IET Communications

Research Article

Physical-layer security in Internet of Things based on compressed sensing and frequency selection

ISSN 1751-8628
Received on 14th September 2016
Revised 13th November 2016
Accepted on 8th December 2016
E-First on 7th June 2017
doi: 10.1049/iet-com.2016.1088
www.ietdl.org

Ning Wang[1] ✉, *Ting Jiang*[1], *Weiwei Li*[2], *Shichao Lv*[3]

[1]*Key Laboratory of Universal Wireless Communication Ministry of Education, Beijing University of Posts and Telecommunications, Beijing, People's Republic of China*

[2]*School of Information and Engineering, Hebei University of Engineering, Handan, People's Republic of China*

[3]*Beijing Key Laboratory of IOT Information Security Technology, IIE CAS, Beijing, People's Republic of China*

✉ *E-mail: wangning8566@bupt.edu.cn*

Abstract: Information security is a vital concern in Internet of Things (IoT). Traditional security method based on public or private key encryption scheme is limited by the trade-off between low cost and high level of security. Among different security solutions, utilising compressed sensing (CS) in combination with the physical-layer security to achieve the security is a remarkable method. However, in the current literatures, little attention has been given to the area of static environment, which will lead the risk of information leakage in the CS security model. In this study, the authors propose a new CS security model, in which circulant matrix is exploited to improve the generation efficiency of the measurement matrix, and binary resilient functions are utilised to enhance the security. Furthermore, considering the practical application, they present a feasible framework, named CS security scheme based on frequency-selective, where the frequency-selective feature of the wireless channel is applied to support the static environment. To verify the effectiveness of the proposed scheme, they conducted experiments and numerical simulations to evaluate the performance, and the results are satisfactory.

1 Introduction

The Internet of Things (IoT) is a new generation of information technology, which is an important stage in the development of the information age [1]. Since a wide range of application scenarios, the security of IoT is a critical issue, undoubtedly. However, due to broadcast transmission and stringent resource constraints, data transmission in the perceptual layer of IoT is prone to interception and eavesdropping. In these cases, how to achieve a low cost and complexity but high security policy is still a challenge.

Wireless security in IoT is traditionally achieved on the higher layers of protocol stack through cryptography approaches, however, it is hard to overcome the conflict between the aims of low computational cost and a high level of security, such as lightweight cryptography. To address the weaknesses of traditional security policy, physical-layer security (PLS) technology is emerging to a promising new technique to solve the security problem in wireless communication. However, based on the current state of the development of the PLS technology, there are still some problems in practical application. Such as key-extraction based on physical layer, the generating rate of the extracted key is still not enough to satisfy the requirement for transfer speed. In the case of keyless security in PLS, artificial noise is one of the primary means, which is undesired in IoT because it will occupy the additional communication resources. Therefore, more studies need to be conducted to protect the security of IoT.

Recently, compressed sensing (CS) [2] has been considered in WSNs and IoT [3]. CS changes the rules of data acquisition in the information system, where the data or signals can be efficiently sampled and accurately reconstructed with much fewer samples than Nyquist theory [4]. It can prolong the lifetime of the perceptual node network and increase the transfer efficiency of valid data. Furthermore, the ability of CS to ensure security received some attention in the past, such as in [5, 6]. In general, using the measurement matrix as the symmetric key for encryption and decryption is a most direct and effective method to achieve security. In [7], the author demonstrated that CS-based encryption does not achieve Shannon's definition of perfect secrecy, but can provide a computational guarantee of secrecy. While, Mayiami *et al.* [8] proposed that under some conditions, based on CS the perfect secrecy can be achievable, and these conditions are not difficult to satisfy in the practical application of CS. Therefore, address to the practical application of IoT, the key problem is how to efficiently construct the measurement matrix and guarantee the security of this matrix.

Surprisingly, in combination with the PLS technology, the CS security model can achieve a better effect. Dautov and Tsouri [9, 10] presented that based on one of the PLS technology, i.e. key-extraction based on physical layer, a CS security framework was established, which can realise securing while sampling in wireless body area networks (WBANs). However, there still exist some problems in the practical application, which has been given limited attention in previous works. In IoT, static environment is the main application scenario, while this situation may lead to some potential harms. First, static environments will lead to a stationary signal over a period of time, which may produce a non-random measurement matrix which may dissatisfy the restricted isometry property (RIP) condition. It will result in the failure of compression and reconstruction in CS. Second, lower randomness under the static environments may cause the extracted secret information to become too simple and easy to be imitated. Moreover, Herman and Strohmer [11] showed that an adversary who has no need to obtain a precise legitimate measurement matrix still be able to reconstruct the compressed signal to some extent based on a similar matrix. Therefore, further investigation is needed in the area of the CS security scheme based on PLS.

In this paper, we propose a new CS security model based on resilient functions and circulant matrix, and give a feasible framework based on frequency-selective for IoT, called CS security scheme based on frequency-selective (CSSS-FS). More specifically, we exploit the circulant matrix to improve the generating efficiency of the measurement matrix, and utilise a binary resilient function to guarantee the security. Considering the practical application, we use the frequency-selective feature of the wireless channel as the effective means to overcome the static application scenario. Moreover, by means of experiments and numerical simulation based on real-world database, i.e. MIT-BIH

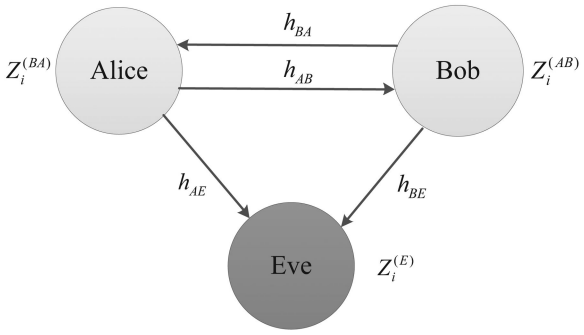

Fig. 1 *General eavesdropping communication system*

arrhythmia database, the performance of the proposed scheme was evaluated.

