一种基于麦克风和扬声器的轻量级认证协议

曹明生[1]，王惟一[2]，王 韬[2]，徐海津[2]，陈大江[1]，秦志光[1*]

(1. 电子科技大学信息与软件工程学院 网络与数据安全四川省重点实验室 成都 611731；
2. 电子科技大学计算机科学与工程学院 成都 611731)

【摘要】随着网络与通信技术的快速发展，移动终端的身份认证已经成为信息安全中不可或缺的一部分。该系统利用移动设备自带的音频收发硬件(扬声器/麦克风)获取音频物理指纹，实现基于音频物理指纹的设备认证。该系统提出的基于音频物理指纹的认证方法，具有良好的普适性、可靠性和稳定性，同时对硬件要求低，可广泛应用于各种设备认证如无线接入、近场通信等场景。针对上述协议，设计了可用于Android系统的应用软件。实验分析发现该协议具有较好的安全性、普适性和鲁棒性，认证的准确率达到99%以上。

关 键 词 认证协议； 设备到设备； 物理层安全； 无线通信
中图分类号 TP309 **文献标志码** A doi:10.3969/j.issn.1001-0548.2019.04.015

A Lightweight Authentication Protocol with Microphone and Loudspeaker

CAO Ming-sheng[1], WANG Wei-yi[2], WANG Tao[2], XU Hai-jin[2], CHEN Da-jiang[1], and QIN Zhi-guang[1*]

(1. Network and Data Security Key Laboratory of Sichuan Province, School of Information and Software Engineering, University of Electronic Science and Technology of China Chengdu 611731;
2. School of Computer Science and Engineering, University of Electronic Science and Technology of China Chengdu 611731)

Abstract With the rapid development of network and communication technologies, wireless devices authentication has become an indispensable part of information security. In this paper, a novel and lightweight device to device authentication approach is proposed by using the audio transceiver hardware (speaker/microphone) equipped on the mobile devices. Our method has strong universality, reliability and stability, and has low hardware requirements, which can be widely applied to device access scenarios such as wireless access control and near field communication. Additionally, an Android application is developed based on these work. The experimental analysis shows that the protocol has good security and robustness, and the authentication accuracy is over 99%.

Key words authentication protocol; device to device; physical layer secure; wireless communication

物联网的广泛应用，带来了更多的智能化和便捷化[1-3]。同时物联网设备也成为黑客的主要攻击目标。以可穿戴计算为例，手环或者其他人体嵌入式设备实时采集人体的生理特征数据。这些信息如果被恶意用户窃听或者篡改，将会给用户带来很大损失。因此做好设备间的认证是物联网环境的关键问题之一[4-11]。

物联网在电池能力、计算能力、存储能力等方面受限，传统的密码学机制如三方认证、密钥存储、分发等无法适用。此外网络传输如WI-FI、蓝牙、ZigBee等存在缺陷，如需要广播通信，基础协议栈存在漏洞等，很容易遭受攻击。

近年来，物理层认证[11-13]被广泛研究，如基于无线信道特征(received signal strength, RSS)[14]、CSI(channel state information)[15]的测量生成共享密钥等。与RSS、CSI相比，声波[16]、WI-FI[17]、射频[18]、磁场[19]、加速度传感器[20]等信号更易采集。其中基于音频指纹的认证利用无线设备麦克风和扬声器对频率响应的物理不可克隆性，通过提取音频射频麦克风和扬声器对不同频率的频率响应作为物理指纹，实现基于音频物理指纹的设备认证。基于文献[16]中的工作，本文提出了一种新的指纹匹配算法。该算法充分考虑了音频信号回声对指纹提取的影响，极大地提高了认证协议的安全性和认证性。

收稿日期：2018－08－21；修回日期：2018－11－20
作者简介：曹明生(1986－)，男，博士生，主要从事信息安全、智慧城市等方面的研究.
通信作者：秦志光，教授，E-mail: qinzg@uestc.edu.cn

1 基于音频物理指纹认证协议概述

由于音频设备(如手机、笔记本电脑等)的麦克风和扬声器是专为人们通话设计的，因此，不可避免会出现音频硬件对中低频频率的损耗较小，而对较高的频段损耗较大。同时，由于硬件在制造工艺上的差异，以及不可控的人为因素，每个音频设备的麦克风和扬声器都具有自身的缺陷和特点。换句话讲，世界上没有完全一样的麦克风和扬声器。除此之外，把这些硬件装备到不同的无线设备上，会进一步加大这种差异性。因此，可利用这种差异性来识别不同的音频设备。

