IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received January 28, 2019, accepted February 28, 2019, date of publication March 7, 2019, date of current version April 1, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2903574

A Unified Approach for Compression and Authentication of Smart Meter Reading in AMI

YONGGU LEE[1], EUISEOK HWANG[2], AND JINHO CHOI[3], (Senior Member, IEEE)
[1]Research Institute for Solar and Sustainable Energies, Gwangju Institute of Science and Technology, Gwangju 61005, South Korea
[2]School of Mechatronics, Gwangju Institute of Science and Technology, Gwangju 61005, South Korea
[3]School of Information Technology, Deakin University, Burwood, VIC 3125, Australia

Corresponding author: Yonggu Lee (yglee1096@gist.ac.kr)

This work was supported in part by Institute for Information and Communications Technology Promotion (IITP) grant funded by the Korean Government (MSIT) (No.2017-0-00413, Streamlined Secure Communications by Physical Layer Identification in Cellular IoT), and in part by Institute for Information and Communications Technology Promotion (IITP) grant funded by the Korean Government (MSIT) (No.2018-0-00170, Virtual Presence in Moving Objects through 5G).

ABSTRACT In this paper, we propose a unified approach for compression and authentication of smart meter reading in advanced metering infrastructure (AMI). In general, smart meters are urged to send sampled reading signals at a high rate for high-quality services. Meanwhile, power reading signals have to be authenticated to prevent impersonation attacks, which can cause serious economic loss. However, the security in smart grids faces more challenges than conventional human-type communications because of limited hardware resources of a smart meter (e.g., small memory). Motivated by these problems, we study simultaneous compression and authentication for power reading signals in multicarrier systems based on the notion of compressive sensing (CS). The CS-based compression and authentication method are applied to empirically modeled signals with a shared secret key, a measurement matrix in CS between a data concentrator unit (DCU) and a legitimate smart meter. In particular, for authentication, the residual error of a received signal at the DCU is used as a test statistic for hypothesis testing, which determines whether the signal is a legitimate signal or an intrusion signal in the proposed approach. Through the analysis and simulation results, we demonstrate that the CS-based compression approach can be applied to smart meter reading with good energy efficiency. In addition, it is shown that the proposed scheme can obtain a low authentication error probability under reasonable conditions. For example, when the number of subcarriers is 64, the DCU can distinguish legitimate and intrusion smart meters with a probability of $1 - P_E$, where $P_E \leq 10^{-4}$.

INDEX TERMS Authentication, compression, compressive sensing, advanced metering infrastructure.

I. INTRODUCTION

In general, a traditional power grid is an electric system that carries power from electric power generators to a large number of consumers. In a power grid, consumers might be uninformed and non-participative with the power system. While there are some drawbacks (e.g., poor visibility, slow response time, lack of situational awareness) caused by the closed nature of the power system [1] in a power grid, the demand for electricity has gradually increased. To satisfy the demand and address these challenges, smart grid which is an evolved power grid that integrates advanced computing and communication technologies into power grid has emerged [1].

Advanced metering infrastructure (AMI) plays a crucial role in smart grid by enabling bidirectional communications of precise measurements and fast reports between energy consumers and producers [2], [3]. But, there might be challenges for AMI due to the smart meters with limited resources (e.g., small memory, limited bandwidth). For example, it might be difficult for smart meters to support fast sampled power signals, i.e., ∼1 kHz in real-time, necessary for power disaggregation, due to the undesired high energy consumption and the broad bandwidth requirement [4]. In addition, assuming that there exist malicious devices which perform impersonation attacks, it is vulnerable to the attacks due to the nature of wireless broadcast [5]. It may cause serious economic loss and instability of the power system by manipulating power signals. For example, a utility of Puerto Rican suffered huge

The associate editor coordinating the review of this manuscript and approving it for publication was Bin Zhou.

economic losses which are about $400 million by malicious manipulations of smart meter readings [6].

To obtain detailed energy usage patterns in the appliance level, a large amount of power information which enable power disaggregation are necessary. Power disaggregation techniques attempt to identify the individual power contributions of the appliances in the whole power consumption [7]–[11]. However, for those techniques, if the aggregated power signals are sampled at a low rate, it is hard to exactly disaggregate the power signals of individual appliances. While high sampling rates of the power signals make an accurate power analysis for appliances by capturing more characteristics of power signals in AMI, they induce a burden of increasing computational complexity and transmission power for the power meter which has limitations of hardware resources.

To address the high sampling problem, data compression techniques have been extensively studied for many applications [12]–[19]. There are a few lossless compression techniques for power data [12]–[15]. In [12], it is shown that a two-dimensional representation based compression outperforms a one-dimensional representation based compression for compression of power quality event data in terms energy compaction efficiency. In addition, in [15], a new lossless compression method for smart meter readings is proposed using Gaussian approximation, based on a dynamic-nonlinear learning technique. In lossy compression, various power signal compression methods have been studied. A wavelet based data compression approach is presented in [17], where a wavelet transform based multiresolution analysis which enables power signal compression and denoising is investigated. In [18], a new compression method for power signals using singular value decomposition (SVD) is presented. This method is superior to wavelet based compression techniques in terms of the reconstruction performance, while it has high complexity for data compression which may impose a heavy burden on a smart meter. In [19], using the notion of compressive sensing (CS), the power signals of a smart meter are compressed to reduce the computational complexity of a smart meter in AMI.

CS has attracted considerable attention in electrical engineering, applied mathematics, statistics and computer science [20]. Recently, CS has been applied to the information security field in [21]–[36]. In particular, two aspects have been studied: the theoretical aspect and the application aspect [37]. In the theoretical aspect, CS based encryption, which uses a measurement matrix as a secret key, has been theoretically analyzed in [21]–[27]. In [21], CS based encryption was found to achieve computational secrecy, while it cannot provide perfect secrecy in terms of information security because of the linear property of measurements. To achieve perfect secrecy, some unfeasible assumptions used in [24] have to be made. In addition, CS based encryption is analyzed in terms of robustness and security in [22]. Particularly, as shown in [25], CS cannot guarantee the security of cryptographic standards, but may provide a useful built-in data obfuscation layer. In the application aspect, CS has been applied to various applications (e.g., image, biometric data) for security [28]–[36]. Particularly, image encryption methods using the notion of CS have been studied to enable simultaneous image compression and encryption in [28]–[31]. Furthermore, an image authentication scheme based on CS is considered in [36]. In [36], a CS based authentication mechanism that uses a tag signal for image authentication is presented, but the study does not provide any theoretical basis. In particular, a threshold for hypothesis testing and theoretical performance analysis is not given in [36].

In smart grid, authentication techniques have been studied mostly in network and application layers [38]–[40]. However, these techniques may not be suitable for some smart meters with limited resources such as low computational capability and limited bandwidth due to high complexity. To address this issue, physical layer authentication methods based on the dynamic physical characteristics (e.g., channel, analog front-end (AFE)) have been considered. Channel based physical layer authentication [41] uses the time-variant channel state information (CSI). In addition, the phases of multicarrier channels are used for secure physical-layer challenge-response authentication with a shared secret key in multicarrier systems [42]. In [43], channel coding is employed to mitigate the difference between the two estimated channels, which are used for physical-layer challenge-response authentication. The hardware imperfection of AFE which causes input and output (I/O) imbalance, phase offset error, carrier frequency offset error can be used for physical layer authentication [44]. However, if the characteristics of AFE and channel for an intrusion node are close to those of a legitimate transmitter, the authentication techniques may result in relatively poor authentication performance. In [45], a tag signal is concurrently transmitted for a stealth authentication with a message signal.

In this paper, we propose a unified approach of compression and authentication for smart meter readings based on the notion of CS in a multicarrier system. The power reading signal of a smart meter is considered as an aggregated signal of power signals of home appliances. Based on the models for power signals of home appliances in [46], empirical modeling of the aggregated signal is considered for the smart meter reading. To this end, we collect power consumption data for home appliances to empirically model using the smart plugs (PM-B310-W2) in the laboratory. Unlike [32] where a CS based compression is applied to power signals of home appliances, a CS based compression is applied to the aggregated power reading signal based on the empirical modeling. Furthermore, we consider an efficient authentication scheme in conjunction with CS based compression to reduce a burden of the signal processing of the smart meter for compression and authentication in the proposed scheme. In particular, by using a measurement matrix as a secret key for the authentication, we show that the power reading signal can be authenticated in physical layer. It means that a data concentrator unit (DCU) which collects the power data from the smart meter can discard an intrusion signal in physical layer directly without upper layer processing, which can effectively reduce the burden of network due to unnecessary traffics caused by malicious intrusion meters. Through the theoretical analysis and numerical simulations, we can show that the CS based compression for smart meter readings has a better performance than wavelet based compression in terms of energy efficiency. In addition, the proposed scheme can obtain a low authentication error probability under various conditions (e.g., the number of subcarriers, signal to noise ratio (SNR)).

The main contributions of this paper are summarized as follows. First, a unified approach for compression and authentication is proposed under a transmission framework for smart meter reading in AMI. Any unified approach for compression and authentication of power reading signals has not been studied in the literature yet, while a similar work can be found in [32], where the compression is combined with encryption. Compared with the approach in [32], the proposed approach focuses more on authentication to prevent impersonation attacks. Secondly, for the authentication performance analysis, we derive a guaranteed authentication error probability for the proposed scheme. In this paper, the authentication error probability is the probability that the DCU makes an incorrect decision of hypothesis testing for the authentication. Here, a power of a residual error for a received signal at the DCU is used for hypothesis testing. Note that this paper is an extension of its conference version in [19] with new material, including the discussion on CS based authentication, which is our main contribution (in Section IV).

The remainder of this paper is organized as follows. In Section II, we explain motivation of our works in this paper. Section III presents a CS based compression scheme based on an empirical modeling of a power reading signal at a smart meter in AMI. We propose a unified approach of compression and authentication for smart meter reading in Section IV. The simulation results to evaluate the performance of the proposed scheme are presented in Section V. Finally, concluding remarks are given in Section VI.

Notation: Upper-case and lower-case boldface letters are used for matrices and vectors, respectively. $\mathbf{A}^{\mathrm{H}}$ and $\mathbf{A}^{\mathrm{T}}$ denote the Hermitian and transpose of $\mathbf{A}$, respectively. For a matrix $\mathbf{X}$ (a vector $\mathbf{x}$), $[\mathbf{X}]_m$ ($[\mathbf{x}]_m$) and $[\mathbf{X}]_{m,n}$ represent the m-th row (element, resp.) and the m-th row and the n-th column element of $\mathbf{X}$, respectively. The p-norm of a vector $\mathbf{a}$ is denoted by $\|\mathbf{a}\|_p$ (If $p = 2$, the norm is denoted by $\|\mathbf{a}\|$ without the subscript). $\mathcal{CN}(\mathbf{a}, \mathbf{R})$ represents the distribution of circularly symmetric complex Gaussian (CSCG) random vectors with mean vector $\mathbf{a}$ and covariance matrix $\mathbf{R}$.