The remaining of this paper is organised as follows: Section 2 contains an introduction to CS and security. We introduce the proposed CS security model in Section 3, and describe the proposed framework, i.e. CSSS-FS, in Section 4. Section 5 is devoted to an account of experiments and simulations to evaluate our method, followed by the discussion and conclusion of this study in Section 6.

2 Preliminaries and problem statement

2.1 Compressed sensing

CS is one of the important progress in the field of signal processing, and it is expected to break through the Nyquist–Shannon sampling theorem. In general, CS contains two stages, i.e. sensing and reconstruction stages, and the sensing stage includes the process of sparse transformation and compressed measurement. In this study, we focus on compressed measurement in CS. Herein, we simply introduce this technology.

Suppose $\boldsymbol{\omega} \in R^{N\times 1}$ is an unknown source vector with only a few non-zero entries; then, $\boldsymbol{\omega}$ through noisy measurements is given by

$$\boldsymbol{t} = \boldsymbol{\Phi}\boldsymbol{\omega} + \boldsymbol{v} \tag{1}$$

where $\boldsymbol{t} \in R^{M\times 1}$ is an available measurement vector, $\boldsymbol{\Phi} \in R^{M\times N}$ ($M \ll N$) is a known measurement matrix, and any M columns of $\boldsymbol{\Phi}$ are linearly independent (i.e., satisfies the RIP condition [2]), $\boldsymbol{v} \in R^{M\times 1}$ is an unknown noise vector. Furthermore, if signal $\boldsymbol{\omega}$ is not sparse itself, it may be represented as a sparse signal in some orthonormal basis $\boldsymbol{\Psi}$, i.e. $\boldsymbol{\omega} = \boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{x}$ is a sparse signal.

While the sampling process is simply a random linear projection, the reconstruction process to recover the original signal $\boldsymbol{x}$ from the received measurements $\boldsymbol{y}$ is highly non-linear. More precisely, the CS theory suggests the following reconstruction algorithm based on the l_1-minimisation of the transform coefficient vector $\boldsymbol{u}$:

$$\min \| \boldsymbol{u}' \|_1 \quad \text{s.t.} \quad \boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{u}' \tag{2}$$

This is a convex optimisation problem that conveniently reduces to a linear program known as basis pursuit or other related reconstruction algorithms.

CS can significantly reduce the number of sampling points and improve the system efficiency. In addition, many signal processing methods can be utilised to achieve the sparse transformation, such as Fourier transform and wavelet transform, hence, in practice, many popular signals can be easily transformed to sparse domain. Therefore, in wireless sensor networks (WSNs) and the IoT, there is ample research for leveraging this technology to improve the performance of the system. Next, we will discuss the security in CS.

2.2 Achieve security by CS

In general, CS does not provide information theoretic security (or called perfect secrecy) but can be viewed as a computationally secure cryptosystem [7]. Thus, as the computational power increased, there is the risk for information leaking in this model. Fortunately, under some conditions, based on CS, the perfect secrecy condition introduced by Shannon can be achievable [8]. These conditions contain: the number of measurements is equal or greater than two times of sparsity level of the messages, the measurement matrix satisfies RIP, and the number of source messages goes to infinity [8]. In fact, these conditions are not difficult to satisfy in some practical application of CS, such as our previous work [12]. Thus, considering the application in IoT, the security model based on CS is focused on how to construct an efficient and secure measurement matrix under various scenarios, such as static environment.

3 CS security model based on resilient functions and circulant matrix

3.1 System setup

Consider a wireless communication system consisting of two legitimate nodes (Alice and Bob) and an eavesdropper (Eve), as shown in Fig. 1. Let h_{AB} and h_{BA} denote the legitimate channels, and h_{AE} and h_{BE} denote the wiretap channels. According to the key extraction technology which is one of PLS, Alice and Bob can extract the similar information $Z_i^{(AB)}$ and $Z_i^{(BA)}$ based on the channel features, respectively. Through quantisation and reconciliation, they can acquire a same bit sequence, denoted as $(s_1, s_2, \ldots, s_m)$. Similarly, based on the wiretap channels h_{AE} and h_{BE}, Eve can receive the correlation information, and finally obtains an estimated sequence $(\tilde{s}_1, \tilde{s}_2, \ldots, \tilde{s}_m)$. The system shown in Fig. 1 can apply to the most of wireless communication scenarios. Next, based on this system, we use the resilient functions to establish a security vector.

3.2 Security vector based on resilient functions

Resilient functions were introduced in [13] for key distribution and generation of random strings in the presence of faulty processors. Based on the threshold characteristic, we exploit resilient function to enhance the security of the measurement matrix. Herein, a binary resilient function is used, and the definition is as follows.

Definition 1: [13]: Let $n \geq m \geq 1$ be integers and suppose

$$f: \{0,1\}^n \to \{0,1\}^m \tag{3}$$

where f is a function that accepts n input bits and produces m output bits. Suppose $t \leq n$ be an integer, and t arbitrary input bits out of n are fixed by an adversary, and the remaining $n - t$ input bits are chosen independently at random. Then f is said to be t-resilient by which an output of every possible m-tuple is equally likely to occur.

