor authentication based on PUF and voiceprint. The function of the browser side is implemented for the login and registration, environment sound recording, audio file encryption and data transmission. The server side is used for the authentication of mobile phones, database access, and data transmission. At the same time, the server stores the PUF's challenge-response pairs or PUF's delay parameters which are obtained by machine learning such as deep neural network for authentication. The mobile phone is used for recording environment sound, audio decryption, data transmission and audio similarity comparison. The whole authentication flow is shown in Figure 2.

1) The browser sends the username and password to the server.
2) The server verifies whether the user name and password are legal. If legal, the server will send a PUF challenge to the mobile phone for authentication. If the PUF response equals with the response stored in the server, the record command will be sent to the browser and mobile phone. If illegal, the server will give a warning message to the browser.
3) The browser and phone start recording 3 seconds environment sound, respectively.
4) The browser encrypts and sends the recorded environment sound file to the server, and then the server transmits the encrypted environment sound file to the mobile phone.
5) The mobile phone decrypts the received environment sound file and compares the similarity between the two audio files.
6) The mobile phone sends the result of audio similarity comparison to the server. Then the server transmits the result to the browser.
7) The browser determines whether the login is admitted.

The microphone can sample the audio with the WebRTC API in HTML5. Once the PC starts recording, the back-end application of user's phone is activated to sample the audio for the same length of time. A time synchronization protocol runs on the two devices and server. After the recording ends, the devices adjust the time-stamp of their respective audio to synchronize the time. Then the computer encrypts the audio and sends it to the phone through the server. The phone decrypts the audio and compares it with the local sampled audio. When the similarity is higher than the threshold, the mobile phone will notify the server that the login is legal. This process is completely transparent to the user. In the authentication of mobile phones, we propose to use the lightweight and highly secure PUF as the authentication source.

B. VOICEPRINT

The similarity computation for two sound segments is similar to the audio fingerprint and automatic media retrieval. For media retrieval, it may be detrimental to the match of the similar audio in the database when using a noisy audio to compare. Therefore, we can only extract the specific attributes related to the noisy audio and use them to compare with the sample that needs to be matched. When the background sound is noisy or thin, the extracted attributes must be robust. In the automatic media retrieval, high-frequency spectral coefficients, microwaves, and peak frequencies can be used as robust attributes. Because of the time misplacement problem, these attributes should focus on the frequency domain in which the audio samples are located. We compare two aligned audio samples by extracting relevant information from their time domain, and calculate the similarity between two audio samples using cross correlation and error energy.

1) CROSS-CORRELATION

Cross-correlation is a common method used to calculate the similarity of audios sampled at the same time. We use x and y to represent n discrete points of the same two consecutive signals. The cross-correlation $C_{x,y}(\ell)$ of the two audio segments is calculated by the function $\log \ell \in [0, n-1]$:

$$C_{x,y}(\ell) = \sum_{i=0}^{n-1} x(i) \cdot y(i-1) \; (i<0 \mid i > n-1, y(i)=0) \qquad (1)$$

In order to accommodate the signal with different amplitudes, the expression of cross-correlation can be normalized to the Equation 2:

$$C'_{x,y}(\ell) = \frac{C_{x,y}(\ell)}{\sqrt{C_{x,x}(0) \cdot C_{y,y}(0)}} \qquad (2)$$

where $C_{x,y}(\ell)$ represents the autocorrelation coefficient. The range of normalization unit $C'_{x,y}$ is $[-1,1]$.

$C'_{x,y} = 1$ means that even if the amplitudes of the two signals are different, the two audio segments have the same graphic structure; $C'_{x,y} = -1$ means that the two audio segments have the same graphic structure but they are completely opposite signals; $C'_{x,y} = 0$ means that the two signals are irrelevant.

If the values of the two signals are unknown, we can use the absolute of the cross-correlation maximum value $\hat{C}_{x,y}(\ell) = \max_{\ell}(|C'_{x,y}(\ell)|)(0 \leqslant \hat{C}_{x,y}(\ell) \leqslant 1)$ as the similarity. In addition, the complexity of calculating $C_{x,y}(\ell)$ can be reduced to calculating $C_{x,y}(\ell) = F^{-1}(F(x)^* \cdot F(y))$, where $F()$ represents a Fourier transform, and * represents a complex pairing operation.