II. MOTIVATION

A key element in AMI is smart meters which continuously record power consumptions of users and send power reading signals to a DCU to permit accurate power estimation and control of an administrator. According to forecast in Frost & Sullivan, it is predicted that 126 million smart meters will be installed annually until 2024 [47]. As the number of smart meters rapidly increases, simple and low-cost smart metering systems for large scale distributed monitoring become an important issue [48]. In the system, a large number of smart meters deployed in AMI collect real time power reading information from home appliances. In this paper, we consider a large scale smart metering system which supports power disaggregation with smart meters of low-cost.

In general, low-cost smart meters have limited hardware resources (e.g., small memory, low computational capacity). However, the conventional cryptographic mechanisms for smart grid require high complexity and large signaling overhead which puts too much burden on the constrained smart meter resources. For example, if a smart meter uses MSP430F471 microcontroller with 16 MHz central processing unit (CPU) and 8 kB random access memory and 120 kB flash memory, conventional authentication schemes such as RSA and elliptic curve digital signature algorithm (ECDSA) would not be recommended for the smart meter due to required resources for computation [38]. In addition, low energy consumption is one of the primary requirements to a smart meter since its energy cost for the utility under main power becomes non-negligible by the massive deployment of the meters running all days, and it needs to run for prolonged periods under battery power for reporting the main power loss or other emergency events [49]. Furthermore, the energy cost and battery limit can be more critical for the multi-utility smart meter which measures water, gas as well as electricity by the extra energy consumption with cryptographic mechanisms. Therefore, a lightweight authentication scheme is preferred for smart meter due to the lowered energy consumption in AMI by reducing the computational complexity. Motivated by the problem, we propose a unified approach for compression and authentication in physical layer. The procedures for compression can be used for authentication based on the notion of CS. The details are provided in Section IV.

III. CS BASED COMPRESSION IN AMI

As illustrated in Fig. 1, a smart meter captures aggregated power consumption signals of home appliances, and transmits the aggregated signal to the DCU [46]. In addition, as mentioned earlier, the power reading signal has to be compressed with a high sampling for an accurate power analysis in AMI. In this section, we study the application of CS to the compression of power reading signals at smart meters after an empirical modeling.

A. EMPIRICAL MODELING OF POWER READING SIGNALS IN AMI

A power reading signal transmitted from a smart meter is empirically modeled by analysis of power profiles of individual home appliances. Home appliances have different power consumption profiles due to their unique electric loads. In addition, the average hours and intervals of the use of home appliances are also quite different in accordance with the purpose of the devices. Thus, it is an intractable problem

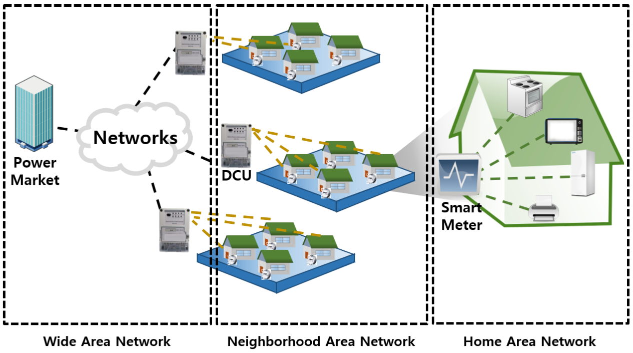


FIGURE 1. Advanced metering infrastructure.

to model the power signal for all different appliances, and an empirical analysis based modeling scheme as in [46] is employed in this study. In addition, proper parameter values and probability distributions for the modeling are determined through empirical analysis of the collected data in this paper. Power waveforms of individual home appliances are captured by a set of smart plugs (PM-B310-W2) simultaneously, and collected through the wireless connections. These waveforms are jointly analyzed for the empirical modeling [19].

Four different models are used for modeling associated with power signals of the home appliances, which are on-off, on-off decay, stable min-max, and random models [46]. We determine parameter values and probability distributions in the models by analyzing the collected power data. The details on the models can be found in [19].

The power reading signal of a smart meter, $p(t)$, can be regarded as an aggregation of power consumptions of active home appliances at time t and expressed as

$$p(t) = \sum_i \epsilon_i(t) p^{(i)}(t), \tag{1}$$

where $\epsilon_i(t) \in \{0, 1\}$ is the activation of the i-th home appliance and $p^{(i)}(t)$ is the power consumption for the i-th home appliance. For simplicity, we assume that the i-th home appliance is used from $t^{(i)}_{init}$ to $t^{(i)}_{end}$. In addition, an activation time of the i-th home appliance, $t^{(i)}_{active} (= t^{(i)}_{end} - t^{(i)}_{init})$ is assumed to be decided by the statistical analysis of the collected power data in this paper.

In Fig. 2 [19], we show the empirical modeling of power signals and measured power signals from smart plugs. The measurement errors due to imperfect hardware are also shown in the captured data. Based on the modeling for power signals of home appliances as described in (1), we can easily obtain the modeling information for smart meter signals.

B. CS BASED POWER READING SIGNAL COMPRESSION IN AMI

To apply the notion of CS to the compression of power reading signals, their sparsity is necessary. In [32], power reading signals of home appliances can be transformed to sparse signals with a proper representation matrix by using empirical power consumptions. But, it considers a compression of power consumptions of home appliances, not a power reading signal of a smart meter. In addition, it only relies on the collected data, not modeled data. On the other hand, in this paper, we design a representation matrix to obtain high sparsity for a power reading signal based on the modeling. In this section, we omit the time index, t for convenience, and represent a power reading signal of a smart meter as follows:

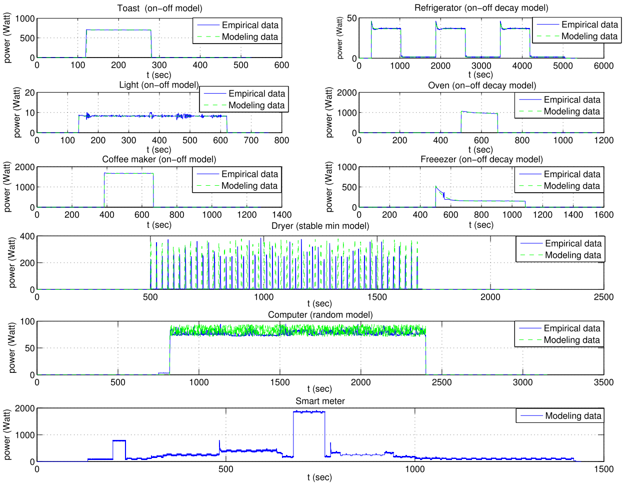


FIGURE 2. Comparison between empirical collected power signal and modeling power signal for home appliances and a smart meter [19].

$$\mathbf{p} = \mathbf{\Psi s}, \tag{2}$$

where $\mathbf{\Psi} \in \mathbb{R}^{N \times N}$ is a representation matrix. Here, $\mathbf{s} \in \mathbb{R}^{N \times 1}$ is Q-sparse. There exists a trade-off between sparsity and the accuracy of the approximation. Ultimately, it is related to a trade-off between a compression ratio and reconstruction error. In this paper, Q is the minimum q subject to the energy of $\mathbf{s}_{(q)}$ contains more than $\kappa(\%)$ of energy of $\mathbf{s}$ (i.e., $\|\mathbf{s}_{(q)}\|^2 \geq \kappa\|\mathbf{s}\|^2$), where κ is a certain threshold. Here, κ is determined according to a target compression ratio.

There are general representation matrices which make a signal transformed to a sparse signal. For example, discrete cosine transform (DCT) and discrete Fourier transform (DFT) matrices [50] can efficiently compress a periodic signal. In addition, adjacent difference transform (ADT) matrix is suitable for a signal with duty-cycled property [32]. Wavelet transform [51] matrix generally is used for other applications (e.g., image compression). In [32], it is shown that Haar wavelet transform (HWT) is a best default basis without prior knowledge of the signal structure for a power signal of a home appliance. In this paper, to compare the performances of the representation matrices for a power reading signal of a smart meter, the sparsity of the compressed power reading signal is calculated by using modeled data.

We collect L samples by a simple projection with measurement vectors $\{\boldsymbol{\phi}_l\}_{1 \leq l \leq L}$ as $x_l = \boldsymbol{\phi}_l^{\mathrm{T}} \mathbf{p}$ (note that the samples are transmitted through L subcarriers.). Thus, the compressed power vector is given by

$$\mathbf{x} = \mathbf{\Phi}\mathbf{\Psi}\mathbf{s}, \tag{3}$$

where $\mathbf{\Phi} = [\boldsymbol{\phi}_1^{\mathrm{T}}; \ \boldsymbol{\phi}_2^{\mathrm{T}}; \ \cdots ; \ \boldsymbol{\phi}_L^{\mathrm{T}}]$ is the measurement matrix. A design of the measurement matrix plays an important role in CS to obtain a good reconstruction performance. The coherence is one of measures to evaluate whether or not a measurement matrix is good [52]. That is, as the coherence between a measurement matrix and a representation matrix is small, it is expected to reconstruct the original signal well in CS. As shown in [52], if $\mathbf{\Psi}$ is the Fourier basis and $\mathbf{\Phi}$ is the canonical or spike basis, e.g., $\phi_{l,k} = \delta(k-l)$, the coherence between $\mathbf{\Phi}$ and $\mathbf{\Psi}$ is given by $\mu(\mathbf{\Phi}, \mathbf{\Psi}) = 1$, where $\mu(\mathbf{\Phi}, \mathbf{\Psi}) \in [1, \sqrt{N}]$. It means that if the representation matrix is a DCT matrix, the measurement matrix using spike basis becomes a best measurement matrix in terms of the coherence. However, for the other representation matrices, a Gaussian measurement matrix (e.g., $\phi_{l,k} \sim \mathcal{CN}(0, \frac{1}{L})$) or a Bernoulli measurement matrix (e.g., $\phi_{l,k} \in \{-\frac{1}{\sqrt{L}}, \frac{1}{\sqrt{L}}\}$) may have a better reconstruction performance, compared to a spike measurement matrix. Meanwhile, for the CS recovery, we use orthogonal matching pursuit (OMP) algorithm [53] which is a low complexity greedy algorithm. In most CS recovery algorithms, the sufficient number of measurements is an important parameter for the successful recovery. At a high SNR, the required number of measurements, L, for the successful recovery is bounded as follows: $L \geq CQ \ln\left(\frac{N}{\delta}\right)$, where C is a constant and $\delta \in (0, 0.36)$ [53]. Then, OMP can reconstruct the signal with probability exceeding $1-\delta$.