The process of establishing the security vector based on the binary resilient functions, can be stated as Theorem 1 as follows:

Theorem 1: Let us consider the communication system in Fig. 1. Let f be an (n, m, t) resilient function, and set $f(z_1^{AB}, z_2^{AB}, \ldots, z_n^{AB}) = f(z_1^{BA}, z_2^{BA}, \ldots, z_n^{BA}) = (s_1, s_2, \ldots, s_m)$ and $f(z_1^E, z_2^E, \ldots, z_n^E) = (\tilde{s}_1, \tilde{s}_2, \ldots, \tilde{s}_m)$. Let the eavesdrop probability be $1 - p$, i.e. $P_r(z_i^E \neq z_{i1}^{AB}) = P_r(z_i^E \neq z_i^{BA}) = p$. If

$$p \geq \lim_{n\to\infty} \frac{(n-t)}{2n} \tag{4}$$

We have $\Pr((\tilde{s}_1, \tilde{s}_2, \ldots, \tilde{s}_m) = (s_1, s_2, \ldots, s_m)) \to 1/2^m$, i.e. $(s_1, s_2, \ldots, s_m)$ is independent of $(\tilde{s}_1, \tilde{s}_2, \ldots, \tilde{s}_m)$.

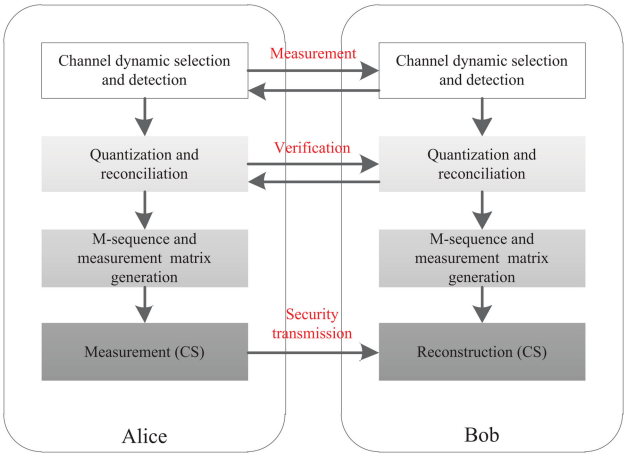

Fig. 2 *Block diagram of the proposed framework*

Proof: Suppose Eve takes exactly arbitrary t bits from the communication of Alice and Bob. Obviously, there are randomly and independently t bits correctly taken in $(z_1^E, z_2^E, \ldots, z_n^E)$. About half of the remaining $n-t$ bits will be correct and the other bits remain to incorrect on average when n approaches to infinity or many frames are repeated. Therefore, the total number of error bits in $(z_1^E, z_2^E, \ldots, z_n^E)$ will be $(n-t)/2$ on average, which leads to $\Pr(z_i^{AB} \neq z_i^E) \to (n-t/2n)$ ($z_i^{AB} \neq z_i^E$ and $z_i^{BA} \neq z_i^E$ are statistically equal). If $\Pr(z_i^{AB} \neq z_i^E) = \Pr(z_i^{BA} \neq z_i^E) \geq \lim_{n\to\infty}(n-t/2n)$, it means no more than t bits in $(z_1^E, z_2^E, \ldots, z_n^E)$ are correctly taken from $(z_1^{AB}, z_2^{AB}, \ldots, z_n^{AB})$ or $(z_1^{BA}, z_2^{BA}, \ldots, z_n^{BA})$. Based on Definition 1, the output $(\tilde{s}_1, \tilde{s}_2, \ldots, \tilde{s}_m)$ recovered from $(z_1^E, z_2^E, \ldots, z_n^E)$ tends to uniformly distributed over m vector space, i.e. $\Pr(f(z_1^E, z_2^E, \ldots, z_n^E) = (\tilde{s}_1, \tilde{s}_2, \ldots, \tilde{s}_m) | s_{i_j} = \tilde{s}_{i_l}) = 2^{-m}$, where $0 \leq j, l \leq t$. Thus, we have $(s_1, s_2, \ldots, s_m)$ is independent of $(\tilde{s}_1, \tilde{s}_2, \ldots, \tilde{s}_m)$ when $p \geq \lim_{n\to\infty}((n-t)/2n)$. □

3.3 Security measurement matrix based on circulant matrix

Up to now, we have obtained a security vector from Theorem 1. Based on this vector, measurement matrix can be established. Furthermore, in order to improve the generation efficiency, we consider circulant matrix as the structure of measurement matrix, where forming a matrix only needs a small amount of vectors by a simple operation. Formally, this process can be stated as Inference 1 as follows:

Inference 1: Let f be an (n, m, t) resilient function, and set $f(z_1^{AB}, z_2^{AB}, \ldots, z_n^{AB}) = f(z_1^{BA}, z_2^{BA}, \ldots, z_n^{BA}) = (s_1, s_2, \ldots, s_m)$ and $f(z_1^E, z_2^E, \ldots, z_n^E) = (\tilde{s}_1, \tilde{s}_2, \ldots, \tilde{s}_m)$. They can form the circulant matrix

$$\boldsymbol{S}_m = \begin{bmatrix} s_1 & s_2 & \cdots & s_m \\ s_m & s_1 & \cdots & s_{m-1} \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}, \quad \tilde{\boldsymbol{S}}_m = \begin{bmatrix} \tilde{s}_1 & \tilde{s}_2 & \cdots & \tilde{s}_m \\ \tilde{s}_m & \tilde{s}_1 & \cdots & \tilde{s}_{m-1} \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \tag{5}$$

These circulant matrices can be used as the measurement matrix in CS, and if $p \geq \lim_{n\to\infty}((n-t)/2n)$, depending on $\tilde{\boldsymbol{S}}_m$, an adversary cannot recover the compressed data when the measurement matrix is $\boldsymbol{S}_m$.