2) ERROR ENERGY

The error energy can be used to measure the similarity of waveforms, which is similar to those used to determine the orthogonality of functions in advanced mathematics. Assuming that the two signals are $x(t)$ and $y(t)$, we can choose a constant a to make $a \cdot y(t)$ approach $x(t)$. The error energy can be expressed using the integral of $(x(t) - a \cdot y(t))$ in the time domain. The constant a must be selected to ensure that the energy error is minimized. By taking the derivative and extreme value of the functions, we can know that it can meet the condition when a is the integral ratio of $x(t) \cdot y(t)$ to $y(t) \cdot y(t)$ in the time domain. The correlation coefficient between $x(t)$ and $y(t)$ is defined as P_{xy}. And the difference between $(P_{xy})^2$ and 1 is the relative error energy, i.e., the ratio of error energy to the integral ratio of $x(t) \cdot x(t)$ in the time domain. The numerator of the equation for P_{xy} is the integral of $x(t) \cdot y(t)$ in the time domain. The denominator is the square root of the integral of the $x(t)^2$ and $y(t)^2$ in the time domain. It can be mathematically proved that the modulus of the numerator is smaller than the denominator, that is, the modulus of the correlation coefficient P_{xy} will not be greater than 1. Since the energy is fixed for a signal, the magnitude of the correlation coefficient P_{xy} is only determined by the integral of $x(t) \cdot y(t)$ in the time domain. If the amplitude and time of two waveforms are independent, we can get $x(t) \cdot y(t) = 0$, and the integral result is also 0. Therefore, when the correlation coefficient is 0, the similarity is the worst. When the correlation coefficient is 1, the error energy is 0, which indicates that the similarity between the two signals is good and linearly related.

C. PUF FOR AUTHENTICATION

As a new hardware security primitive, Physical Unclonable Functions (PUFs) have been extensively studied in academia in recent years [8] [9]. The industry has also developed corresponding products based on PUF [18]. PUF makes use of the fabrication variation to uniquely identify chips produced by the same manufacturing process. When a challenge is input to PUF, it will generate a response. Even with the complex manufacturing facilities, it is impossible to create another system with the same challenge-response behavior. Although

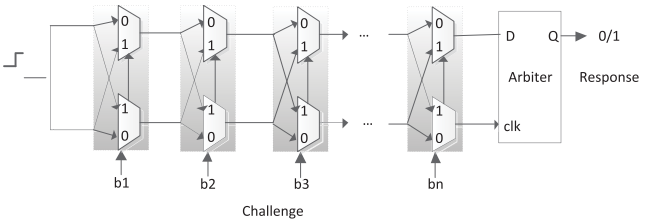


FIGURE 3. The structure of Arbiter PUF.

it is difficult to define all kinds of PUFs at the moment, the PUF should meet the following requirements [8]:

- Reliability and unpredictability. PUF responses are random and unpredictable, but remain unchanged with repeated measurements under the same challenge.
- Unclonable. It is impossible to get the response when given a specific challenge without physically accessing the PUF. In other words, given a PUF, it is not feasible for an attacker to rebuild a PUF that satisfies all of the challenges mapped to the corresponding responses. This is determined by the uncontrollable process variation.
- Tamper-proof. Invasive attacks on PUF will destroy the structure of PUF, which can be easily detected.

Due to these unique attributes, PUFs have been proposed for IP protection [19] [20], FPGA security [21], device authentication [22] and software security [32] [33]. Since the first optical PUF [23] was proposed in 2002, many PUFs such as Arbiter PUF [24], RO PUF [22] [25] and SRAM PUF [26] have been proposed. Arbiter PUF, as a classic strong PUF, has a large number of challenge-response pairs (CRPs), which is suitable for device authentication. The structure of Arbiter PUF is shown in Figure 3, two parallel multiplexer chains share the same input, and the outputs are respectively connected to the D and the clock input of a flip-flop. The input uses the step signal, and the selection of the multiplexer chain forms the challenge bits $b_1 \sim b_n$. The signal b_i determines whether the step signal at the i-th order is transmitted along the original multiplexer chain or interchanged to another multiplexer chain. The delay difference between the two multiplexer chains determines whether the step signal will reach the D input or the clock input first, resulting in logic 1 or logic 0 being latched, respectively. The latched value can be used as a PUF response bit.

The most common way to authenticate device with the PUF is to collect a large number of CRPs at the test step and then store the CRPs on the server [22]. As shown in Figure 4, a PUF is embedded into the device A, and CRPs are collected and stored in a secure database. As the PUF response is unique and unpredictable for each device, we can simply compare a regenerated PUF response with prestored response in database when using with same challenge for authentication. In order to protect against man-in-the-middle attacks, the used CRPs will be deleted from database. However, each PUF has a large number of CRPs, which will have a very high demand on server storage to store the CRPs for all devices. Therefore, we do not recommend using this common

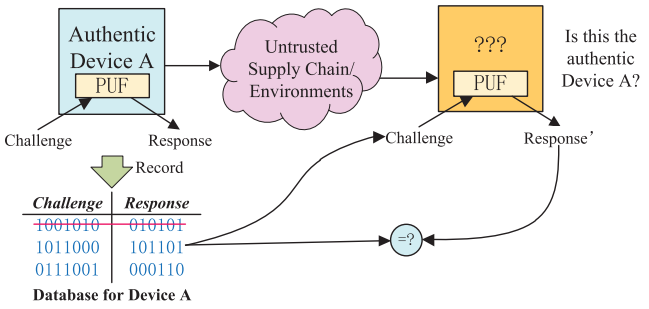


FIGURE 4. Traditional PUF-based authentication.

authentication method. In what follows, we will introduce a low storage overhead PUF-based authentication method.