It is important to evaluate the compression technique with a proper performance metric. In this paper, sparsity, the percentage root mean square difference (PRD) and energy consumption are considered for the performance metrics. Sparsity, which influences CS recovery performance, is given by $\zeta = \frac{N-Q}{N}$. The recovery problem in CS is underdetermined because there are fewer measurements than unknowns in CS. However, the problem in CS can be solved by using the sparsity. Here, $\mathbf{s}$ belongs to the set obtained by the union of all the $\binom{N}{Q} = \frac{N!}{Q!(N-Q)!}$ Q-dimensional subspaces in $\mathbb{R}^{N \times 1}$. The sparsity information (i.e., large number of null entries) is powerful a-priori information that can be exploited in the solution of an underdetermined linear system. Then, we can consider a simple example with $L = 2$ and $N = 3$ as follows:

$$x_1 = \frac{1}{\sqrt{2}} s_1 + \frac{1}{\sqrt{3}} s_2 + \frac{1}{\sqrt{6}} s_3,$$
$$x_2 = \frac{1}{\sqrt{3}} s_1 + \frac{1}{2} s_2 + \frac{\sqrt{5}}{\sqrt{12}} s_3.$$

Thus, if $Q = 2$, a Q-sparse vector can belong to only one of the three coordinate planes. On the other hand, if $Q = 1$, the sparse vector can belong to only one of the three axes. Then, if Q is small (i.e., higher sparsity), the search space can be reduced. The strongly reduced search space is the key that allows the retrieval of sparse vectors using a small number of measurements. The PRD that quantifies the reconstruction quality is defined as $\mathrm{PRD} = \frac{\|\mathbf{s}-\hat{\mathbf{s}}\|_2}{\|\mathbf{s}\|_2} \times 100$, where $\hat{\mathbf{s}}$ is the reconstructed signal. The energy consumption is a very important metric to increase energy efficiency of a smart meter. To evaluate the energy efficiency, an energy model [16] for compression and transmission is considered, as follows: $E_{total} = E_{comp} + E_{tx}$, where E_{comp} and E_{tx} are amounts of energy for compression and transmission, respectively. As shown in [54], $E_{comp} = B_{flop}E_{flop}$, where B_{flop} and E_{flop} are the number of flops for compression and the amount of energy for processing 1 flop, respectively. In addition, $E_{tx} = B_{bit}E_{bit}$, where B_{bit} and E_{bit} are the number of bits for transmission and the amount of energy spent in transmitting 1 bit, respectively. Assuming a simple energy consumption model for transmission in [55], an energy consumption for 1 bit transmission is given by $E_{bit} = E_{t_{elec}} + \epsilon_{amp}d^2$, where $E_{t_{elec}}$ is the amount of energy for transmitter circuitry and ϵ_{amp} is the amount of energy for a transmit amplifier and d is the communication distance, respectively.

As mentioned earlier, the compressed power reading signal should be authenticated to prevent an intrusion meter (IM) forging power consumptions in AMI. To this end, we also use the notion of CS which is used for the power reading signal compression in this section. In Section IV, we will investigate details on the CS based authentication which are our main contributions in this paper.

IV. CS BASED AUTHENTICATION IN AMI

In this section, we consider CS based physical layer authentication for a power metering system consisting of a DCU and multiple smart meters in the presence of malicious intrusion smart meters which are called intrusion meters. For simplicity, we only consider one legitimate meter (LM) and one IM in this paper. Then, the CS based authentication which enables simultaneous compression and authentication is considered as illustrated in Fig. 3.

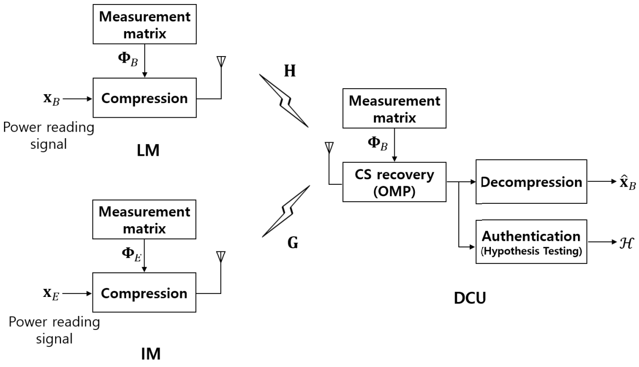


FIGURE 3. System model of the CS based authentication in AMI.

A. AUTHENTICATION SCENARIO

We present a physical layer authentication scheme that enables simultaneous compression and authentication based on the notion of CS in AMI. In the CS based authentication, a measurement matrix is used as a secret key. In general, linear feedback shift registers (LFSRs) which are typically used for a low-complexity implementation of a pseudo-random generator can be employed to generate the measurement matrix for security. In this case, an initial vector in the LFSRs becomes a secret key to generate the measurement matrix. We consider the following scenario for the physical layer authentication:

- The DCU performs an initial authentication by using conventional cryptography based authentication.
- The DCU generates a measurement matrix as a secret key by using LFSRs.
- By using a physical layer security scheme [56] based on CSI between the DCU and the LM, the DCU securely transmits an initial vector of LFSRs to generate measurement matrix.
- Based on the shared measurement matrix, the DCU determines whether or not a received signal is from the LM by using the CS based authentication.

As mentioned earlier, we consider a physical layer security scheme in [56] to share a measurement matrix between a DCU and an LM. As [56], subcarriers are interleaved according the sorted order of their channel gains in time division duplex (TDD) mode. Then, based on the channel reciprocity, the LM can derive the interleaving pattern initiated by the DCU, while the interleaving pattern is unknown to the IM. Then, the DCU can securely send an initial vector which is used to generate the measurement matrix to the LM. Based on the scenario, the CS based authentication between the DCU and the LM is considered in the presence of the IM.

B. CS BASED PHYSICAL LAYER AUTHENTICATION

We consider a multicarrier system for transmissions from an LM to a DCU with L subcarriers. Throughout the paper, the LM sends a block of signals denoted by $\mathbf{x}_B \in \mathbb{C}^{L\times 1}$ over L subcarriers, where $\mathbf{x}_B$ contains a message of smart meter reading. We also consider the case that an IM transmits a forged signal block, denoted by $\mathbf{x}_E \in \mathbb{C}^{L\times 1}$, over L subcarriers with the aim at impersonating the LM. Then, the received signal at the DCU is given by

$$\mathbf{y} = \begin{cases} \mathbf{H}\mathbf{x}_B + \mathbf{n}, & \text{if the LM transmits,} \\ \mathbf{G}\mathbf{x}_E + \mathbf{n}, & \text{if the IM transmits,} \end{cases} \tag{4}$$

where $\mathbf{n} \sim \mathcal{CN}(0, \sigma^2\mathbf{I})$ is the background noise term. Here, $\mathbf{H} = diag(H_0, \ldots, H_{L-1})$ and $\mathbf{G} = diag(G_0, \ldots, G_{L-1})$ are the diagonal channel matrices from the LM and IM to the DCU, respectively, where H_l and G_l are the l-th frequency-domain channel coefficients of the LM and IM, respectively, which are assumed to be CSCG random variables, i.e., $H_l \sim \mathcal{CN}(0, \sigma_h^2)$ and $G_l \sim \mathcal{CN}(0, \sigma_g^2)$.

As mentioned in Section III, power reading signals can be transformed into sparse signals with a proper representation matrix (e.g., HWT). It means that power reading signals can be compressed using the notion of CS. Meanwhile, for the authentication, we can design an authentication matrix, $\mathbf{\Phi}_B$ which has to be shared between the LM and the DCU as a secret key for the authentication. As mentioned earlier, we consider that the DCU transmits an initial vector used to generate the measurement matrix to the LM based on a physical layer security scheme [56]. Then, the authentication matrix is unknown to the IM, while the representation matrix, $\mathbf{\Psi}$, is known. Let $\mathbf{x}_B = \mathbf{\Phi}_B\mathbf{\Psi}\mathbf{s}_B$ and $\mathbf{x}_E = \mathbf{\Phi}_E\mathbf{\Psi}\mathbf{s}_E$, where $\mathbf{s}_B$ and $\mathbf{s}_E$ are the sparse signals at the LM and IM, respectively, and $\mathbf{\Phi}_B$ and $\mathbf{\Phi}_E$ are the authentication matrices for the LM and the IM, respectively. From (4), the received signal at the DCU can be represented as follows:

$$\mathbf{y} = \begin{cases} \mathbf{H}\mathbf{\Phi}_B\mathbf{\Psi}\mathbf{s}_B + \mathbf{n}, & \text{if the LM transmits,} \\ \mathbf{G}\mathbf{\Phi}_E\mathbf{\Psi}\mathbf{s}_E + \mathbf{n}, & \text{if the IM transmits.} \end{cases} \tag{5}$$

Here, $\mathbf{\Phi}_B$ used as a secret key is known to the DCU, while $\mathbf{\Phi}_E$ is unknown to the DCU. Meanwhile, the DCU should know the channels to detect the transmitted signals. To this end, the smart meters transmit pilot signals before the power reading signals. Throughout the paper, we assume that the channels for the LM and IM are perfectly estimated, respectively. For the LM, the IM which performs impersonation attacks also transmits a pilot signal to estimation channels between the DCU and the IM. Then, as shown in [57], the channel estimation errors which can be influenced by noises can sufficiently be reduced by using channel estimation techniques based on denoising strategies. So, the channels can be estimated with negligibly small error. For convenience, let $\mathbf{D}$ be the estimated channel at the DCU, i.e., $\mathbf{D} = \mathbf{H}$ or $\mathbf{G}$ if the LM or IM transmits, respectively.

The measurement matrix which is unknown to the IM can be estimated by the IM's known plaintext attacks [58] due to the linear property of CS. So, as [58], if an artificial noise is used, it makes the attacks difficult. To minimize the performance degradation at the DCU, the artificial noise is selectively transmitted in the frequency domain based on known CSI. It becomes difficult for the IM to perform attack with a fraction of the received signals due to artificial noise [58]. Then, it ensures a guarantee of secrecy in terms of a probability of successful attack under certain conditions. For example, if SNR = 8dB, $L = 64$, a successful attack probability, denoted by P_{SA}, can be as low as $P_{SA} \approx 10^{-12}$. In this paper, the method using artificial noise is not studied due to the page limitation.