Proof: Firstly, the proof of that the circulant matrix can be used as the measurement matrix in CS, has been given in [14]. According to [14], the circulant matrix does not reduce the quality of the recovered data in CS. Then, in the process of CS, the lower the correlation between the measurement matrices in compression and reconstruction phase, the harder the data to be currently recovered. This property has been stated by Herman and Strohmer [11]. As a result, if the CS is security, i.e. the compressed data cannot be recovered by an adversary, it must satisfy that the legitimate measurement matrix is independent of the illegitimate one. From Theorem 1, we can see that $(s_1, s_2, \ldots, s_m)$ is independent of $(\tilde{s}_1, \tilde{s}_2, \ldots, \tilde{s}_m)$ when $p \geq \lim_{n\to\infty}((n-t)/2n)$. Obviously, constructed by $(s_1, s_2, \ldots, s_m)$ and $(\tilde{s}_1, \tilde{s}_2, \ldots, \tilde{s}_m)$, $\boldsymbol{S}_m$ and $\tilde{\boldsymbol{S}}_m$ are mutual independent under the same condition. This means that the illegitimate measurement matrix which is obtained by adversary is independent of the legitimate measurement matrix in the proposed CS model. Thus, the CS model in Inference 1 is security.

In addition, considering the communication environment in IoT and the characteristic of key-extraction based on physical layer, adversary is hard to obtain more correct bits, which is protected by channel reciprocity and randomness, i.e. the condition $p \geq \lim_{n\to\infty}((n-t)/2n)$ is easy to achieve. □

4.4 Proposed security framework

4.1 Scheme overview

As shown in Fig. 2, our mechanism includes three phases: channel dynamic selection and detection, quantisation and reconciliation, m-sequence and measurement matrix generation.

i. *Channel dynamic selection and detection*. This phase is the initial stage, where the selected channel (or carrier frequency) is determined and the initial RSSI is collected by Alice and Bob.
ii. *Quantisation and reconciliation*. After collecting a set of RSSI, Alice and Bob compute the value of RSSI to obtain a series of quantitative values. Through quantisation and reconciliation, Alice and Bob can acquire a same bits sequence.
iii. *M-sequence and cyclic matrix generation*. Using the sequence obtained from ii Alice and Bob generate m-sequence and construct a security measurement matrix based on resilient function and circulant matrix.

In the following, we describe our technical details.

4.2 Channel dynamic selection and measurement

Inspired by the authors in [15, 16], we exploit the frequency-selection channel to increase the entropy of the measured channel to improve the rate of generated keys from physical layer. Moreover, considering the actual communication system with multiple communication channels (such as Zigbee, which is one of the most used communication technology in IoT and has 16 channels in the 2.4 GHz range), we use dynamic-selection channel method which is based on the frequency-selective feature to overcome the application problem of static environment.

Without loss of generality, we consider a system with n different carrier frequencies, i.e. channels, $\boldsymbol{F} = \{f_1, \ldots, f_n\}$. First, Alice and Bob need a pre-set index $CH_1 \in \{1, \ldots, n\}$ as the initial channel. CH_1 can be encrypted for further enhance the security. Then the following dynamic selection sequence is computed by the formula:

$$\mathrm{CH}_{j+1} = ((\mathrm{CH}_j - E) \bmod F) + G \tag{6}$$

where E, F and G are integer parameters, which can be regulated according to the channel number.

As a result, the sequence $\mathbf{CH} = \{\mathrm{CH}_1, \ldots, \mathrm{CH}_n\}$ can be formed. For example, zigbee communication system has $n = 16$ different carrier frequencies, then $E = 6$, $F = 16$ and $G = 1$. Here, it is noteworthy that Alice and Bob only pre-set very few parameters, i.e. index CH_1, even if augmenting the encryption algorithm, the computational cost is very small.

According to sequence **CH**, Alice and Bob can choose channel f_i, $i = \mathrm{CH}_j$ at the same time to detect the wireless channel at one slot. By using training sequence or preamble, they can obtain a set

of RSSI values $\boldsymbol{M} = (m_1, \ldots, m_j)$, and this process can be denoted as $\mathbf{CH} \rightarrow \boldsymbol{M}$.

As noted, formula (6) ensures that Alice and Bob have the same detection channel, and all of the candidate channel can be used.

4.3 Quantisation and reconciliation scheme

After detecting the channels, the value of the RSSI can be collected, i.e. $M = (m_1, \ldots, m_j)$. Then, through quantising, we can obtain a set of secure and random bit sequences. In this study, we used a quantiser which has multiple quantisation levels and guard bands, developed in [16, 17], as the quantisation scheme. The quantisation intervals and guard band values can be determined by

$$\int_{v_i - 1}^{v_i - g_i} U_h \, \mathrm{d}h = \frac{1-\alpha}{m}; \quad \int_{v_i - g_i}^{v_i} U_h \, \mathrm{d}h = \frac{m}{1-\alpha} \tag{7}$$

where v_i and g_i are thresholds of the quantisation and guard band, α is the ratio of the sampling values that are discarded, m is the number of discrete quantisation values, and h is the sampling values with probability distribution U.

Meanwhile, in order to extract the same bits, the cascade protocol [18] of reconciliation can be used, and the hash function can enhance this process. Finally, Alice and Bob can obtain a series of identical bits $\boldsymbol{w}$ with entries $\boldsymbol{w}_i \in \{0, 1\}$ from the physical layer of wireless communication. This process can be denoted as $\boldsymbol{M} \rightarrow \boldsymbol{w}$.