First, we model the PUF with a machine leaning algorithm and then store the parameters instead of CRPs on the server, which incurs negligible storage overhead. If the delay of all blocks in a PUF path is known, it is easy to calculate the response after a given challenge. However, in practice, it is difficult to measure the delay of each block. Therefore, we use machine learning algorithms to simulate the delay of each block in the path of the PUF, i.e., the machine learning technique is used to build a software model for the PUF to simulate it's challenge-response behavior so as to predict the random response of the PUF. In this paper, Arbiter PUF is taken as an example and logistic regression is used to model the PUF. The details are as follows.

The structure of Arbiter PUF is shown in Figure 3. Challenge C is generated by external control bits, and the output 0 or 1 is usually used as the response R. The function of an Arbiter PUF can be expressed with the linear delay model [27]. The total delay of the signal is the accumulated delay of each stage. In this model, we can define the final delay difference Δ between the upper and lower parts as follows:

$$\Delta = (\vec{\omega})^T \vec{\phi} \tag{3}$$

The dimensions of $\vec{\omega}$ and $\vec{\phi}$ are both $k+1$. The parameter vector $\vec{\omega}$ represents the delay of each stage in the Arbiter PUF, while the feature vector $\vec{\phi}$ represents $(k+1)$-bit C in the PUF circuit. We use $\delta_i^{0/1}$ to represent the delay of stage$_i$, where 1 indicates that a cross occurred in the multiplexer, and 0 indicates that no cross occurred. Therefore, we can get the following vector:

$$\vec{\omega} = (\omega^1, \omega^2, \cdots \omega^k, \omega^{k+1})^T \tag{4}$$

where the $\omega^1 = \frac{\delta_1^0 - \delta_1^1}{2}$, $\omega^i = \frac{\delta_{i-1}^0 + \delta_{i-1}^1 + \delta_1^0 - \delta_1^1}{2}$, $i = 2, 3, \cdots, k$ and $\omega^{k+1} = \frac{\delta_k^0 - \delta_k^1}{2}$.

$$\vec{\phi}(\vec{C}) = (\vec{\phi}^{\,1}(\vec{C}), \cdots \vec{\phi}^{\,k}(\vec{C}), 1)^T \tag{5}$$

where the $\vec{\phi}^{\,\ell}(\vec{C}) = \Pi_{i=l}^{k}(1-2b_i)$, $\ell = 1, \cdots k$

Finally, the output value t of the Arbiter PUFs is determined by the sign function of the final total delay difference Δ. Here, the output of the PUF is 0 when $t = -1$; the output of the PUF is 1 when $t = 1$,

$$t = sgn(\Delta) = sgn((\vec{\omega})^T \vec{\phi}) \tag{6}$$

The Equation 6 indicates that vector $\vec{\omega}$ determines a separate hyperplane in all feature vector $\vec{\phi}$ space when $\vec{\omega}^T \vec{\phi} = 0$. When $t = -1$, all the feature vectors are on one side of the hyperplane. Conversely, when $t = 1$, all the feature vectors are on the other side. And the PUF can be predicted from the obtained hyperplane.

Logistic Regression (LR) [28] is a widely used machine learning algorithm. When the PUF is modeled with LR, each challenge C is assigned a probability $P(C, t|\vec{\omega})$ to produce 1 or -1. In order to facilitate modeling, here we will use -1 and 1 instead of 0 and 1. The probability is determined by the sigmoid function:

$$P(C, t|\vec{\omega}) = sigmoid(tf) = (1 + e^{-tf})^{-1} \tag{7}$$

Therefore, $f = 0$ can be used to determine the decision boundaries for equal output probabilities. For the training set M that gives the CRPs, we can minimize the log-likehood of the negative by constantly training the parameter $\vec{\omega}$:

$$\hat{\vec{\omega}} = arg\ min_{\vec{\omega}} \ell(M, \vec{\omega}) = arg\ min_{\vec{\omega}} \sum -\ln(\delta(tf(\vec{\omega}, C))) \tag{8}$$