C. HYPOTHESIS TESTING

For the authentication, we consider two hypotheses: $\mathcal{H}_1$ is the hypothesis that the received signal is transmitted by the LM with $\mathbf{\Phi}_B$ and $\mathcal{H}_0$ is the other hypothesis that the received signal is transmitted by the IM with $\mathbf{\Phi}_E$. In this subsection, for tractable analysis, we assume $[\mathbf{\Phi}_B]_{l,n}, [\mathbf{\Phi}_E]_{l,n} \sim \mathcal{CN}(0, \frac{1}{L})$. Note that if multiple LFSRs can be used, each element of a measurement matrix approximately becomes a Gaussian random variable [59]. Since $\mathbf{\Phi}_E$ is randomly generated for the impersonation attack by the IM without any information of $\mathbf{\Phi}_B$, the binary hypothesis problem can be formulated as follows:

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{\Phi} = \mathbf{\Phi}_E \\ \mathcal{H}_1 &: \mathbf{\Phi} = \mathbf{\Phi}_B, \end{aligned} \tag{6}$$

where $\mathbf{\Phi}$ is an authentication matrix used in a smart meter, i.e., $\mathbf{\Phi} = \mathbf{\Phi}_B$ or $\mathbf{\Phi}_E$ if the LM or IM transmits, respectively.

Let $U = (Y, \Phi)$, where Y is a random vector of a received signal and Φ is a random matrix of the authentication matrix used as a secret key. At the DCU, $u = (\mathbf{y}, \mathbf{\Phi}_B)$. Then, in the case of $\mathcal{H}_1$, the realization (i.e., outcome of measurements), u is generated by the joint probability $p(Y, \Phi)$, where Y depends on Φ. On the other hand, in the case of $\mathcal{H}_0$, u is generated by the distribution $p(Y)p(\Phi)$ because the realization of the received signal, $\mathbf{y}$ is generated with $\mathbf{\Phi}_E$ which is perfectly independent of $\mathbf{\Phi}_B$. Then, according to [60], the optimal binary hypothesis testing can be formulated as follows:

$$\begin{aligned}\Lambda &= \log \frac{P_{U|\mathcal{H}_1}(u)}{P_{U|\mathcal{H}_0}(u)} \\ &= \log \frac{p(\mathbf{y}|\mathbf{\Phi} = \mathbf{\Phi}_B)}{\sum_{\mathbf{\Phi}_E} p(\mathbf{y}|\mathbf{\Phi} = \mathbf{\Phi}_E)p(\mathbf{\Phi} = \mathbf{\Phi}_E)}. \end{aligned} \tag{7}$$

Unfortunately, since the DCU does not know $\mathbf{\Phi}_E$, a different test statistic has to be used. With CS recovery at the DCU to find the sparse vector $\mathbf{s}$ that minimizes $\|\mathbf{y} - \mathbf{D\Phi}_B\mathbf{\Psi s}\|^2$, the following test statistic can be used:

$$\begin{aligned}\alpha &= \|\mathbf{y} - \hat{\mathbf{\Theta}}\hat{\mathbf{s}}\|^2, \\ &= \|\mathbf{D}(\mathbf{\Phi\Psi s} - \mathbf{\Phi}_B\mathbf{\Psi}\hat{\mathbf{s}}) + \mathbf{n}\|^2, \end{aligned} \tag{8}$$

where $\hat{\mathbf{\Theta}} \in \mathbb{C}^{L\times Q}$ represents the submatrix of $\mathbf{D\Phi}_B\mathbf{\Psi}$ obtained by the Q column vectors corresponding to the support of the recovered sparse signal by a CS algorithm (e.g., the OMP algorithm) and $\hat{\mathbf{s}} = (\hat{\mathbf{\Theta}}^{\mathrm{H}}\hat{\mathbf{\Theta}})^{-1}\hat{\mathbf{\Theta}}^{\mathrm{H}}\mathbf{y}$. Note that in the case of $\mathcal{H}_1$, α becomes small, where $\mathbf{\Phi} = \mathbf{\Phi}_B$, while α becomes relatively large in the case of $\mathcal{H}_0$, where $\mathbf{\Phi} = \mathbf{\Phi}_E$ due to a difference of $\mathbf{\Phi\Psi s} - \mathbf{\Phi}_B\mathbf{\Psi}\hat{\mathbf{s}}$ in (8). Then, α is compared to a threshold, η for the final decision.

For convenience, let $\alpha_B = \alpha|(\mathbf{\Phi} = \mathbf{\Phi}_B)$ and $\alpha_E = \alpha|(\mathbf{\Phi} = \mathbf{\Phi}_E)$. Then, the resulting hypotheses are

$$\begin{aligned}\mathcal{H}_0 &: \alpha = \alpha_E \\ \mathcal{H}_1 &: \alpha = \alpha_B. \end{aligned} \tag{9}$$

Thus, it is important to derive the conditional distributions of α, denoted by $f(\alpha_B)$ and $f(\alpha_E)$ for the LM and IM, respectively, which are used to analyze the security performance of the proposed scheme in terms of detection and false alarm probabilities. However, since $\mathbf{\Phi}_E$ is unknown to the DCU, $\mathbb{E}_{\mathbf{\Phi}_E}[f(\alpha_E)]$ is alternatively used as the distribution of α_E.

1) PROBABILITY DENSITY FUNCTION OF α_B

To obtain the distribution of α_B, we assume that when the LM transmits, the support set of a Q-sparse signal, $\mathbf{s}$ is perfectly recovered by using the OMP algorithm in a high SNR. As shown in [61], the power of residual error becomes

$$\alpha = \mathbf{y}^{\mathrm{H}}\mathbf{y} - \mathbf{y}^{\mathrm{H}}\hat{\mathbf{\Theta}}(\hat{\mathbf{\Theta}}^{\mathrm{H}}\hat{\mathbf{\Theta}})^{-1}\hat{\mathbf{\Theta}}^{\mathrm{H}}\mathbf{y}. \tag{10}$$

Thus, from (5) and (10), the power of residual error for the legitimate signal is given by

$$\alpha_B = \mathbf{s}_B^{\mathrm{H}}\mathbf{\Theta}_B^{\mathrm{H}}\mathbf{U}_B\mathbf{\Theta}_B\mathbf{s}_B + \mathbf{s}_B^{\mathrm{H}}\mathbf{\Theta}_B^{\mathrm{H}}\mathbf{U}_B\mathbf{n} + \mathbf{n}^{\mathrm{H}}\mathbf{U}_B\mathbf{\Theta}_B\mathbf{s}_B + \mathbf{n}^{\mathrm{H}}\mathbf{U}_B\mathbf{n}, \tag{11}$$

where $\mathbf{U}_B = \mathbf{I} - \mathcal{R}_B$ and $\mathbf{\Theta}_B = \mathbf{H\Phi}_B\mathbf{\Psi}$. Here, $\mathcal{R}_B = \hat{\mathbf{\Theta}}(\hat{\mathbf{\Theta}}^{\mathrm{H}}\hat{\mathbf{\Theta}})^{-1}\hat{\mathbf{\Theta}}^{\mathrm{H}}|\mathcal{H}_1$ and trace$\{\mathcal{R}_B\} = Q$ [62]. In the OMP, a support set of the recovered sparse signal which corresponds to $\hat{\mathbf{\Theta}}$ can be determined with the column indices associated with the Q largest correlation coefficients between $\mathbf{y}$ and the columns of $\mathbf{D\Phi}_B\mathbf{\Psi}$. For a high SNR, if the LM transmits, the estimated support set becomes $\hat{T}|\mathcal{H}_1 = T_B$, where T_B is the support set of $\mathbf{s}_B$. Then, $\hat{\mathbf{\Theta}}\hat{\mathbf{s}}|\mathcal{H}_1 = \mathbf{\Theta}_B\mathbf{s}_B$.

Thus, from (11), we can rewrite the power of the residual error as follows:

$$\alpha_B = \mathbf{n}^{\mathrm{H}}\mathbf{U}_B\mathbf{n}, \tag{12}$$

where $\mathbf{s}_B^{\mathrm{H}}\mathbf{\Theta}_B^{\mathrm{H}}\mathbf{n} = \mathbf{s}_B^{\mathrm{H}}\mathbf{\Theta}_B^{\mathrm{H}}\mathcal{R}_B\mathbf{n}$, $\mathbf{n}^{\mathrm{H}}\mathbf{\Theta}_B\mathbf{s}_B = \mathbf{n}^{\mathrm{H}}\mathcal{R}_B\mathbf{\Theta}_B\mathbf{s}_B$ and $\mathbf{s}_B^{\mathrm{H}}\mathbf{\Theta}_B^{\mathrm{H}}\mathbf{\Theta}_B\mathbf{s}_B = \mathbf{s}_B^{\mathrm{H}}\mathbf{\Theta}_B^{\mathrm{H}}\mathcal{R}_B\mathbf{\Theta}_B\mathbf{s}_B$. Then, $\frac{\alpha_B}{\iota_B}$ follows the Chi-square distribution with $2L$ degrees of freedom, and the probability density function of α_B is given by

$$f(\alpha_B) = \frac{1}{2^L\Gamma(L)}\left(\frac{\alpha_B}{\iota_B}\right)^{L-1} e^{-\frac{\alpha_B}{2\iota_B}}, \tag{13}$$

where $\Gamma(x) = (x-1)!$ and $\iota_B = \frac{(L-Q)\sigma_h^2\sigma^2}{2L}$. In addition, the mean and variance of α_B are given by $\mu_B = (L-Q)\sigma_h^2\sigma^2$ and $\sigma_B^2 = \frac{(L-Q)^2\sigma_h^4\sigma^4}{L}$, respectively.

2) PROBABILITY DENSITY FUNCTION OF α_E

Unlike the power of residual error for the legitimate signal, it is hard to draw the exact probability density function of α_E due to the unknown $\mathbf{\Phi}_E$. From (5) and (10), the power of residual error for the intrusion signal can be represented as

$$\begin{aligned}\alpha_E = \mathbf{s}_E^{\mathrm{H}}\mathbf{\Theta}_E^{\mathrm{H}}\mathbf{U}_E\mathbf{\Theta}_E\mathbf{s}_E + \mathbf{s}_E^{\mathrm{H}}\mathbf{\Theta}_E^{\mathrm{H}}\mathbf{U}_E\mathbf{n} \\ + \mathbf{n}^{\mathrm{H}}\mathbf{U}_E\mathbf{\Theta}_E\mathbf{s}_E + \mathbf{n}^{\mathrm{H}}\mathbf{U}_E\mathbf{n}, \end{aligned} \tag{14}$$

where $\mathbf{\Theta}_E = \mathbf{G\Phi}_E\mathbf{\Psi}$ and $\mathcal{R}_E = \hat{\mathbf{\Theta}}(\hat{\mathbf{\Theta}}^{\mathrm{H}}\hat{\mathbf{\Theta}})^{-1}\hat{\mathbf{\Theta}}^{\mathrm{H}}|\mathcal{H}_0$.