It is noteworthy that based on the existing physical layer key-extraction technology, the rate of extracted bits is generally lower on the condition of RSSI. However, in this study, we can utilise M-sequence and circulant matrix to achieve the CS security model even if the rate of extracted bits is not extremely high, since the number of the required security bits in the generation of the security measurement matrix is small.

4.4 M-sequence and partial circulant generation

To implement the proposed CS security model, m-sequence and circulant matrix are generated. M-sequence is a pseudo random sequence, which can be generated by linear feedback shift register (LFSR). Given the state of the previous output (called a seed), the LFSR uses this seed as the input again, and shifts it into adjacent positions to produce a single output bit. In this study, we integrate LFSR and the physical layer key-extraction scheme to produce m-sequence.

First, based on Theorem 1, we consider $\boldsymbol{w}$ as the input vector, then $\boldsymbol{w}^{(s)}$ is the output vector, denoted as $\boldsymbol{w} \rightarrow \boldsymbol{w}^{(s)}$. Next, we see $\boldsymbol{w}^{(s)}$ as the initial state of LFSR, then the m-sequence can be generated, denoted as $\boldsymbol{w}^{(s)} \rightarrow \boldsymbol{w}^{(m)}$. Furthermore, to balance the number of zeros and ones, we map the binary sequence $\boldsymbol{w}^{(m)}$ to the Bernoulli sequence $\boldsymbol{B} = (b_0, b_1, \ldots, b_{N-1})$ with entries $b \in \{\pm 1\}$, denoted as $\boldsymbol{w}^{(m)} \rightarrow \boldsymbol{B}$.

Subsequently, sequence $\boldsymbol{B}$ can be easy to construct a circulant matrix, denoted as $\boldsymbol{B} \rightarrow \boldsymbol{S}$. As stated by [14] the partial circulant matrix (i.e., a part of a circulant matrix) whose entries are ± 1 Bernoulli variables can competent as measurement matrix in CS, and according to Inference 1, this measurement matrix can achieve the security of transmission.

Furthermore, in order to enhance security, we use the arithmetic sum of $\boldsymbol{S}$ to enhance the measurement matrix, given by

$$\boldsymbol{\Phi} = \sum_{i=1}^{L} \boldsymbol{S}_i \tag{8}$$

where L is the number of iterations.

Obviously, the superposition of multiple matrices $\boldsymbol{S}$ can increase the difficulty for adversary to obtain the correct measurement matrix. This process can be denoted as $\boldsymbol{S} \rightarrow \boldsymbol{\Phi}$.

As a result, from channel selection, i.e. $\mathbf{CH} \rightarrow \boldsymbol{M}$, to security measurement matrix construction i.e. $\boldsymbol{S} \rightarrow \boldsymbol{\Phi}$, we establish a feasible security framework, i.e. CSSS-FS.

5 Evaluation

5.1 Efficiency analysis by experiments

i. *Experiment setup:* To test our claim, we performed channel measurements in the time domain using Crossbow's TelosB motes (MicaZ) as the working devices. The MPR2400 uses the Chipcon CC2420, IEEE 802.15.4 compliant, ZigBee ready radio frequency transceiver integrated with an Atmege128L micro-controller. Here, we consider IEEE 802.15.4 channels which are divided on 16 channels in the 2.4 GHz range. To present a static environment, experiments were conducted in a small quiet conference room, and the execution time was at the middle of the night. The experiment considers bits extraction in close proximity of a sensor node (Bob), an AP (Alice) and an attacker (Eve). Eve is placed on the conference table and Alice and Eve are placed about 10 inches to the right and to the left of the Bob, respectively. Using the communication between a MicaZ node and a programmer board MIB520, we can observe the value of RSSI. The transmission time per packet is 1.2 ms, and the test lasts for 10,000 message exchanges.

ii. *Data acquisition:* To verify the performance of our scheme to the static environment, we conduct two sets of experiments: single channel samples (fixed one of 16 channels) and dynamic channel samples (hopping among 16 channels) between Alice and Bob. Fig. 3 shows a snapshot of the obtained data under the two operating modes. It is evident that the value of the RSSI is relatively stable under the single channel sampling, while in the case of dynamic channel, the RSSI still have a better randomness even under a static environment.

iii. *Efficiency of generating matrix:* We validate the generation efficiency of the measurement matrix in the proposed security framework relative to the reference framework. The proposed method is given in Section 4, i.e. CSSS-FS, and the reference method is presented in [9], where the reconstruction matrix is formed by the Gaussian random matrix. Due to they have different application scenarios (dynamic and static environment, respectively), for consistency, we suppose that they have obtained a same length bits from wireless channel.

iv. Herein, we conduct two sets of experiments to test both types of matrices corresponding to the proposed and the reference framework, respectively. In [10], we assumed that there is a sparse signal R of length $N = 256$ with $t = 30$ non-zero elements that randomly take on value from $\{\pm 1\}$ at random locations. A compression ratio (CR) of 50% defines the compressed signal of length $M = 128$, and according to formulae (1) and (3), the measurement matrix needs an $M \times N$, i.e.128×256 matrix.

v. To simplify the analysis, the security bits extracted from physical layer are divided into equal parts (known as 'seed') and the length of every seed is $m = 15$. Under the reference framework, LFSR can produce 32,768 (128×256) bit m-sequences when seed length is 15 and the primitive polynomial is $x^{15} + x^{14} + 1$. Thus, the required measurement matrix (i.e., Gaussian random matrix) can be reorganised. By contrast, under the same conditions, the proposed framework can generate 128 required measurement matrices (i.e., partial circulant matrix). Fig. 4 shows the number of generated measurement matrix when the seed number is increasing. We can see that compared with the reference framework, the ability of generating the measurement matrix of CSSS-FS obtains a significant promotion.