Since there is no direct way to calculate $\vec{\omega}$, we use the iterative way to calculate $\vec{\omega}$. RProp [28] [29] performs the best for different gradient descent modes in machine learning. Therefore, here RProp gradient descent method is used:

$$\nabla \ell = (M, \vec{\omega}) = \sum_{(C,t)\in M} t(\delta(tf(\vec{\omega}, C)) - 1)\nabla f(\vec{\omega}, C) \tag{9}$$

As discussed above, Arbiter PUF response can be expressed as a linear function of the challenge. Therefore, we can model the Arbiter PUF with the machine learning algorithms. However, attackers can also model the PUF with collected CRPs [30]. In order to resist the modeling attack, many obfuscation techniques [31] have been proposed to obfuscate the map relationship between challenges and responses and hence increase the difficulty of modeling attacks. Obfuscation techniques are beyond the scope of this paper, more details please refer to obfuscation-related work.

D. SECURITY ANALYSIS

1) GUESSING ATTACKS

The security of T2FA stems from the inability of attackers to guess the sound in the victim's environment at the time of the attack. It is difficult for attackers to guess the sound recorded by the victims phone since the recorded sound is dependent on the environment and recorded time. The experimental results shown in Section IV prove that T2FA can discriminate the legitimate and fraudulent logins in noisy indoors, noisy outdoors and even in quiet indoor/outdoor areas. Therefore, T2FA is immune to guessing attacks.

TABLE 1. Test Environment.

Hardware Environment	Software Environment
PC: Intel Core^{TM} i7-5820k Memory: 8.00 GB	Windows 10 Professional Edition 64-bit
Mobile phone: Galaxy S6 edge + Android 6.0.1 Kernel Version: 3.10.61-7342950	Android 4.4.2
	Chrome 43.0.2351.3
FPGA: Xilinx Vertex 5	Visual studio 2010
PUF: Arbiter PUF	Eclipse Luna Service Release 1 (v4.4.1)

2) IMPERSONATION ATTACKS IN QUIET ENVIRONMENTS

In order to prevent the impersonation attack in quiet environments, we require the login user humming when the victim's environment is quiet (e.g., sleeping at night), which can improve the similarity significantly in the quiet environment. In Section IV, the experimental results show that the login success rate in quiet environment is 0%, but after humming in this environment, login success rate becomes 100%. Therefore, T2FA is immune to impersonation attacks in quiet environments.

3) MODELING ATTACKS

Machine-learning
based modeling attacks are one of well-known attacks for PUFs. They are exclusively applicable to strong PUFs such the Arbiter PUF. Strong PUFs has a publicly accessible CRP interface, which allows the simple collection of the large numbers of CRP that are required in this attack type during their learning phase [8]. Since the Arbiter PUF response can be expressed as a linear function of the challenge, it is possible for attackers to collect a huge number of CRPs to model the PUF behavior. However, current obfuscation techniques [31] can be used to obfuscate the map relationship between challenges and responses to resist modeling attacks.

IV. EXPERIMENTAL RESULTS

A. TEST ENVIRONMENT

We have implemented our proposed PUF and voiceprint-based transparent two-factor authentication system, where A 64 × 64 Arbiter PUF is implemented in a Xilinx Vertex 5 FPGA and consumes 184 LUTs. The hardware and software environment are shown in Table 1.

B. AUDIO SIMILARITY ALGORITHM TEST

1) AUDIO SIMILARITY

We use Node.js to build a simple server that supports the test of audio similarity algorithm. During the test, every time the user logs in, the mobile phone and the computer will record 3s audio, and the audio files collected in different environments are processed by the algorithm to obtain the similarity values of the audio files.

In what follows, we try to login in different scenarios to collect data such as noisy indoors (e.g., people playing music, video or chatting in the dorm room), noisy outdoors (e.g., eating in the canteen) and quiet indoor/outdoor (e.g., sleeping at night).

In the above three scenarios, we consider the following cases:

(1) For noisy indoors:

- The mobile phone is close to the computer.
- The mobile phone is in the pocket of the tester who tries to log in the computer.
- The distance between the phone and the computer is 1m.
- The distance between the phone and the computer is 2m.
- The distance between the phone and the computer is 3m.
- The distance between the phone and the computer is more than 4m.

(2) For noisy outdoors:

- The mobile phone is close to the computer.
- The mobile phone is in the pocket of the tester who tries to log in the computer.
- The distance between the phone and the computer is more than 1m.

(3) For quiet indoor/outdoor:

- The mobile phone and the computer are in any position.
- The mobile phone is close to the computer (with the login user humming).

In the noisy indoor environment, we tested six different cases separately. We got the similarity values between the mobile phone and the computer in the six cases of close contact, 1m, 2m, 3m, more than 4m and the mobile phone in the pocket. The test results are given in Figure 5, which shows that close contact can obtain the best similarity, and similarity decreases as the distance increases.