As $\hat{T}|\mathcal{H}_1$, when the IM transmits, the estimated support set, denoted by $\hat{T}|\mathcal{H}_0$, can be determined with the column indices of the Q largest correlation coefficients between $\mathbf{y}$ and the columns of $\mathbf{D\Phi}_B\mathbf{\Psi}$. However, unlike $\hat{T}|\mathcal{H}_1$, for a high SNR, $\hat{T}|\mathcal{H}_0 \neq T_E$, where T_E is the support set of $\mathbf{s}_E$ because $\mathbf{\Theta}_E$ is different with $\mathbf{D\Phi}_B\mathbf{\Psi}$. In the OMP, among N columns of $\mathbf{D\Phi}_B\mathbf{\Psi}$, the Q columns of the largest correlations with $\mathbf{y}$ are selected for the estimated support set. Let ν_q denote the mean of the q-th largest correlation coefficient among N correlation coefficients between $\mathbf{y}$ and the columns of $\mathbf{D\Phi}_B\mathbf{\Psi}$. Here, $\nu_1 \geq \nu_2 \geq \cdots \geq \nu_N$. Let $\mathcal{L} = \{\nu_1, \nu_2, \cdots, \nu_Q\}$ denote the set of correlation coefficients corresponding to the columns of $\hat{T}|\mathcal{H}_0$. Thus, if the elements of $\mathcal{L}$ are sufficiently large, the resulting α_E becomes small. Then, if we assume that all of the columns of $\mathbf{D\Phi}_B\mathbf{\Psi}$ have ν_1 correlation coefficients with $\mathbf{y}$ (i.e., the all elements of $\mathcal{L}$ are assumed to be ν_1), a lower bound on α_E, which is denoted by $\underline{\alpha}_E$, can be obtained. Then, for the case of $\underline{\alpha}_E$, $\mathbb{E}_{\mathbf{\Phi}_E}\left[\hat{\mathbf{\Theta}}\hat{\mathbf{s}}|\mathcal{H}_0\right]$ can be represented as

$$\mathbb{E}_{\mathbf{\Phi}_E}\left[\hat{\mathbf{\Theta}}\hat{\mathbf{s}}|\mathcal{H}_0\right] = \nu_1\varepsilon(\mathbf{\Theta}_E\mathbf{s}_E + \mathbf{n}) + \sqrt{1-\nu_1^2}\mathbf{E}\hat{\mathbf{s}}_E, \tag{15}$$

where $\varepsilon = \sqrt{\frac{\frac{Q}{L}}{\frac{Q}{L}+\sigma^2}}$, $\hat{\mathbf{s}}_E = (\mathbf{E}^{\mathrm{H}}\mathbf{E})^{-1}\mathbf{E}^{\mathrm{H}}\mathbf{y}$, $\nu_1 \in [0, 1]$ and $[\mathbf{E}]_{l,n} \sim \mathcal{CN}\left(0, \frac{1}{L}\right)$ is independent of $\mathbf{\Theta}_E$. Then, from (8) and (15), the lower bound on the power of residual error for the intrusion signal, $\underline{\alpha}_E$, is given by

$$\begin{aligned}\underline{\alpha}_E &= (1-\nu_1\varepsilon)(\mathbf{s}_E^{\mathrm{H}}\mathbf{\Theta}_E^{\mathrm{H}}\mathbf{\Theta}_E\mathbf{s}_E + \mathbf{s}_E^{\mathrm{H}}\mathbf{\Theta}_E^{\mathrm{H}}\mathbf{n} + \mathbf{n}^{\mathrm{H}}\mathbf{\Theta}_E\mathbf{s}_E + \mathbf{n}^{\mathrm{H}}\mathbf{n}) \\ &\quad -\sqrt{1-\nu_1^2}(\mathbf{s}_E^{\mathrm{H}}\mathbf{\Theta}_E^{\mathrm{H}}\mathcal{R}_E\mathbf{\Theta}_E\mathbf{s}_E + \mathbf{s}_E^{\mathrm{H}}\mathbf{\Theta}_E^{\mathrm{H}}\mathcal{R}_E\mathbf{n} \\ &\quad +\mathbf{n}^{\mathrm{H}}\mathcal{R}_E\mathbf{\Theta}_E\mathbf{s}_E + \mathbf{n}^{\mathrm{H}}\mathcal{R}_E\mathbf{n}), \qquad (16)\end{aligned}$$

where $\mathcal{R}_E = \mathbf{E}(\mathbf{E}^{\mathrm{H}}\mathbf{E})^{-1}\mathbf{E}^{\mathrm{H}}$. From (16), $\mathbb{E}[\mathbf{s}_E^{\mathrm{H}}\mathbf{\Theta}_E^{\mathrm{H}}\mathbf{\Theta}_E\mathbf{s}_E] = Q\sigma_g^2$ and $\mathbb{E}[\mathbf{s}_E^{\mathrm{H}}\mathbf{\Theta}_E^{\mathrm{H}}\mathcal{R}_E\mathbf{\Theta}_E\mathbf{s}_E] = \frac{Q^2}{L}\sigma_g^2$ because $\mathbf{\Theta}_E\mathbf{s}_E$ is independent of $\mathbf{E}\hat{\mathbf{s}}_E$. In addition, $\mathbb{E}[\mathbf{n}^{\mathrm{H}}\mathcal{R}_E\mathbf{n}] = Q\sigma^2$ due to $\mathrm{trace}\{\mathcal{R}_E\} = Q$. The means of the other terms in (16) are zero. Then, the mean and variance of $\underline{\alpha}_E$, respectively, are given by

$$\begin{aligned}\underline{\mu}_E &= (1-\nu_1\varepsilon)Q\sigma_g^2 - \sqrt{1-\nu_1^2}\frac{Q^2}{L}\sigma_g^2 \\ &\quad + ((1-\nu_1\varepsilon)L - \sqrt{1-\nu_1^2}Q)\sigma^2 \qquad (17)\end{aligned}$$

and

$$\begin{aligned}\underline{\sigma}_E^2 &\approx \left(1-\nu_1\varepsilon-\sqrt{1-\nu_1^2}\frac{Q}{L}\right)^2\frac{3Q^2\sigma_g^4}{L} \\ &\quad + \frac{\left(\left((1-\nu_1\varepsilon)L-\sqrt{1-\nu_1^2}Q\right)\sigma^2\right)^2}{L} \\ &\quad + \left(\sqrt{2}(1-\nu_1\varepsilon)-\sqrt{2(1-\nu_1^2)}\frac{Q}{L}\right)^2 Q\sigma_g^2\sigma^2, \qquad (18)\end{aligned}$$

where $\mathrm{Var}[\mathbf{s}_E^{\mathrm{H}}\mathbf{\Theta}_E^{\mathrm{H}}\mathbf{\Theta}_E\mathbf{s}_E - \nu_1\varepsilon\mathbf{s}_E^{\mathrm{H}}\mathbf{\Theta}_E^{\mathrm{H}}\mathbf{\Theta}_E\mathbf{s}_E - \sqrt{1-\nu_1^2}\mathbf{s}_E^{\mathrm{H}}\mathbf{\Theta}_E^{\mathrm{H}}\mathcal{R}_E \mathbf{\Theta}_E\mathbf{s}_E] = (1-\nu_1\varepsilon-\sqrt{1-\nu_1^2}\frac{Q}{L})^2\frac{3Q^2\sigma_g^4}{L}$, $\mathrm{Var}[(1-\nu_1\varepsilon)\mathbf{n}^{\mathrm{H}}\mathbf{n} - \sqrt{1-\nu_1^2}\mathbf{n}^{\mathrm{H}}\mathcal{R}_E\mathbf{n}] = (((1-\nu_1\varepsilon)L-\sqrt{1-\nu_1^2}Q)\sigma^2)^2\frac{1}{L}$ and $\mathrm{Var}[(1-\nu_1\varepsilon)(\mathbf{s}_E^{\mathrm{H}}\mathbf{\Theta}_E^{\mathrm{H}}\mathbf{n}+\mathbf{n}^{\mathrm{H}}\mathbf{\Theta}_E\mathbf{s}_E) - \sqrt{1-\nu_1^2}(\mathbf{s}_E^{\mathrm{H}}\mathbf{\Theta}_E^{\mathrm{H}}\mathcal{R}_E\mathbf{n} + \mathbf{n}^{\mathrm{H}}\mathcal{R}_E\mathbf{\Theta}_E\mathbf{s}_E)] = (\sqrt{2}(1-\nu_1\varepsilon)-\sqrt{2(1-\nu_1^2)}\frac{Q}{L})^2Q\sigma_g^2\sigma^2$. Here, ν_1 is obtained by the Monte-Carlo simulations and $\underline{\sigma}_E^2$ is approximated under the assumption that the covariances of the terms in (16) are negligibly small. By using the Gaussian approximation for a large L, the distribution of $\underline{\alpha}_E$ is given by

$$\mathbb{E}_{\mathbf{\Phi}_E}\left[f(\underline{\alpha}_E)\right] = \varphi\left(\frac{\underline{\alpha}_E - \underline{\mu}_E}{\underline{\sigma}_E}\right), \qquad (19)$$

where $\varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$ is a normal distribution.

The difference of the distributions for the LM and IM makes that the DCU can distinguish whether a received signal is a legitimate signal or an intrusion signal. As Q and SNR increase, the distributions are far away. In particular, from (13) and (17), as SNR $\to \infty$, the mean of α_B approaches zero, while the mean of $\underline{\alpha}_E$ approaches $\left((1-\nu_1\varepsilon)Q-\sqrt{1-\nu_1^2}\frac{Q^2}{L}\right)\sigma_g^2$.