vi. *Security strength:* In this study, we consider the mutual information as the degree of security. Herein, to estimate the entropy and mutual information from the small-sample data in experiments, we use the James–Stein (JS) shrinkage estimator [19] as the estimate method. JS-type shrinkage is a simple

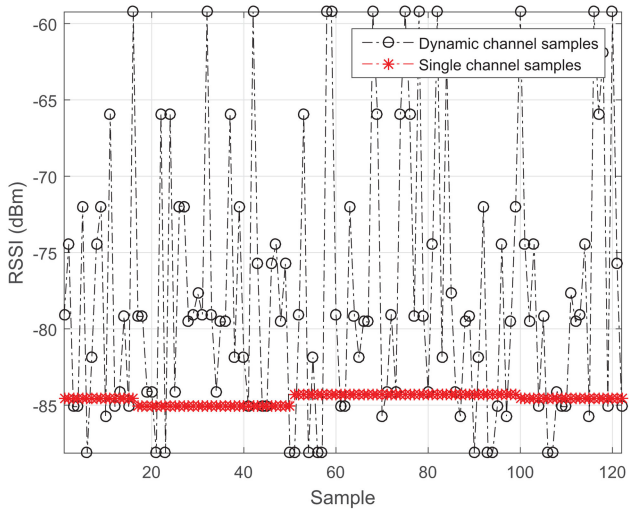

Fig. 3 *Sampling over signal and dynamic channel*

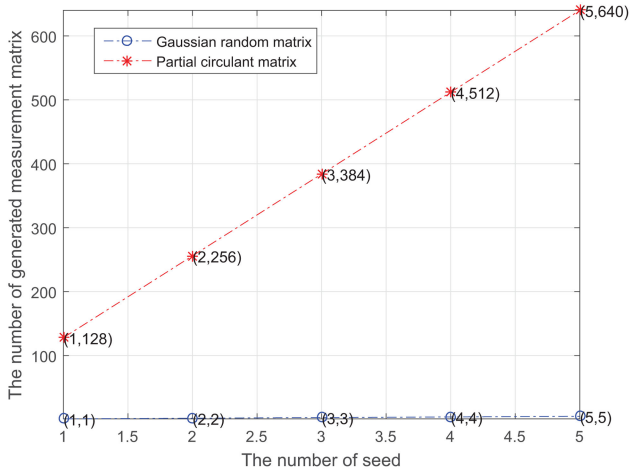

Fig. 4 *Number of generated measurement matrices* (128×256) *under the case of increasing seed*

analytic device to perform regularised high-dimensional inference, and ideally suited for small-sample settings.

In our experiments, the transmitted data before compressing and the decoded data which is reconstructed in CSSS-FS are quantised to obtain the identical finite set. The measurement set contains the mutual information between Alice and Bob while between Alice and Eve. By using the JS shrinkage estimator, the corresponding data are estimated under different participants and different distances.

Table 1 reports the statistics of the results. Each entry in the table is an average value over 1000 records. It shows that when the distance between Alice and Eve is reduced, the corresponding mutual information is increasing. While, the value of the mutual information of Alice–Bob is pretty stable, which is always considerably larger than that of Alice–Eve. It implies that Eve is hard to obtain more useful information about Alice–Bob, where the gap is protected by physical layer property.

We then verify the performance of the proposed and reference framework [9, 10] under the different arithmetic sum of the

Table 1 PDR and reconstructions signal quality class

Distance	Alice–Bob (mean)	Alice–Bob (variance)	Alice–Eve (mean)	Alice–Eve (variance)
10 inches	2.56	0.22	0.02	10.24×10^{-4}
6 inches	2.56	0.22	0.09	15.11×10^{-4}
4 inches	2.56	0.22	0.18	49.35×10^{-4}
2 inches	2.56	0.22	0.25	76.29×10^{-4}

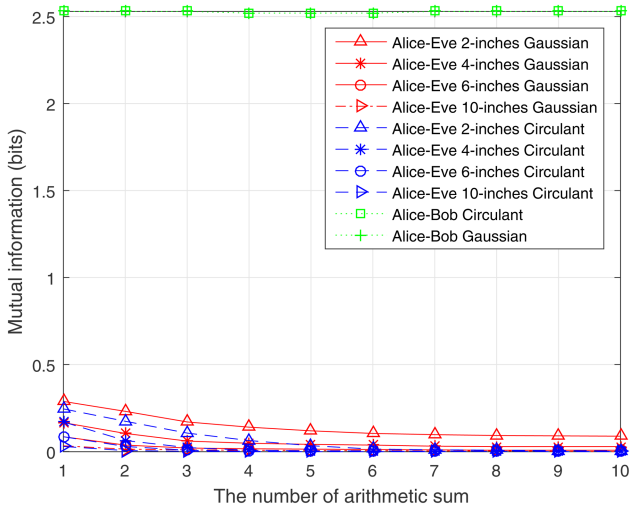

Fig. 5 *Experiment results under different distances and different arithmetic sums*

measurement matrix and the different matrix type. Fig. 5 shows the intuitive results, and it is clear that the legitimate users, i.e. Alice and Bob, have maximum mutual information for all considered conditions. While the mutual information between the illegitimate users is decreased when the number of arithmetic sum is increasing. Moreover, the type of the measurement matrix can affect the performance of the illegitimate users. More specifically, in the case of the partial circulant matrix, all of the curve between Alice and Eve decrease faster and converges to the minimum value. While, in the case of the Gaussian random matrix, these curves are more sensitive to the distance. It shows that the performance of the partial circulant matrix is better than that of the Gaussian random matrix under the same conditions, and through the arithmetic sum of the measurement matrix, the gap between legitimate and illegitimate users can be enlarged.

