Diverse Grouping-Based Aggregation Protocol With Error Detection for Smart Grid Communications

Zhiguo Shi, *Member, IEEE*, Ruixue Sun, Rongxing Lu, *Member, IEEE*, Le Chen, Jiming Chen, *Senior Member, IEEE*, and Xuemin (Sherman) Shen, *Fellow, IEEE*

***Abstract*—Smart grid, as the next generation of power grid characterized by "two-way" communications, has been paid great attention to realizing green, reliable, and efficient electricity delivery for our future lives. In order to support the two-way communications in smart grid, a large number of smart meters (SMs) should be deployed to customers to report their near real-time data to control center for monitoring purpose. However, this kind of real-time report could disclose users' privacy, bringing down the users' willingness to participate in smart grid. In order to address the challenge, in this paper, by considering the lifetime of SMs as exponential distribution, we propose a diverse grouping-based aggregation protocol with error detection (DG-APED), which employs differential privacy technique into grouping-based private stream aggregation for secure smart grid communications. DG-APED can not only achieve privacy-preserving aggregation, but also perform error detection efficiently when some SMs are malfunctioning. Detailed security analysis shows that DG-APED can guarantee the security and privacy requirements of smart grid communications. In addition, extensive performance evaluation also verifies the effectiveness and efficiency of the proposed DG-APED.**

***Index Terms*—Differential privacy, error detection, exponential distribution, privacy-preserving aggregation, smart grid.**

NOMENCLATURE

N	Number of residential users.
$x_i(t)$	Residential user U_i's individual data.
r_i	Random noise with unbiased binary distribution generated by user U_i.
sk_i	Private key of user U_i used to encrypt $x_i(t)$.
$sk_{u0}^{v,l}$	Private key to aggregate the data of residential users in $\mathbb{P}_v^l$, the vth group set in smart meter (SM) type SM_l.
sk_{l0}^{v}	Private key to aggregate the data of residential users in SM_l of the vth group set.
sk_0	Private key to aggregate the data of all residential users.
$\rho_0(\Delta t)$	Probability that a SM is functioning.
$\rho_1(\Delta t)$	Probability that a SM is malfunctioning.
z_l	Size of residential users group in SM type SM_l.
w_l	Rounds of random residential users grouping in SM type SM_l.
m_l	Number of malfunctioning SMs in SM type SM_l.

Manuscript received June 7, 2014; revised November 5, 2014 and March 15, 2015; accepted May 30, 2015. Date of publication July 28, 2015; date of current version October 17, 2015. This work was supported by the National Natural Science Foundation of China under Grant 61171149. Paper no. TSG-00545-2014.

Z. Shi was with the University of Waterloo, Waterloo, ON N2L 3G1, Canada. He is now with the Faculty of Information Technology, Zhejiang University, Hangzhou 310027, China (e-mail: shizg@zju.edu.cn).

R. Sun and J. Chen are with the Faculty of Information Technology, Zhejiang University, Hangzhou 310027, China (e-mail: sunrx@zju.edu.cn; jmchen@zju.edu.cn).

R. Lu and L. Chen are with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798.

X. Shen is with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSG.2015.2443011

I. INTRODUCTION

RECENTLY, there have been increasing demands for improving green, reliable, and efficient power transmission. As the next generation of power grid, smart grid is characterized by integrating various technologies, e.g., advanced control and sensing, data communication, and information processing technologies, into the traditional power grid. This characteristic enables the power distribution to be more efficient and reliable from power generation, transmission, and distribution to customers, and supports the renewable energy [1]–[3]. Due to these merits, smart grid has been paid great attention from not only government but also industry and academia. For example, as shown in Fig. 1, a conceptual smart grid model was defined by National Institute of Standards and Technology (NIST) in 2010, where the smart grid is divided into seven domains, and each domain encompasses one or more smart grid actors, including devices, systems, or programs that make decisions, and exchange information necessary to perform typical applications [4].

One of the most vital components in smart grid is SM, which could record the consumption of electric energy in intervals of an hour or less and communicate the information at least daily back to the operation center for monitoring purposes. Such an advanced SM differs from traditional automatic meters in that, it enables two-way communications between the meter and operation center. Owing to this characteristic, the SM is promoted around the world. According to the analyst firm Berg Insight, the installed base of SMs in Europe at the end of 2008 was about 39 million units. Globally, Pike Research found that SM shipments were 17.4 million units for the first quarter of 2011, and the value of the global SM market reached 7 billion in 2012 [5]. By equipping with plenty of SMs, the smart grid is able to aggregate

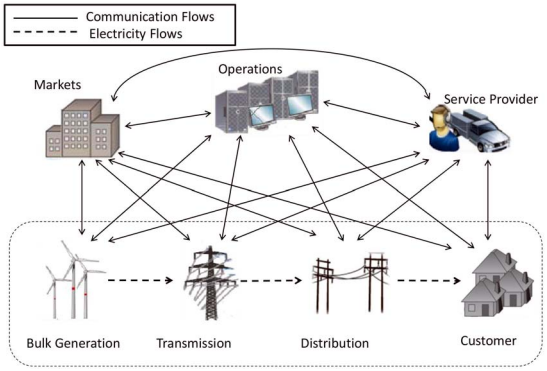


Fig. 1. Conceptual smart grid model from NIST.

real-time information related to energy consumption, thus performs intelligent balancing of consumption between peak and off-peak periods [6]–[9].