In the noisy outdoor environment, we tested three different cases (close contact, more than 1m and the mobile phone in the pocket) separately to compute the similarity values. We can see from the test results in Figure 6 that in pocket and close contact have good similarity.

Similarly, in a quiet environment, the test results are shown in Figure 7, which shows that humming can improve the similarity significantly in the quiet environment.

2) THRESHOLD FOR SIMILARITY ALGORITHM

Taking into account the indoor environment such as in your own home or office, the mobile phone may not carry with you. Therefore, we set 2m as a legitimate login range. In the outdoors, such as canteens, lecture hall and so on, the mobile phone should be carried around. And we set the login range within 1m as the legal distance. The quiet environment login should be rejected because it is easily guessed and imitated by

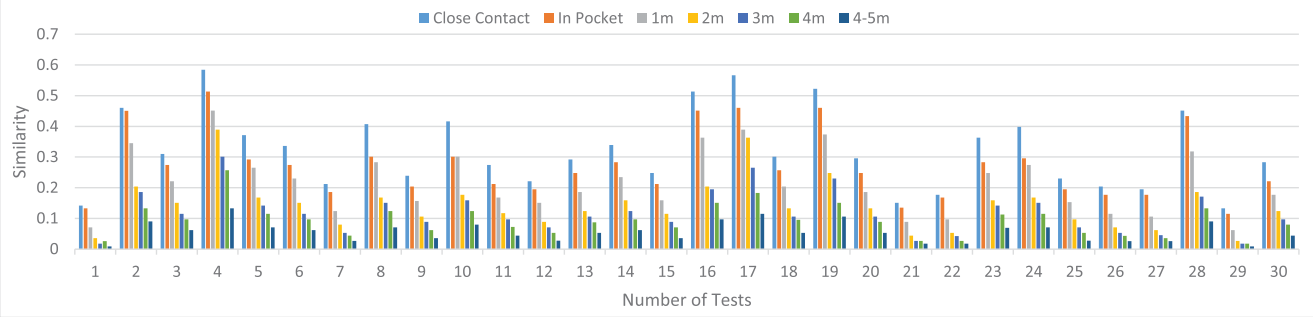


FIGURE 5. Similarity data in noisy indoors.

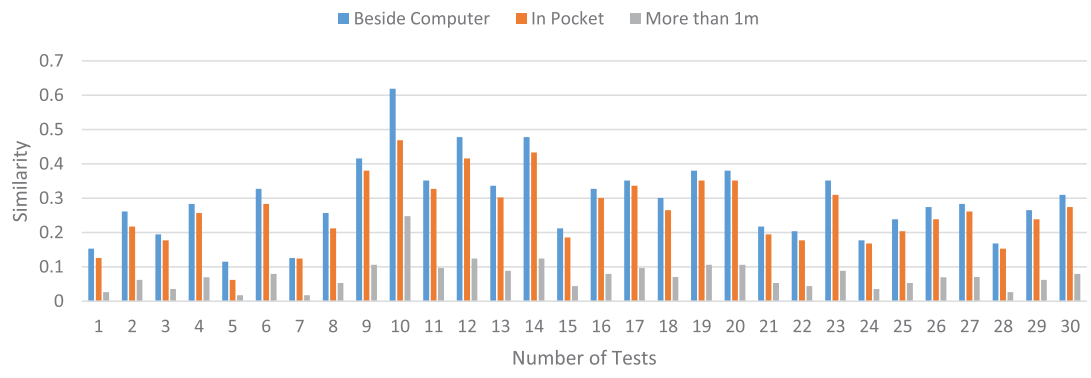


FIGURE 6. Similarity data in noisy outdoors.

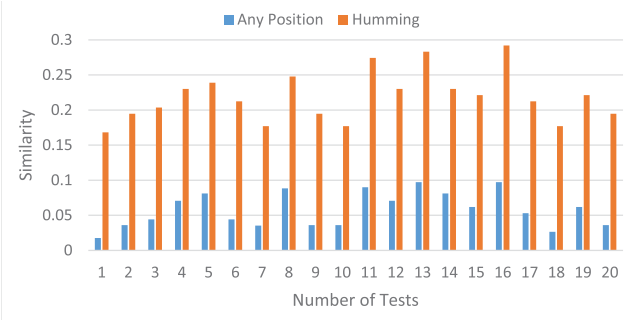


FIGURE 7. Similarity data in quiet indoor/outdoor.