D. AUTHENTICATION ERROR PROBABILITY

To evaluate the proposed authentication scheme, we consider an authentication error probability which is an incorrect decision probability at the DCU and given by

$$P_E = \varrho(1-P_D) + (1-\varrho)P_F, \qquad (20)$$

where P_D and P_F are the detection and false alarm probabilities at the DCU, respectively, and ϱ is a weighting factor ($0 \le \varrho \le 1$). Here, P_D is a probability that when the LM transmits, the DCU decides that the signal is a legitimate signal and P_F is a probability that when the IM transmits, the DCU decides that the signal is a legitimate signal. Thus, P_D and P_F are determined by the probability distributions for α_B and α_E and an authentication threshold denoted by η. It means that $P_D = P(\alpha < \eta|\mathcal{H}_1)$ and $P_F = P(\alpha < \eta|\mathcal{H}_0)$. In this paper, for a fixed target detection probability, P_D°, η is decided using the probability density function of α_B. For a fixed η, the detection probability is given by

$$P_D = F\left(\frac{2L\tau}{(L-Q)\sigma_h^2\sigma^2}, 2L\right), \qquad (21)$$

where $F(x, a)$ is the cumulative distribution function of the Chi-square distribution. Then, for a fixed P_D°, the required threshold becomes

$$\eta = \arg\min_{\eta^\circ} F\left(\frac{2L\eta^\circ}{(L-Q)\sigma_h^2\sigma^2}, 2L\right) \ge P_D^\circ. \qquad (22)$$

Once η is determined for a given P_D°, the false alarm probability can be calculated using the distribution of α_E. As mentioned earlier, it is hard to obtain a closed-form expression for the distribution of α_E. So, we draw an upper bound on the false alarm probability with the distribution of $\underline{\alpha}_E$ and η. Then, for a fixed η, an upper bound on the false alarm probability, denoted by $\bar{P}_F$, is given by

$$\bar{P}_F = 1 - Q\left(\frac{\eta - \underline{\mu}_E}{\underline{\sigma}_E}\right), \qquad (23)$$

where $Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-\frac{u^2}{2}}du$. Then, the upper bound on the authentication error probability can be obtained by applying (21) and (23) to (20).

V. SIMULATION RESULTS

In this section, we present the simulation results for the proposed scheme with the empirical model for smart meter readings. For simulations, we assume $\sigma_h^2 = \sigma_g^2 = 1$. In addition, the SNR is defined as $\frac{\|\mathbf{\Phi\Psi s}\|^2}{\|\mathbf{n}\|^2} = \frac{Q}{L\sigma^2}$. As we mentioned earlier, $Q = \min q$ subject to $\|\mathbf{s}_{(q)}\|^2 \ge \kappa\|\mathbf{s}\|^2$.

Fig. 4 depicts the sparsity of $\mathbf{s}$ denoted by ζ, for the various thresholds of sparse signal approximation, κ, where $N = 128$. For simulations, we generate power signals of smart meters through the model. In addition, we evaluate the sparsity of different representation matrices (e.g., ADT, DCT, DFT, HWT). In this figure, HWT has the highest sparsity, because the power reading signal of a smart meter is not

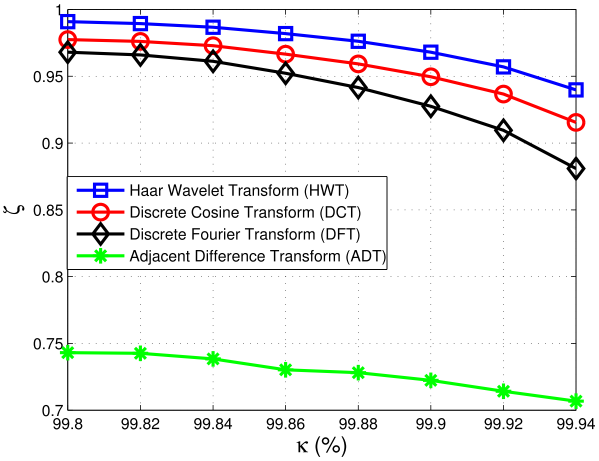


FIGURE 4. Sparsity, denoted by ζ for κ over various representation matrices, where $N = 128$.

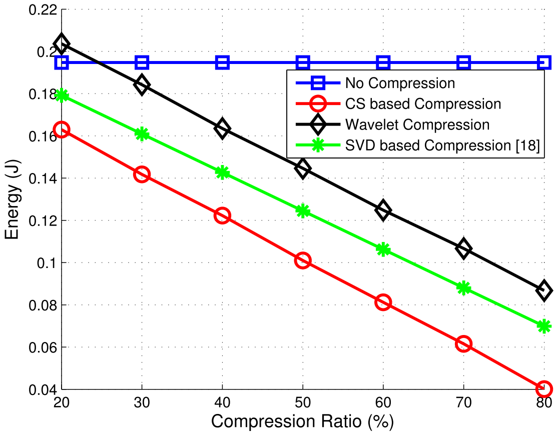


FIGURE 6. Energy consumption for compression and transmission over the compression ratio, where $N = 128$, $E_{flop} = 371$pJ [63], $E_{t_{elec}} = 50$nJ, $\epsilon_{amp} = 0.1$nJ/m^2 [55] and $d = 1000$m.

periodic, and its duty-cycled signals are suitable for DCT, DFT, and ADT, respectively. In summary, the CS based compression and authentication scheme can be applied to power reading signals of smart meters with an HWT representation matrix.

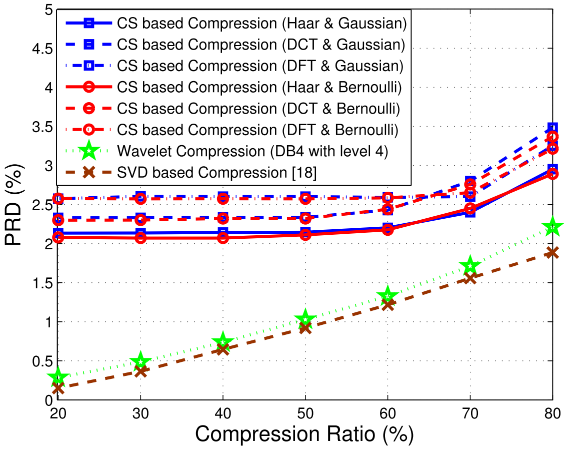


FIGURE 5. PRD of the reconstructed power signal with various compression ratio, where $N = 128$.

In Fig. 5, we show the PRD of the reconstructed power signal with various compression ratios when $N = 128$. Here, the compression ratio is given by CR $= \frac{B_{orig} - B_{comp}}{B_{orig}} \times 100$, where B_{orig} and B_{comp} are the number of bits for the original signal and compressed signal, respectively. In simulations, we compare the PRD of the CS based compression scheme with the existing compression schemes (the Daubechies wavelet (DB4) compression and SVD based compression [18]). To compare the compression performance of the proposed scheme with those of the existing methods, B_{orig} and B_{comp} in the compression ratio are obtained by calculating the entropies for the original and compressed power signals, respectively. The existing schemes have lower PRDs than the CS based compression method. Thus, it is important to find the best representation and measurement matrices in the CS based compression scheme that make the PRD low for AMI power signals. Based on the real-time power reading signal, an adaptive selection of the best representation and measurement matrices can be considered an optimal selection in practice. However, it is out of our scope in this paper (considered future work). Alternatively, in this study, based on empirically modeled signals, the statistically optimal representation matrix and measurement matrix are investigated to improve the reconstruction performance of the proposed scheme. From this figure, we can see that if a representation matrix and a measurement matrix are HWT and Gaussian (or Bernoulli) matrix, respectively, we can obtain the lowest PRD in the CS based compression, which is slightly higher than those of the existing compression schemes.

In Fig. 6, we show the simulation results for energy consumption for compression and transmission over the compression ratio, when $N = 128$, $E_{flop} = 371$pJ [63], $E_{t_{elec}} = 50$nJ, $\epsilon_{amp} = 0.1$nJ/m^2 [55] and $d = 1000$m. In simulations, four compression cases ('no compression', 'CS based compression', 'wavelet compression', 'SVD based compression [18]') are considered. The case of 'no compression' has no energy consumption for compression because it directly transmits power signals without compression. As shown in this figure, we can find that the case of 'CS based compression' can reduce energy consumption, compared to those of the other cases. For a low compression ratio (i.e., CR $< 20\%$), the case of 'wavelet compression' has poorer energy efficiency than that of 'no compression' because of energy consumption for compression, while the case of 'CS based compression' has a high energy efficiency regardless of the compression ratio. Thus, it is shown that the proposed CS based compression scheme is suitable for smart meters that have limited hardware resources (e.g., low computational capability) with large energy efficiency.

Fig. 7 depicts the probability density functions of the power of residual errors for legitimate and intrusion signals, in which $N = 128$, $L = 64$, $\kappa = 0.9985$, and SNR $= 10$dB.

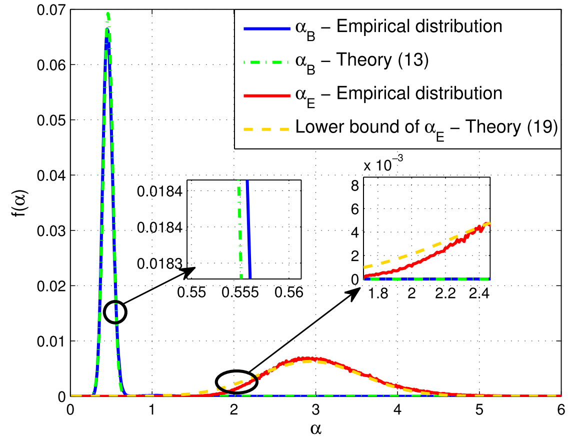
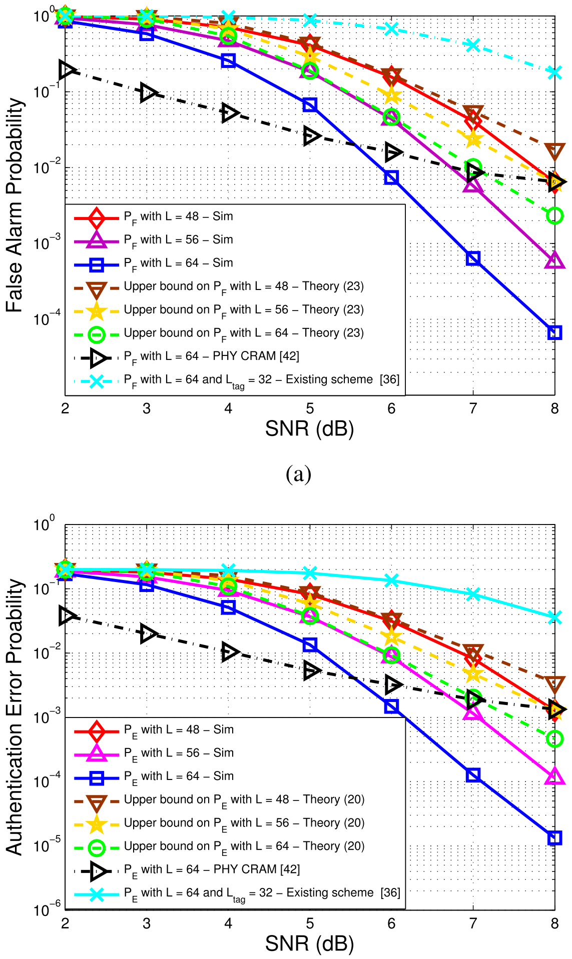


FIGURE 7. Comparison of the probability density functions between the LM and the IM, where $N = 128$, $L = 64$, $\kappa = 0.9985$ and SNR = 10dB.