5.2 Performance analysis in practical application

In this subsection, we use a real-world database, i.e. MIT-BIH arrhythmia database [20], to validate the compression and reconstruction performance in CSSS-FS. MIT-BIH arrhythmia database is most commonly used to study electrocardiograph (ECG) signal, and it consists of two-lead ambulatory ECG recordings from 47 subjects. Without loss of generality, we choose one set of data as our testing object. Moreover, for a quantitative description, we employ two of the most widely used metrics to measure the recovery performance in CS [21], and they contain CR and percentage root-mean-squared difference (PRD).

The CR is defined as

$$\mathrm{CR} = \frac{\theta_{\mathrm{orig}} - \theta_{\mathrm{comp}}}{\theta_{\mathrm{orig}}} \times 100 \tag{9}$$

where θ_{orig} and θ_{comp} represent the number of bits required for the original and compressed signals, respectively. Here the CR is the compressibility of the ECG data, and it also indicates the ratio of radio energy consumption saving [21].

The PRD quantifies the error between the original signal vector $\boldsymbol{x}$ and the reconstructed signal vector $\hat{\boldsymbol{x}}$, given by

$$\mathrm{PRD} = \frac{\| \boldsymbol{x} - \hat{\boldsymbol{x}} \|_2}{\| \boldsymbol{x} \|_2} \times 100 \tag{10}$$

The relationship between the PRD and the diagnostic distortion has been established in [21], and Table 2 lists the resulting classes of very good quality, good quality, uncertain quality, and the corresponding PRDs.

Based on CSSS-FS, we tested heartbeats for 6 s ($N = 2161$) from raw ECG signals. Beat records 103, 114, and 207 were processed by the proposed algorithm and Fig. 6 shows the

Table 2 PRD and reconstructions signal quality class

PRD, %	Reconstructed signal quality
0–2	'very good' quality
2–9	'good' quality
>9	impossible to determine the quality

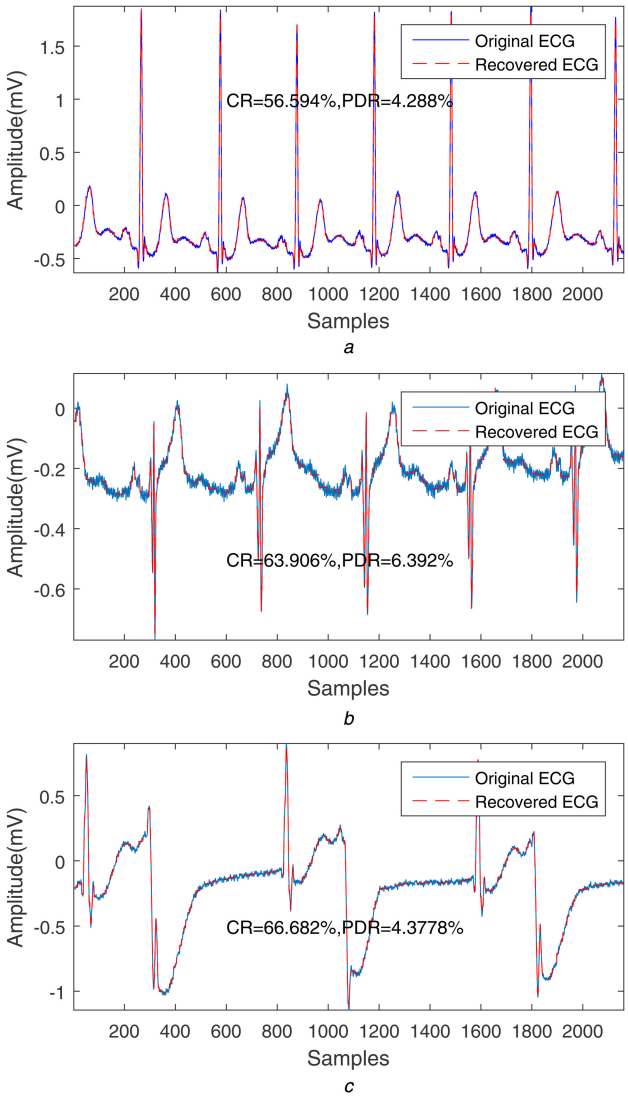

Fig. 6 *Performance of consistency simulation*
(a) Based on record 103, ***(b)*** Based on record 114, ***(c)*** Based on record 207

simulation results. Fig. 6*a–c* illustrate the recovery signal simulation performance compared with the original signal of record-103, 114, and 207, respectively. From these figures, we can see that in case of high CR (CR = 56.59, 63.9, and 66.68%), signal reconstruction quality was impressive (PRD = 4.2, 6.3, and 4.3%, respectively).

Furthermore, the signal reconstruction quality with a larger scale of the CR is shown in Fig. 7. It shows that all of the signal reconstruction quality, i.e. the curves of PRDs, show an upward trend when the CR increases. Specifically, when the CR is in the vicinity of 80%, the PRD curves rise sharply, where the reconstructed signals are unavailable due to the PRD exceeds 10%. Furthermore, from Fig. 7[AQ3] , we can see that the reconstructed signal curves based on the proposed security model, i.e. CSSS-FS, are extremely close to the original curves which are based on normal CS. It implies that our proposed security model is almost no impact on the quality of signal reconstruction. These results show that CSSS-FS can apply to the practical application, and it does not reduce the algorithm performance in CS.