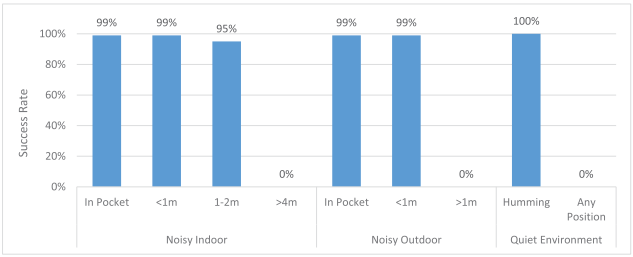


FIGURE 8. Success rate in the noisy indoor, noisy outdoor and quiet environment.

illegal visitors in this environment. While refusing to login in the quiet environment, we ask the user to hum a song to make noise in order to guarantee a legitimate login.

Therefore, the ideal threshold should meet following requirements:

- In the indoor environment, when the distance between the mobile phone and the computer is less than 2m, the success rate of login is 100%; and the success rate of login is 0% when the distance is more than 4m.
- In the outdoor environment, when the distance between the mobile phone and the computer is less than 1m, the success rate of login is 100%; and the success rate of login is 0% when the distance is more than 1m.
- Regardless of indoor and outdoor, the quiet environment login success rate is 0%. After humming in this environment, login success rate becomes 100%.

We conducted 50 tests separately for noisy indoors, noisy outdoors and quiet indoor/outdoor areas, and the similarity threshold t is set to 0.11504425. As shown in Figure 8, the test results satisfies the following conditions:

- In the noisy indoor environment, when the distance between the phone and the computer is less than 1m or the mobile phone is in the pocket, the success rate is 99%; when the distance is greater than 1m and less than 2m, the success rate is 95%; when the distance is greater than 4m, the success rate is 0%.
- In the noisy outdoor environment, when the distance between the phone and the computer is less than

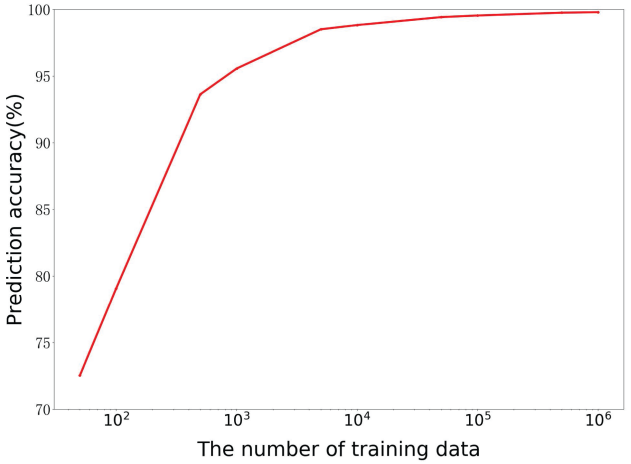


FIGURE 9. LR on a 64 × 64 Arbiter PUF.

1m or the mobile phone is in the pocket, the success rate of login is 99%; when the distance is greater than 1m, the success rate is 0%.

- Regardless of indoor or outdoor, the success rate for quiet environment is 0%, and after humming in this environment, the success rate becomes 100%.

C. MODELING ARBITER PUF

As shown in Figure 9, for a 64 × 64 arbiter PUF, we can achieve 95% prediction accuracy with only 650 CRPs and the training time is less than 1s, and achieve 99% prediction accuracy with less than 3000 CRPs in 1s, and 99.9% prediction accuracy with about 20000 CRPs in about 2s. Therefore, in our proposed T2FA, storing the model parameters of PUFs on the server for device authentication not only reduces the storage overhead but also improves the authentication efficiency.

V. CONCLUSION

This paper proposes the concept of transparent two-factor authentication and then proposes the first PUF and voiceprint-based transparent two-factor mechanism. The first authentication factor is still the traditional username and password. The second authentication factor is the mobile phone, which includes two novel authentication ways: 1) PUFs are used to authenticate the mobile phone; 2) the same environment verification between the computer and the mobile phone. When the user logs in, the browser sends a login request to activate the mobile phone. Both of them use their respective microphones to record the environment noise, and adjust the time-stamp to synchronize the time. The mobile phone compares the similarity of the two audios to determine whether the browser is in the same environment as the mobile phone and then determines whether the login is valid. The second factor authentication is completely transparent to the user, avoiding the tedious interaction between the user and the device. Therefore, our proposed T2FA exhibits the anti-fraud ability with good application prospect. The effectiveness of our proposed T2FA is further proved with detailed experiments.

ACKNOWLEDGMENT

Special thanks to 2017 CCF-IFAA RESEARCH FUND for greatly supporting the writing of the paper. To know more about IFAA: www.ifaa.org.cn/en.