In simulations, HWT is used for the representation matrix, $\boldsymbol{\Psi}$, because it is shown that HWT is the best representation matrix, as shown in Fig. 4. Thus, as shown in this figure, the empirical distribution of the legitimate signal closely agrees with the theoretical distribution of (13), while the mean of the theoretical distribution of α_B is very slightly lower than that of the empirical distribution, because we assume that $\boldsymbol{\Theta}_B\mathbf{s}_B = \hat{\boldsymbol{\Theta}}_B\hat{\mathbf{s}}_B$ in the theoretical analysis. In addition, the empirical distribution of the intrusion signal is close to the theoretical distribution of $\underline{\alpha}_E$, (19). We can see that the distribution of α_B is sufficiently distinguishable from that of α_E. Since the authentication matrix (i.e., key of the authentication) is unknown to the IM, the power of the residual error is relatively higher than that of the LM. This means that the difference between authentication matrices of the DCU and IM induces a sufficiently large residual error in the proposed scheme. Therefore, from this figure, we can find that if we set a proper threshold for the hypothesis testing in the proposed scheme, the signal from the IM can be detected with a high probability.

In Fig. 8, we show the impact of SNR and L on the performance, when $N = 128$, $P_D^\circ = 1 - 10^{-15}$, $\kappa = 0.9985$, and $\varrho = 0.8$. In general, the signal from the LM should be authenticated with a high probability for efficient power reporting in AMI. So, in simulations, the proposed scheme is evaluated with a high target detection probability ($P_D^\circ = 1 - 10^{-15}$), in terms of false alarm probability. Note that it is an unfavorable condition because, for a high P_D°, the false alarm probability becomes high due to the high η, determined by P_D°. Based on the different statistics of α_B and α_E in Fig. 7, the false alarm probability can be seen in Fig. 8(a) with different SNR and L. As mentioned earlier, as SNR increases, the false alarm probability dramatically decreases because a moderate SNR is necessary to guarantee $\boldsymbol{\Theta}_B\mathbf{s}_B = \hat{\boldsymbol{\Theta}}_B\hat{\mathbf{s}}_B$ in the LM, but it is not in the IM. In addition, as L increases, the correlation coefficient between measurement matrices of the DCU and IM, ν_1, becomes low, which makes the false alarm probability low. However, an existing CS based authentication method [36] that uses a tag signal for image authentication maintains a relatively high false alarm probability with $L_{tag} = 32$, which denotes the length of a tag signal due to short packet size. This means that it is not suitable for authentication of smart meter readings. Meanwhile, the proposed scheme has a lower false alarm probability for a high SNR (i.e., SNR $\geq$ 6dB) than an existing physical layer authentication method [42]. In a low SNR, we cannot obtain good authentication performance in the proposed scheme because an exact CS recovery cannot be guaranteed. In addition, while the existing scheme (for a physical layer challenge-response (PHY-CRAM) in multicarrier systems) needs to transmit the subsequent transmission of the power reading signal, the proposed scheme does not need to transmit a power reading signal and an authentication sequence separately. In Fig. 8(b), the authentication error probabilities are shown, in which a low authentication error probability can be obtained for a high SNR and a large L. In summary, although P_D° is high, the proposed scheme can guarantee good authentication performance (e.g., $P_E = 10^{-4}$) with a moderate SNR (e.g., 8dB) and L (e.g., 64).



(a)

(b)

FIGURE 8. Authentication performances with various SNRs, where $N = 128$, $P_D^\circ = 1 - 10^{-15}$, $\kappa = 0.9985$ and $\varrho = 0.8$; (a) a false alarm probability; (b) an authentication error probability.

VI. CONCLUSION

Lightweight authentication is a major concern for low-cost smart meters in AMI. In this paper, we proposed a unified approach for power reading signal compression and authentication in AMI by using the notion of CS. The proposed scheme, which enables simultaneous compression and authentication for smart meter readings can reduce the burden of computational complexity for low-cost smart meters in AMI. To this end, an aggregated power reading signal of smart meters is transformed into a sparse signal for CS based compression. Meanwhile, a power reading signal can be efficiently authenticated without additional signal processing by using a residual error, which is an output of decompression. Furthermore, we derived a theoretical threshold for hypothesis testing and theoretical analysis for an authentication error probability. From the analysis and simulation results, we showed that the power reading signal can be authenticated under reasonable conditions (e.g., $L = 64$ and SNR $= 8$dB), with a probability of $1 - P_E$, where $P_E \leq 10^{-4}$.

VII. CHALLENGES AND FUTURE WORK

There are several avenues in which to explore this research further. For example, we can consider the optimization of representation and measurement matrices to improve compression and authentication performance. In particular, it is necessary to a secure measurement matrix against intelligent attacks from an IM. In addition, there may be potential to enhance CS based authentication by using CSI between a DCU and an LM as not only another signature for authentication but also an obstacle to known plaintext attacks in which an IM tries to estimate a measurement matrix.

REFERENCES

[1] V. C. Gungor *et al.*, "Smart grid technologies: Communication technologies and standards," *IEEE Trans. Ind. Informat.*, vol. 7, no. 4, pp. 529–539, Nov. 2011.

[2] C. Selvam, K. Srinivas, G. S. Ayyappan, and M. V. Sarma, "Advanced metering infrastructure for smart grid applications," in *Proc. Int. Conf. Recent Trends Inf. Technol. (ICRTIT)*, Apr. 2012, pp. 145–150.

[3] J. Zhou, R. Q. Hu, and Y. Qian, "Scalable distributed communication architectures to support advanced metering infrastructure in smart grid," *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 9, pp. 1632–1642, Sep. 2012.

[4] A. Zoha, A. Gluhak, M. A. Imran, and S. Rajasegarar, "Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey," *Sensors*, vol. 12, no. 12, pp. 16838–16866, 2012.

[5] P. McDaniel and S. McLaughlin, "Security and privacy challenges in the smart grid," *IEEE Security Privacy*, vol. 7, no. 3, pp. 75–77, May/Jun. 2009.

[6] B. Krebs, "FBI: Smart meter hacks likely to spread," *Krebs Secur.*, 2012. [Online]. Available: http://krebsonsecurity.com/2012/04/fbi-smart-meter-hacks-likely-to-spread/

[7] J. Froehlich, E. Larson, S. Gupta, G. Cohn, M. Reynolds, and S. Patel, "Disaggregated end-use energy sensing for the smart grid," *IEEE Pervasive Comput.*, vol. 10, no. 1, pp. 28–39, Mar. 2011.

[8] M. Zeifman and K. Roth, "Nonintrusive appliance load monitoring: Review and outlook," *IEEE Trans. Consum. Electron.*, vol. 57, no. 1, pp. 76–84, Feb. 2011.

[9] M. Zeifman, "Disaggregation of home energy display data using probabilistic approach," *IEEE Trans. Consum. Electron.*, vol. 58, no. 1, pp. 23–31, Feb. 2012.

[10] Z. Guo, Z. J. Wang, and A. Kashani, "Home appliance load modeling from aggregated smart meter data," *IEEE Trans. Power Syst.*, vol. 30, no. 1, pp. 254–262, Jan. 2015.

[11] D. Setlhaolo, X. Xia, and J. Zhang, "Optimal scheduling of household appliances for demand response," *Electr. Power Syst. Res.*, vol. 116, pp. 24–28, Nov. 2014.

[12] Ö. N. Gerek and D. G. Ece, "Compression of power quality event data using 2D representation," *Electr. Power Syst. Res.*, vol. 78, no. 6, pp. 1047–1052, 2008.

[13] D. Zhang, Y. Bi, and J. Zhao, "A new data compression algorithm for power quality online monitoring," in *Proc. Int. Conf. Sustain. Power Gener. Supply (SUPERGEN)*, Apr. 2009, pp. 1–4.

[14] M. Ringwelski, C. Renner, A. Reinhardt, A. Weigel, and V. Turau, "The Hitchhiker's guide to choosing the compression algorithm for your smart meter data," in *Proc. IEEE Int. Energy Conf. Exhib. (ENERGYCON)*, Sep. 2012, pp. 935–940.

[15] A. Abuadbba, I. Khalil, and X. Yu, "Gaussian approximation-based lossless compression of smart meter readings," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 5047–5056, Sep. 2018.

[16] G. Da Poian, D. Brandalise, R. Bernardini, and R. Rinaldo, "Energy and quality evaluation for compressive sensing of fetal electrocardiogram signals," *Sensors*, vol. 17, no. 1, p. 9, 2016.

[17] M. P. Tcheou *et al.*, "The compression of electric signal waveforms for smart grids: State of the art and future trends," *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 291–302, Jan. 2014.

[18] J. C. S. de Souza, T. M. L. Assis, and B. C. Pal, "Data compression in smart distribution systems via singular value decomposition," *IEEE Trans. Smart Grid*, vol. 8, no. 1, pp. 275–284, Jan. 2017.

[19] Y. Lee, E. Hwang, and J. Choi, "Compressive sensing based power signal compression in advanced metering infrastructure," in *Proc. 23rd Asia–Pacific Conf. Commun. (APCC)*, Dec. 2017, pp. 1–6.

[20] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2012.

[21] Y. Rachlin and D. Baron, "The secrecy of compressed sensing measurements," in *Proc. 46th Annu. Allerton Conf. Commun., Control, Comput.*, Sep. 2008, pp. 813–817.

[22] A. Orsdemir, H. O. Altun, G. Sharma, and M. F. Bocko, "On the security and robustness of encryption via compressed sensing," in *Proc. IEEE Mil. Commun. Conf. (MILCOM)*, Nov. 2008, pp. 1–7.

[23] S. A. Hossein, A. Tabatabaei, and N. Zivic, "Security analysis of the joint encryption and compressed sensing," in *Proc. 20th Telecommun. Forum (TELFOR)*, Nov. 2012, pp. 799–802.

[24] M. R. Mayiami, B. Seyfe, and H. G. Bafghi, "Perfect secrecy via compressed sensing," in *Proc. Iran Workshop Commun. Inf. Theory*, 2013, pp. 1–5.