1.1 系统模型

假设有两个无线设备：认证发起方Alice(简记为A)和认证方Bob(简记为B)。如图1所示，Alice和Bob通过以下交互实现Bob对Alice的认证。

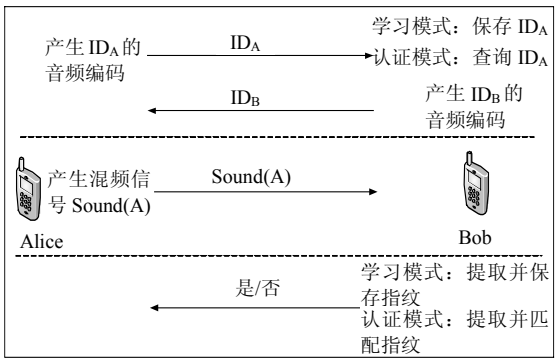


图1 设备认证流程图

1) 认证声源产生阶段：A与B首先通过音频握手建立连接，A利用指纹生成算法产生一段音频信号，并通过扬声器向B发送；

2) 音频指纹提取阶段：当B用麦克风收到A发音频信号以后，利用恰当的指纹提取算法获取A的音频物理指纹；

3) 指纹匹配阶段：在指纹学习模式下，B将获取的物理指纹与A的ID关联并存储；在设备认证模式下，B将收集到的指纹与本地存储的A的指纹作匹配，若匹配成功则通过认证，否则，认证失败。

1.2 认证声源产生阶段

A与B首先实现音频握手。本文采用1 575，1 764，2 004，2 321，2 756，3 392，4 410，6 300，1 025，14 700 Hz的音频信号分别表示0,1,2,…,9。认证的发起者A通过上述音频编码方式将自己的ID(注：在实际应用中，这里的ID是设备的惟一标识，比如手机的序列号)编码成音频信号发送给B；认证者B收到A的ID音频编码后，解码该信号。在学习模式下，B将A的ID存到本地认证设备列表；在认证模式下，B对照本地认证设备列表查找A的ID，如果A的ID不在本地列表，认证失败。认证者B将自己的ID音频编码发送给A，若A在规定时延内接受到B的ID则交互成功，否则，重复上述过程。

为了排除环境噪声的干扰，提高认证效率，这里采用4 000～20 000 Hz频段(以100 Hz为步长)的160个频率混合在一起，即：

$$\mathrm{Sound(A)}=\sum_{i=1}^{n+1}\frac{1}{n+1}\sin(2\pi(4\ 000+100(i-1)T)) \quad (1)$$

式中，n=160，Sound(A) 为混频信号。在实际系统中的发送时长为T=2 s，并且规定在学习模式和认证模式下发送的音量相同。

1.3 音频指纹提取阶段

在利用麦克风记录下A发送的混频信号(该信号记为 $\mathrm{Sound(A}\rightarrow\mathrm{B)}$)后，认证方B执行下列步骤。B利用FFT(fast Fourier transformation)将时域的音频信号转换成频率上的音频信号，并对振幅做20log(·)的数值处理，作为本次获取的音频指纹，即：

$$O_{\mathrm{A}}=20\log(\mathrm{FFT(Sound(A}\rightarrow\mathrm{B)))}=(\xi_1,\xi_2,\cdots,\xi_n) \quad (2)$$

1.4 指纹匹配阶段

在学习模式下，B将该指纹与A的ID相关联，并将 $(\mathrm{ID_A},O_{\mathrm{A}})$ 存入指纹库。在认证模式下，B从指纹库中调出与A的ID相关联的指纹样本 O_{A}，该指纹样本记为：

$$O'_{\mathrm{A}}=(\xi'_1,\xi'_2,\cdots,\xi'_n) \quad (3)$$

B调用匹配算法 $\mathrm{MA}(O_{\mathrm{A}},O'_{\mathrm{A}})$:若 $\mathrm{MA}(O_{\mathrm{A}},O'_{\mathrm{A}})=1$，则匹配成功；否则，认证失败。