REFERENCES

[1] (2016). *Statistical Report on Internet Development in China*. [Online]. Available: https://cnnic.com.cn/IDR/ReportDownloads/201604/P020160419390562421055.pdf

[2] *Chinese Internet Suffers the Most Serious User Data Leak in History*. Accessed: Jun. 2018. [Online]. Available: https://blogs.forcepoint.com/security-labs/chinese-internet-suffers-most-serious-us

[3] *Google Plans to Bring Password-Free Logins to Android Apps by Year-End*. Accessed: Jun. 2018. [Online]. Available: https://techcrunch.com/2016/05/23/google-plans-to-bring-password-free-logins-to-android-apps-by-year-end/

[4] S. Kumari and M. K. Khan, "More secure smart card-based remote user password authentication scheme with user anonymity," *Secur. Commun. Netw.*, vol. 7, no. 11, pp. 2039–2053, Nov. 2014.

[5] Z. Siddiqui, A. H. Abdullah, M. K. Khan, and A. S. Alghamdi, "Smart environment as a service: Three factor cloud based user authentication for telecare medical information system," *J. Med. Syst.*, vol. 38, no. 1, p. 9997, Jan. 2014.

[6] *Dissecting Operation High Roller*. Accessed: Jun. 2018. [Online]. Available: https://www.finextra.com/finextra-downloads/newsdocs/high-roller.pdf

[7] *Verification Code Scam in Telecommunication Fraud*. Accessed: Jun. 2018. [Online]. Available: http://tv.cctv.com/2016/04/26/VIDE6W0VUOnzLvWsYeLrMDYC160426.shtml

[8] J. Zhang, G. Qu, Y. Lyu, and Q. Zhou, "A survey on silicon PUFs and recent advances in ring oscillator PUFs," *J. Comput. Sci. Technol.*, vol. 29, no. 4, pp. 664–678, Jul. 2014.

[9] C. Herder, M.-D. Yu, F. Koushanfar, and S. Devadas, "Physical unclonable functions and applications: A tutorial," *Proc. IEEE*, vol. 102, no. 8, pp. 1126–1141, Aug. 2014.

[10] E. INC. *RSA SecurID*. Accessed: Jun. 2018. [Online]. Available: https://www.rsa.com/en-us/products/rsa-securid-suite/rsa-securid-access/securid-hardware-tokens

[11] G. INC. *Google 2-Step Verification*. Accessed: Jun. 2018. [Online]. Available: https://www.google.com/landing/2step/

[12] N. Karapanos, C. Marforio, C. Soriente, and S. Čapkun, "Sound-proof: Usable two-factor authentication based on ambient sound," in *Proc. 24th USENIX Conf. Secur. Symp. (SEC)*, 2015, pp. 483–498.

[13] W. C. INC. *AIRCable*. Accessed: Jun. 2018. [Online]. Available: https://www.aircable.net/extend.php

[14] G. INC. *SlickLogin*. Accessed: Jun. 2018. [Online]. Available: http://www.slicklogin.com

[15] D. A. Russell, J. P. Titlow, and Y.-J. Bemmen, "Acoustic monopoles, dipoles, and quadrupoles: An experiment revisited," *Amer. J. Phys.*, vol. 67, no. 8, pp. 660–664, Aug. 1999.

[16] A. Rodríguez Valiente, A. Trinidad, J. R. García Berrocal, C. Górriz, and R. Ramírez Camacho, "Extended high-frequency (9–20 kHz) audiometry reference thresholds in 645 healthy subjects," *Int. J. Audiol.*, vol. 53, no. 8, pp. 531–545, Aug. 2014.

[17] MOZZILLA. *Location-Aware Browsing*. Accessed: Jun. 2018. [Online]. Available: https://www.mozilla.org/en-US/firefox/geolocation

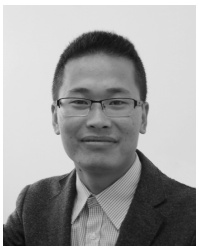
[18] Verayo Technology. (2018). [Online]. Available: http://verayo.com/tech.php

[19] J. Zhang, Y. Lin, Y. Lyu, and G. Qu, "A PUF-FSM binding scheme for FPGA IP protection and pay-per-device licensing," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 6, pp. 1137–1150, Jun. 2015.

[20] J. Zhang, "A practical logic obfuscation technique for hardware security," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 24, no. 3, pp. 1193–1197, Mar. 2016.

[21] J. Zhang, Y. Lin, and G. Qu, "Reconfigurable binding against FPGA replay attacks," *ACM Trans. Des. Autom. Electron. Syst.*, vol. 20, no. 2, pp. 1–20, Feb. 2015.