[25] T. Bianchi, V. Bioglio, and E. Magli, "Analysis of one-time random projections for privacy preserving compressed sensing," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 2, pp. 313–327, Feb. 2016.

[26] R. Dautov and G. R. Tsouri, "Establishing secure measurement matrix for compressed sensing using wireless physical layer security," in *Proc. Int. Conf. Comput., Netw. Commun. (ICNC)*, Jan. 2013, pp. 354–358.

[27] V. Cambareri, M. Mangia, F. Pareschi, R. Rovatti, and G. Setti, "On known-plaintext attacks to a compressed sensing-based encryption: A quantitative analysis," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 10, pp. 2182–2195, Oct. 2015.

[28] P. Lu, Z. Xu, X. Lu, and X. Liu, "Digital image information encryption based on compressive sensing and double random-phase encoding technique," *Optik-Int. J. Light Electron Opt.*, vol. 124, no. 16, pp. 2514–2518, 2013.

[29] J. Lang and J. Zhang, "Optical image cryptosystem using chaotic phase-amplitude masks encoding and least-data-driven decryption by compressive sensing," *Opt. Commun.*, vol. 338, pp. 45–53, Mar. 2015.

[30] N. Zhou, S. Pan, S. Cheng, and Z. Zhou, "Image compression–encryption scheme based on hyper-chaotic system and 2D compressive sensing," *Opt. Laser Technol.*, vol. 82, pp. 121–133, Aug. 2016.

[31] Y. Zhang *et al.*, "A block compressive sensing based scalable encryption framework for protecting significant image regions," *Int. J. Bifurcation Chaos*, vol. 26, no. 11, 2016, Art. no. 1650191.

[32] S.-Y. Chiu, H. H. Nguyen, R. Tan, D. K. Yau, and D. Jung, "JICE: Joint data compression and encryption for wireless energy auditing networks," in *Proc. 12th Annu. IEEE Int. Conf. Sens., Commun., Netw. (SECON)*, Jun. 2015, pp. 453–461.

[33] L.-B. Zhang, Z.-L. Zhu, B.-Q. Yang, W.-Y. Liu, H.-F. Zhu, and M.-Y. Zou, "Medical image encryption and compression scheme using compressive sensing and pixel swapping based permutation approach," *Math. Problems Eng.*, vol. 2015, Jul. 2015, Art. no. 940638.

[34] C. Wang, B. Zhang, K. Ren, J. M. Roveda, C. W. Chen, and Z. Xu, "A privacy-aware cloud-assisted healthcare monitoring system via compressive sensing," in *Proc. IEEE INFOCOM*, Apr. 2014, pp. 2130–2138.

[35] I. Orović and S. Stanković, "Combined compressive sampling and image watermarking," in *Proc. 55th Int. Symp. ELMAR*, Sep. 2013, pp. 41–44.

[36] T. Wu and C. Ruland, "An improved authenticated compressive sensing imaging," in *Proc. IEEE 12th Int. Conf. Semantic Comput. (ICSC)*, Jan. 2018, pp. 164–171.

[37] Y. Zhang, L. Y. Zhang, J. Zhou, L. Liu, F. Chen, and X. He, "A review of compressive sensing in information security field," *IEEE Access*, vol. 4, pp. 2507–2519, 2016.

[38] M. M. Fouda, Z. M. Fadlullah, N. Kato, R. Lu, and X. Shen, "A lightweight message authentication scheme for smart grid communications," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 675–685, Dec. 2011.

[39] H. Nicanfar, P. Jokar, and V. C. M. Leung, "Smart grid authentication and key management for unicast and multicast communications," in *Proc. IEEE PES Innov. Smart Grid Technol.*, Nov. 2011, pp. 1–8.

[40] H. Li, R. Lu, L. Zhou, B. Yang, and X. Shen, "An efficient Merkle-tree-based authentication scheme for smart grid," *IEEE Syst. J.*, vol. 8, no. 2, pp. 655–663, Jun. 2014.

[41] L. Xiao, L. J. Greenstein, N. B. Mandayam, and W. Trappe, "Using the physical layer for wireless authentication in time-variant channels," *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2571–2579, Jul. 2008.

[42] X. Wu and Z. Yang, "Physical-layer authentication for multi-carrier transmission," *IEEE Commun. Lett.*, vol. 19, no. 1, pp. 74–77, Jan. 2015.

[43] J. Choi, "A coding approach with key-channel randomization for physical-layer authentication," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 1, pp. 175–185, Jan. 2019.

[44] W. Hou, X. Wang, J.-Y. Chouinard, and A. Refaey, "Physical layer authentication for mobile systems with time-varying carrier frequency offsets," *IEEE Trans. Commun.*, vol. 62, no. 5, pp. 1658–1667, May 2014.

[45] P. L. Yu, J. S. Baras, and B. M. Sadler, "Physical-layer authentication," *IEEE Trans. Inf. Forensics Security*, vol. 3, no. 1, pp. 38–51, Mar. 2008.

[46] S. Barker, S. Kalra, D. Irwin, and P. Shenoy, "Empirical characterization, modeling, and analysis of smart meter data," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 7, pp. 1312–1327, Jul. 2014.

[47] F. Sullivan, "Could smart meters eventually power themselves?" *Smart Energy Int.*, no. 3, 2016. [Online]. Available: https://www.metering.com/magazine-article/smart-meters-eventually-power/

[48] X. Li, Q. Huang, and D. Wu, "Distributed large-scale co-simulation for IoT-aided smart grid control," *IEEE Access*, vol. 5, pp. 19951–19960, 2017.

[49] J. Harrison, "Choosing an MCU for smart energy meters," *Electron. Products*, 2012. [Online]. Available: https://www.digikey.com/en/articles/techzone/2012/jul/choosing-an-mcu-for-smart-energy-meters

[50] N. Ahmed, T. Natarajan, and K. R. Rao, "Discrete cosine transform," *IEEE Trans. Comput.*, vol. 100, no. 1, pp. 90–93, Jan. 1974.

[51] J. Ning, J. Wang, W. Gao, and C. Liu, "A wavelet-based data compression technique for smart grid," *IEEE Trans. Smart Grid*, vol. 2, no. 1, pp. 212–218, Mar. 2011.

[52] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[53] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.

[54] L. F. Cisneros-Sinencio, A. Diaz-Sanchez, and J. Ramirez-Angulo, "FGMOS flip-flop for low-power signal processing," *Int. J. Electron.*, vol. 100, no. 12, pp. 1683–1689, 2013.

[55] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proc. 33rd Annu. Hawaii Int. Conf. Syst. Sci.*, Jan. 2000, pp. 1–10.

[56] H. Li, X. Wang, and J. Y. Chouinard, "Eavesdropping-resilient OFDM system using sorted subcarrier interleaving," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 1155–1165, Feb. 2015.

[57] P. Sure and C. M. Bhuma, "A survey on OFDM channel estimation techniques based on denoising strategies," *Eng. Sci. Technol., Int. J.*, vol. 20, no. 2, pp. 629–636, 2017.

[58] J. Choi, "Secure transmissions via compressive sensing in multicarrier systems," *IEEE Signal Process. Lett.*, vol. 23, no. 10, pp. 1315–1319, Oct. 2016.

[59] C. Condo and W. J. Gross, "Pseudo-random Gaussian distribution through optimised LFSR permutations," *Electron. Lett.*, vol. 51, no. 25, pp. 2098–2100, 2015.

[60] U. M. Maurer, "Authentication theory and hypothesis testing," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1350–1356, Jul. 2000.

[61] G. Dziwoki, "Averaged properties of the residual error in sparse signal reconstruction," *IEEE Signal Process. Lett.*, vol. 23, no. 9, pp. 1170–1173, Sep. 2016.

[62] G. Strang, *Introduction to Linear Algebra*. Cambridge, MA, USA: Wellesley-Cambridge Press, 2003.

[63] J. W. Choi, D. Bedard, R. Fowler, and R. Vuduc, "A roofline model of energy," in *Proc. IEEE 27th Int. Symp. Parallel Distrib. Process. (IPDPS)*, May 2013, pp. 661–672.

YONGGU LEE received the B.E. degree in electronics engineering from Kyungpook National University, Daegu, South Korea, in 2013, and the M.S. and Ph.D. degrees in electrical engineering and computer science from the Gwangju Institute of Science and Technology (GIST), Gwangju, South Korea, in 2016 and 2019, respectively, where he is currently a Senior Researcher with the Research Institute for Solar and Sustainable Energies. His research interests include the areas of communications and signal processing, with emphasis on physical layer security, 5G communication systems, and machine type communications.

EUISEOK HWANG received the B.S. and M.S. degrees from the School of Engineering, Seoul National University, Seoul, South Korea, in 1998 and 2000, respectively, and the M.S. and Ph.D. degrees in electrical and computer engineering from Carnegie Mellon University (CMU), Pittsburgh, PA, USA, in 2010 and 2011, respectively. He was with the Digital Media Research Center, Daewoo Electronics Co., Ltd., South Korea, from 2000 to 2006, and was with the Channel Architecture Group, LSI Corporation (now Broadcom), San Jose, CA, USA, from 2011 to 2014. Since 2015, he has been an Assistant Professor with the School of Mechatronics, Gwangju Institute of Science and Technology (GIST), South Korea. He has over 70 journal and conference papers on various data channel and signal processing issues, and holds 21 granted U.S. patents. His research interests include signal disaggregation, coding, and read channel for data storage and communication systems, and emerging large-scale information processing applications, such as smart grids.

JINHO CHOI (SM'02) was born in Seoul, South Korea. He received the B.E. degree *(magna cum laude)* in electronics engineering from Sogang University, Seoul, in 1989, and the M.S.E. and Ph.D. degrees in electrical engineering from the Korea Advanced Institute of Science and Technology (KAIST), in 1991 and 1994, respectively. He was with Swansea University, U.K., as a Professor/Chair in wireless, and was with the Gwangju Institute of Science and Technology (GIST), South Korea, as a Professor, in 2018. He is currently a Professor with the School of Information Technology, Deakin University, Burwood, Australia. He has authored two books published by Cambridge University Press, in 2006 and 2010, respectively. His research interests include the Internet of Things (IoT), wireless communications, and statistical signal processing. He received the 1999 Best Paper Award for Signal Processing from EURASIP and the 2009 Best Paper Award from WPMC (Conference). He served as an Associate Editor or an Editor for other journals, including the IEEE Communications Letters, the *Journal of Communications and Networks* (JCN), the IEEE Transctions on Vehicular Technology, and the *ETRI Journal*. He is currently the Editor of the IEEE Transactions on Communications and the IEEE Wireless Communications Letters.