文献[16]提出的指纹匹配算法：设定两个阈值 Γ 和 Δ，初始化S=0; T=0，对于每一个 $i\in\{1,2,\cdots,n\}$，做循环：若 $|\xi_i-\xi'_i|<\Gamma$，则S=S+1，i=i+1；否则，T=T+1，i=i+1；循环结束。最后，计算偏离率：$\mathrm{DR}(O_{\mathrm{A}},O'_{\mathrm{A}})$=$T$/$S$。若 $\mathrm{DR}(O_{\mathrm{A}},O'_{\mathrm{A}})<\Delta$，则输出1；否则，输出0。

2 管状指纹匹配算法

2.1 算法设计

本文设计了一个新的音频指纹匹配算法，该算法被称为管状指纹匹配算法。

输入：$O'_{\mathrm{A}}=(\xi'_1,\xi'_2,\cdots,\xi'_n)$ 为B的指纹库中存储的A的指纹；$O_{\mathrm{A}}=(\xi_1,\xi_2,\cdots,\xi_n)$ 为B在认证模式下获取的待认证指纹；η 为判定阈值；ε 为容错阈值。

对每一个 $i\in\{1,2,\cdots,n\}$，B作如下计算：

如果 $\xi_i \in (\xi_i' - \varepsilon, \xi_i' + \varepsilon)$，那么记 $\Delta O_i = 0$；否则，记 $\Delta O_i = |\xi_n - \xi_n'|$。判定值累加得到累计误差：

$$\Delta O_{\mathrm{A}} = \sum_{i=1}^{n} \Delta O_i \tag{4}$$

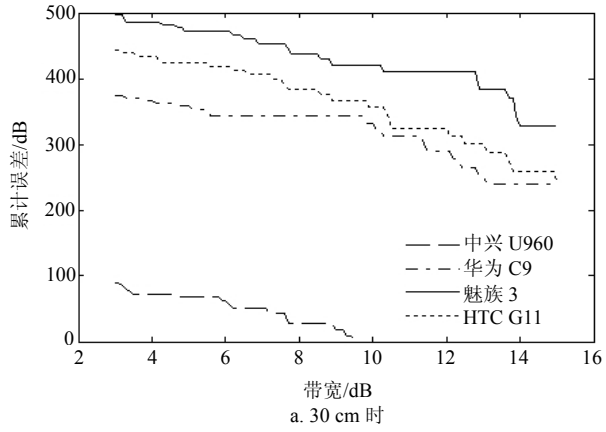
最后将累加和与设定的阈值 η 比较：若 $\Delta O_{\mathrm{A}} < \eta$ 则认为认证通过，返回 $\mathrm{MA}(O_{\mathrm{A}}, O_{\mathrm{A}}') = 1$；否则，返回 $\mathrm{MA}(O_{\mathrm{A}}, O_{\mathrm{A}}') = 0$。

2.2　参数设定

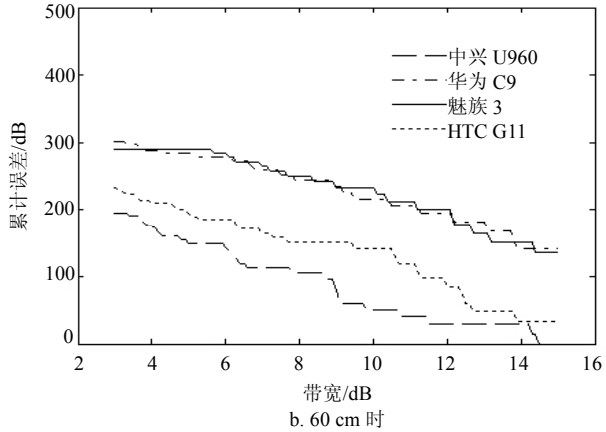
1) 容错误差

为了确定管状匹配算法MA中容错误差 ε 的值，本文做了如下实验：首先，选取30、60、90和600 cm 4个不同的距离进行实验。整个实验中，认证的发起方为中兴U960手机，认证接受方均为HTC T328d手机，同时存在3个敌手：华为C9手机、魅族3手机和HTC G11手机。然后，将4台手机分别按照上述定距的认证协议进行音频握手、混频信号的生成与发送、以及指纹数据的提取。