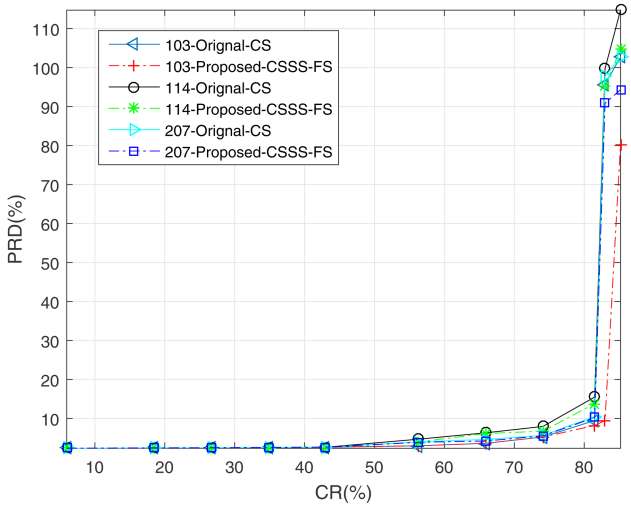

Fig. 7 *Illustration of signal reconstruction quality under original system and CSSS-FS system*

6 Conclusion

In this paper, we concentrated on the CS security model uniting PLS technology for IoT. To solve the problem in the practical application, we proposed a new security model based on resilient functions and circulant matrix, and given a feasible framework, i.e. CSSS-FS. Moreover, by means of experiments and numerical simulation, we validated our presented scheme. In contrast to previous works, our main contribution can be treated as a technology promotion, which is more efficient and applies to more application scenarios.

7 References

[1] Li, S., Xu, L.D., Zhao, S.: 'The internet of things: a survey', *Inf. Syst. Front.*, 2015, **17**, (2), pp. 243–259

[2] Donoho, D.L.: 'Compressed sensing', *IEEE Trans. Inf. Theory*, 2006, **52**, (4), pp. 1289–1306

[3] Li, S., Xu, L.D., Wang, X.: 'Compressed sensing signal and data acquisition in wireless sensor networks and internet of things', *IEEE Trans. Ind. Inf.*, 2013, **9**, (4), pp. 2177–2186

[4] Jafarpour, S., Willett, R., Raginsky, M., *et al.*: 'Performance bounds for expander-based compressed sensing in the presence of Poisson noise', *IEEE Trans. Signal Process.*, 2009, **59**, (9), pp. 513–517

[5] Agrawal, S., Vishwanath, S.: 'Secrecy using compressive sensing'. ITW, 2011, pp. 563–567

[6] Sreedhanya, A., Soman, K.: 'Ensuring security to the compressed sensing data using a steganographic approach', *Bonfring Int. J. Adv. Image Process.*, 2013, **3**, (1), p. 1

[7] Rachlin, Y., Baron, D.: 'The secrecy of compressed sensing measurements'. 46th Annual Allerton Conf. on Communication, Control, and Computing, 2008, 2008, pp. 813–817

[8] Mayiami, M.R., Seyfe, B., Bafghi, H.G.: 'Perfect secrecy via compressed sensing'. Workshop on Communication and Information Theory (IWCIT), 2013, Iran, 2013, pp. 1–5

[9] Dautov, R., Tsouri, G.R.: 'Securing while sampling in wireless body area networks with application to electrocardiography', *IEEE J. Biomed. Health Inform.*, 2016, **20**, (1), pp. 135–142

[10] Tsouri, G.R., Dautov, R.: 'Establishing secure measurement matrix for compressed sensing using wireless physical layer security'. Int. Conf. on Computing, Networking and Communications, 2013, pp. 354–358

[11] Herman, M.A., Strohmer, T.: 'General deviants: an analysis of perturbations in compressed sensing', *IEEE J. Sel. Top. Signal Process.*, 2010, **4**, (2), pp. 342–349

[12] Li, W., Jiang, T., Wang, N.: 'Compressed sensing based on the characteristic correlation of ecg in hybrid wireless sensor network', *Int. J. Distrib. Sens. Netw.*, 2015, **2015**, p. 5

[13] Gopalakrishnan, K.: 'A study of correlation-immune, resilient and related cryptographic functions', 1994

[14] Rauhut, H.: 'Circulant and toeplitz matrices in compressed sensing', *Proc. of SPARS'09*, pp. 1–6

[15] Wilhelm, M., Martinovic, I., Schmitt, J.B.: 'Secure key generation in sensor networks based on frequency-selective channels', *IEEE J. Sel. Areas Commun.*, 2013, **31**, (9), pp. 339–343

[16] Yao, L., Ali, S.T., Sivaraman, V., *et al.*: 'Decorrelating secret bit extraction via channel hopping in body area networks'. IEEE Int. Symp. on Personal Indoor & Mobile Radio Communications, 2012, pp. 1454–1459

[17] Zeng, K., Wu, D., Chan, A., *et al.*: 'Exploiting multiple-antenna diversity for shared secret key generation in wireless networks'. IEEE INFOCOM, 2010, pp. 1–9

[18] Brassard, G., Salvail, L.: 'Secret-key reconciliation by public discussion', *Lect. Notes Comput. Sci.*, 1998, **765**, pp. 410–423
[19] Hausser, J., Strimmer, K.: 'Entropy inference and the James-Stein estimator, with application to nonlinear gene association networks', *Stat. Appl. Genetics Mol. Biol.*, 2008, **10**, (3), pp. 1469–1484
[20] 'Mit-bih arrhythmia database', 2005. Available at: http://www.physionet.org/physiobank/database/mitdb/
[21] Zigel, Y., Cohen, A., Katz, A.: 'The weighted diagnostic distortion (wdd) measure for ecg signal compression', *IEEE Trans. Biomed. Eng.*, 2000, **47**, (11), pp. 1422–1430