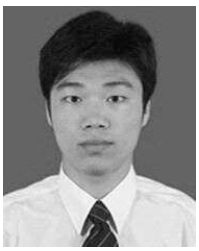
[22] G. E. Suh and S. Devadas, "Physical unclonable functions for device authentication and secret key generation," in *Proc. 44th ACM/IEEE Design Automat. Conf.*, Jun. 2007, pp. 9–14.
[23] R. Pappu, B. Recht, J. Taylor, and N. Gershenfeld, "Physical one-way functions," *Science*, vol. 297, no. 5589, pp. 2026–2030, Sep. 2002.
[24] J. W. Lee, D. Lim, B. Gassend, G. E. Suh, M. van Dijk, and S. Devadas, "A technique to build a secret key in integrated circuits for identification and authentication applications," in *Symp. VLSI Circuits. Dig. Tech. Papers*, Jun. 2004, pp. 176–179.
[25] Z. Pang, J. Zhang, Q. Zhou, S. Gong, X. Qian, and B. Tang, "Crossover Ring Oscillator PUF," in *Proc. 18th Int. Symp. Qual. Electron. Design (ISQED)*, 2017, pp. 237–243.
[26] J. Guajardo, S. S. Kumar, G. Schrijen, and P. Tuyls, "FPGA intrinsic PUFs and their use for IP protection," in *Proc. Cryptograph. Hardw. Embedded Syst.*, 2007, pp. 63–80.
[27] M. Majzoobi, F. Koushanfar, and M. Potkonjak, "Lightweight secure PUFs," in *Proc. IEEE/ACM Int. Conf. Comput.-Aided Design*, Nov. 2008, pp. 670–673.
[28] C. M. Bishop, "Pattern recognition and machine learning," *J. Electron. Imag.*, vol. 16, no. 4, p. 49901, Jan. 2007.
[29] M. Riedmiller and H. Braun, "A direct adaptive method for faster backpropagation learning: The RPROP algorithm," in *Proc. IEEE Int. Conf. Neural Netw.*, Mar. 1993, pp. 586–591.
[30] D. Lim, J. W. Lee, B. Gassend, G. E. Suh, M. van Dijk, and S. Devadas, "Extracting secret keys from integrated circuits," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 13, no. 10, pp. 1200–1205, Oct. 2005.
[31] J. Ye, Y. Hu, and X. Li, "RPUF: Physical unclonable function with randomized challenge to resist modeling attack," in *Proc. IEEE Asian Hardw.-Oriented Secur. Trust (AsianHOST)*, Dec. 2016, pp. 1–6.
[32] J. Zhang, B. Qi, and G. Qu. (2018). "HCIC: Hardware-assisted control-flow integrity checking." [Online]. Available: https://arxiv.org/abs/1801.07397
[33] P. Qiu *et al.*, "Physical unclonable functions-based linear encryption against code reuse attacks," in *Proc. 53rd Annu. Design Automat. Conf.*, Jun. 2016, pp. 1–6.



JILIANG ZHANG received the Ph.D. degree in computer science and technology from Hunan University, Changsha, China, in 2015. From 2013 to 2014, he was a Research Scholar with the Maryland Embedded Systems and Hardware Security Lab, University of Maryland, College Park. From 2015 to 2017, he was an Associate Professor with Northeastern University, China. Since 2017, he has been with Hunan University. He has authored over 30 papers in refereed international conferences and journals such as the IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, the IEEE TRANSACTIONS ON VERY LARGE SCALE INTEGRATION, the *ACM Transactions on Design Automation of Electronic Systems*, and the ACM/IEEE Design Automation Conference. His current research interests include hardware/hardware-assisted security, artificial intelligence security, field programmable gate array, and emerging technologies.

XIAO TAN received the M.S. degree from Hunan University, Changsha, China, in 2012, where he is currently pursuing the Ph.D. degree. His current research interests include information security and Internet of things.

XIANGQI WANG received the Ph.D. degree in computational mathematics from Hunan Normal University, Changsha, China, in 2016. Since 2017, she has been an Assistant Professor with Hunan First normal University, China. She is currently a Visiting Scholar with Peking University. Her current research interests include numerical solution of partial differential equations, large scale parallel computing, and artificial intelligence security.



AIBIN YAN received the Ph.D. degree from the Hefei University of Technology, Hefei, China, in 2015. He is currently an Assistant Professor with the School of Computer Science and Technology, Anhui University, Hefei, China.

ZHENG QIN received the Ph.D. degree in computer science from Chongqing University, Chongqing, China, in 2001. From 2010 to 2011, he served as a Visiting Scholar with the Department of Computer Science, Michigan State University. He is currently a Professor with the College of Computer Science and Electronic Engineering, Hunan University, where he serves as the Vice Dean. He also serves as the Director of the Hunan Key Laboratory of Big Data Research and Application and the Vice Director of Hunan Engineering Laboratory of Authentication and Data Security. He has authored over 80 papers in well-known journals and international conferences. His current research interests include network and data security, data analytics and applications, machine learning, and applied cryptography.