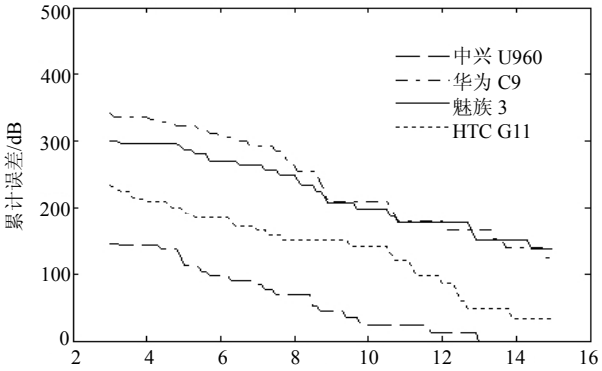
对提取到的指纹数据进行管状匹配分析，为其设置不同的容错阈值，如图2所示。图中的a、b、c、d分别描述的是当认证距离为30、60、90、600 cm时，容错阈值 ε 和累计误差 ΔO_{A} 之间的关系。



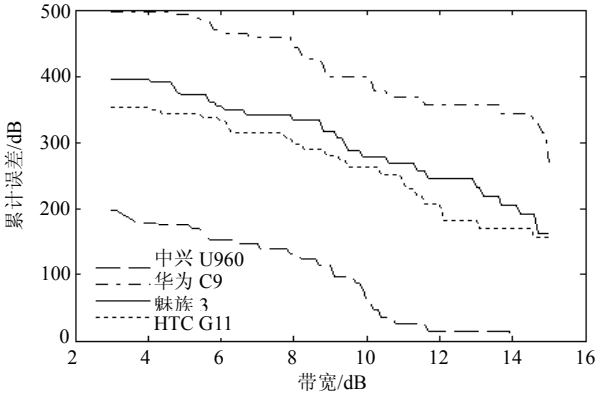
a. 30 cm 时



b. 60 cm 时



c. 90 cm 时



d. 600 cm 时

图2　不同距离时带宽与累计误差分析图

从图2a中可以看出，在相距30 cm认证时，因为中兴U960手机是认证的发起者，当其去认证时，HTC T328d手机从指纹库中提取出对应的样本，通过管状匹配算法得到的累计误差明显要小于其他手机。而其他手机(即敌手)去认证时，因为冒充的是中兴U960，所以认证的接收者提取的仍然是中兴U960的指纹样本，此时再进行管状匹配分析所得到的累计误差要大得多。并且，从图中可以看出，华为C9手机的累计误差相对接近中兴U960，但是当容错阈值超过10之后，中兴U960手机的累计误差已经为0，而华为C9手机的累计误差仍然较大。图2b所描述的是相距60 cm时的认证情况。此时中兴U960手机的累计误差仍然最小，与它最接近的敌手是HTC G11手机。尽管两者在一部分容错阈值下的差别不是很大，但是当容错阈值为10.5左右时，能够很好地区分中兴U960手机和敌手手机。图2c为间距为90 cm的认证情况。本次认证情况和图2b描述的认证情况非常相似，并且当容错阈值为10.8时，中兴U960手机与敌手手机的差别达到最大。图2d显示的是相距600 cm时的认证情况，在这次实验中，中兴U960手机和敌手手机的累计误差差距较大。尽管这次实验中各个手机的表现和之前有所不同，但是当容错阈值为

10.8左右时，仍能很好地区分认证的发起者与敌手。

综上所述，不同手机(中兴U960、华为C9、HTC G11、魅族3)和不同距离(30、60、90、600 cm)对容错阈值的选取影响不大，并且在容错阈值为10.8时能够很好地区分认证的发起者和敌手。因此本文将容错阈值设置为10.8。

2) 阈值的设置

根据图2给出的不同手机在不同距离下容错阈值和累计误差的关系图，已经得到最佳容错阈值为10.8。而且从图2中可以看出当容错阈值为10.8时，敌手的累计误差均大于100，而认证的发起者的累计误差均小于100，所以将阈值设置为100能够很好地判断认证是否成功。关于容错阈值和阈值的确定将在后面的实际测试中，给出实际测试的结果来验证容错阈值和阈值设置的合理性。

3) 距离对认证稳定性的影响

本文给出该认证方法的最大适用距离。按照上文给出的定距认证协议，通过实验的方式采集到认证成功率与距离的关系如表1所示。

表1 距离与认证成功率

设备间距离/m	认证成功率/%
5	99
10	99
15	96
20	90

从表1中可以看出当设备间距离变大时，认证效果有所下降，当相距20 m时的成功率已经下降到90%，即认为此时该方法已经不能稳定使用，所以20 m是该系统能够稳定工作的最大距离。但是当距离进一步变大时，在实验过程中发现本文所提出的设备认证方法仍能适用。经分析得知，在设备音量不变的前提下，只要待认证设备发出的声音不被周围噪音干扰，能够被顺利接收，那么认证就能顺利进行。所以实际有效认证距离可以更远，同时20 m的认证范围也已经适用于绝大多数场合。

这里需要注意的是，如果将认证协议更改成调用3次原始协议，当有一次认证成功就算A通过认证。在此情形下，当设备之间距离为20 m时，认证成功率可提高到99%。

3 实验结果

本节将给出认证协议具体的测试情况。如文献[16]所示，由于同一音频设备的频率响应在不同环境下保持相对稳定，因此实验环境可以任意选取，本文选择在室内进行实验。测试设备为13部手机，分别是中兴U960、华为C9、HTC G11、三星S5560、索尼Z1、三星Note2、华为荣耀3C、HTC ONE、酷派、两部同型号的魅族3和两部同型号的小米2。考虑到某些手机间的频率响应相似，使得敌手在使用这些手机时攻击成功的概率更高，因此本文又做了大量实验，对13部手机，两两组合，一部作为认证发起者一部作为敌手进行实验。总共的10 000次实验中，认证发起者共进行了2 000次认证，敌手共进行了8 000次攻击.

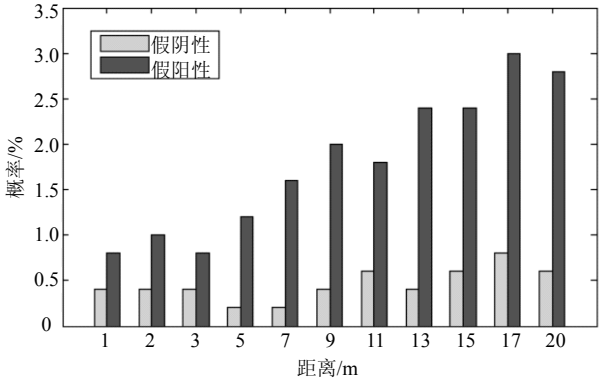


图3 协议的错误率

图3根据实验结果绘制出了在距离变化下该认证协议假阴性和假阳性错误的比率。实验结果表明，在定距认证下，随着认证距离的增大，敌手攻击成功的概率(即假阳性错误率)会有所升高，但在0～10 m的认证距离范围内敌手攻击成功的概率不会超过2%；同时，在0～20 m的认证距离内，对合法用户认证失败的概率(即假阴性错误率)都低于1%。

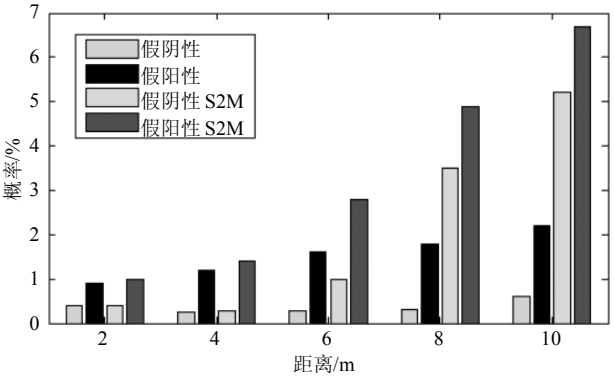


图4 新协议与S2M协议的性能比较

为了进一步说明该协议的稳定性和安全性，在2～10 m的认证距离下，对比了本文协议和S2M协议的性能。如图4所示，在2 m和4 m的认证距离下，两个协议的错误率相差无几；在6、8和10 m的认证距离下，本文协议的效率明显高于S2M协议。因此，

本文设计的认证方法的安全性和鲁棒性得到了很好的验证，同时也证明了该认证方法适用于大多数移动设备，具有良好的推广性。

4 结 束 语

本文通过利用设备自带的音频硬件，提出一种轻量级的基于音频硬件物理指纹的设备认证协议。该协议基于一个新的音频指纹匹配算法，即管状音频指纹匹配算法。针对上述协议设计了可用于Android系统的应用软件，并在真实场景进行了大量实验。实验结果表明，该协议的应用范围可以达到10 m以上，并且认证的平均准确率达到99%以上。与传统密码学方法相比，基于音频物理指纹的设备认证协议具有较好的安全性、普适性和鲁棒性。

参 考 文 献

[1] ZHANG Kuan, LIANG Xiao-hui, LU Rong-xing, et al. Sybil attacks and their defenses in the Internet of things[J]. IEEE Internet of Things Journal, 2014, 1(5): 372-383.

[2] SICARI S, RIZZARDI A, GRIECO L, et al. Security, privacy and trust in Internet of things: The road ahead[J]. Computer Networks, 2015, 76: 146-164.

[3] ZHANG Kuan, YANG Kan, LIANG Xiao-hui, et al. Security and privacy for mobile healthcare networks: From a quality of protection perspective[J]. IEEE Wireless Communication, 2015, 22(4): 104-112.

[4] CHEN Da-jiang, MAO Xu-fei, QIN Zhen, et al. Wireless device authentication using acoustic hardware fingerprints [C]//BigCom2015. Taiyuan, China: Springer, 2015: 193-204.

[5] XI Wei, HE Yuan, LIU Yun-hao, et al. Locating sensors in the wild: Pursuit of ranging quality[C]//ACM SenSys'10. Zürich, Switzerland: ACM, 2010: 295-308.

[6] CHEN Da-jiang, QIN Zhen, MAO Xu-fei, et al. SmokeGrenade: An efficient key generation protocol with artificial interference[J]. IEEE Transactions on Information Forensics and Security, 2013, 8(11): 1731-1745.

[7] WANG Wei, CHEN Ying-jie, Zhang Qian. Privacy-preserving location authentication in Wi-Fi networks using fine-grained physical layer signatures[J]. IEEE Transactions on Wireless Communications, 2016, 15(2): 1218-1225.

[8] DANEY B, LUECKEN H, CAPKUN S, et al. Attacks on physical-layer identification[C]//Proceeding of the Third ACM Conference on Wireless Network Secwrity. Aoboken, New Jersey, USA: ACM, 2010: 89-98.

[9] ZHOU Zhe, DIAO Wen-rui, LIU Xiang-yu, et al. Acoustic fingerprinting revisited: Generate stable device id stealthily with inaudible sound[C]//ACM CCS'14. Scottsdale, AZ, USA: ACM, 2014: 429-440.

[10] ZHU Tong, MA Qiang, ZHANG Shan-feng, et al. Context-free attacks using keyboard acoustic emanations[C]//ACM CCS'14. Scottsdale, AZ, USA: ACM, 2014: 453-464.

[11] CHEN Da-jiang, ZHANG Ning, CHENG Nan, et al. Physical layer based message authentication with secure channel codes[J]. IEEE Transactions on Dependable and Secure Computing, 2018, DOI: 10.1109/TDSC. 2018. 2846258.

[12] CHEN Da-jiang, ZHANG Ning, LU Rong-xin, et al. An LDPC code based physical layer message authentication scheme with prefect security[J]. IEEE Journal on Selected Areas in Communications, 2018, 36(4): 748-761.

[13] CHEN Da-jiang, ZHANG Ning, LU Rong-xin, et al. Channel precoding based message authentication in wireless networks: Challenges and solutions[J]. IEEE Network, 2018, 33(1): 99-105.

[14] AZIMI-SADJADI B, KIAYIAS A, MERCADO A, et al. Robust key generation from signal envelopes in wireless networks[C]//CCS'07. Alexandia, Virginia, USA: ACM, 2007: 401-410.

[15] SAYEED A, PERRIG A. Secure wireless communications: Secret keys through multipath[C]//IEEE International Conference on Acoustics. Las Vegas, USA: Speech and Signal Processing, 2008: 3013-3016.

[16] CHEN Da-jiang, ZHANG Ning, QIN Zhen, et al. S2M: A lightweight acoustic fingerprints based wireless device authentication protocol[J]. IEEE Internet of Things Journal, 2017, 4(1): 88-100.

[17] LIU Yao, NING Peng. Enhanced wireless channel authentication using time synched link signature[C]// INFOCOM'12. Orlando, USA: IEEE, 2012: 2636-2640.

[18] URENTEN O, SERINKEN N. Wireless security through RF fingerprinting[J]. Canadian Journal of Electrical and Computer Engineering, 2007, 32(1): 27-33.

[19] BICHLER D, STROMBERG G, HUEMER M, et al. Key generation based on acceleration data of shaking processes[C]//UbiComp. Berlin, Germany: Springer, 2007: 304-317.

[20] QIU Fu-dong, HE Zheng-xian, KONG Ling-hem, et al. MAGIK: An efficient key extraction mechanism based on dynamic geomagnetic field[C]//INFOCOM 2017. Atlanta, GA, USA: IEEE, 2017: 1-9.

编　辑　刘飞阳