However, the widespread deployment of SMs makes customer privacy a crucial issue during the aggregation process. A typical concern is that the data of energy consumption stored at SMs acts as an information-rich side channel, and can be repurposed by an adversary to reveal customer information such as individuals habits, behaviors, activities, preferences, and even beliefs [10], [11]. As a result, these potential threats targeting user privacy would, without doubt, prevent smart grid from being widely accepted by customers. To combat this, many research works on privacy-preserving aggregation have been carried out for smart grid communications. However, some of these existing works are computational expensive or need interaction among the participants, while others will fail to aggregate related data when there are malfunctioning SMs. Therefore, how to design an efficient, fault tolerant, and error detectable privacy-preserving aggregation protocol for practical smart grid scenarios is very challenging.

To address this problem, in this paper, we propose an efficient diverse grouping-based aggregation protocol with error detection (DG-APED), which employs differential privacy technique into grouping-based private stream aggregation to not only achieve privacy-preserving aggregation but also perform error detection when some SMs are malfunctioning. The main contribution of this paper are summarized as follows.

1) By formulating the lifetime of SMs as exponential distribution, we classify all the SMs into different types accordingly, then the control center (CC) can aggregate related data and detect the malfunctioning SMs with homologous aggregation parameters efficiently in terms of computational complexity.
2) Considering the malicious data mining attack, we propose a novel aggregation protocol that utilizes the differential privacy to enable the CC to get the aggregation results without obtaining the individual value.
3) We analyze the security of the proposed DG-APED to formally prove that the privacy of users can be guaranteed against different attacks. Moreover, we evaluate the performance of DG-APED and show that it is effective for smart grid data aggregation and efficient in terms of computational complexity and communication overhead.

The remainder of this paper is organized as follows. We review the related work and introduce the models and design goals in Sections II and III, respectively. In Section IV, we describe our basic grouping-based aggregation protocol G-PSA. In Section V, we propose the diverse grouping-based aggregation protocol DG-APED, followed by its security analysis and performance evaluation in Sections VI and VII, respectively. Finally, we conclude this paper in Section VIII.

II. Related Work

There have appeared several research works that discuss secure aggregation techniques [12]–[25], and these techniques can be roughly classified into: 1) aggregation using plain data and 2) aggregation using encrypted data with the requirement of decrypting data at aggregators.

The earlier works [12]–[14] are focused on aggregation using plain data. Hu and Evans [12] proposed security protocol based on symmetric cryptography to achieve secure aggregation hop-by-hop and study the problem of aggregation if one meter is compromised. The concept of aggregators is proposed by Przydatek *et al*. [13] to prevent that more than one meter is compromised, and also propose security protocol in which random sampling mechanism and interactive proof are used to check the correctness of the aggregated data at the CC. SecureDAV protocol in [14] provides confidentiality by elliptic curve cryptography and improves the data integrity by signing the aggregated data. In these works, security aggregation is achieved in a hop-by-hop fashion, that is, aggregators must decrypt all the received information for aggregation and encrypt the aggregated result before forwarding it. In addition, to ensure end-to-end confidentiality, the aggregators have to establish secret keys with their neighboring users. Clearly, this is not an efficient way and it may result in considerable delay.

To mitigate the drawbacks of aggregation using plain data, a set of aggregation protocols using encrypted data are proposed in [15]–[25]. The homomorphic encryption technique, which allows one to implement homomorphic operation on ciphertexts under the same key without being able to decrypt, has been adopted widely for privacy-preserving aggregation in [15]–[18]. However, since the homomorphic encryption technique requires complex public key operations, these protocols are not efficient in smart grid scenario. The concept of privacy-preserving aggregation is introduced for smart grid in [19], in which homomorphic properties of Paillier encryption is exploited on additive sharing of each meter's reading. Kursawe *et al.* [20] proposed a protocol based on Diffie–Hellman key exchange to add secret value on each meter's readings for each round such that they add up to zero. References [19] and [20] require interaction between the individual SMs, resulting in relatively expensive cryptography on the SMs. Another privacy-preserving technique is the differential privacy, which is a tradeoff scheme between the utility of statistical data and the privacy of users [21], [26]. Intuitively, differential privacy is achieved by introducing appropriate noise which obeys certain random distribution to guarantee that, the presence or absence of an individual user's value

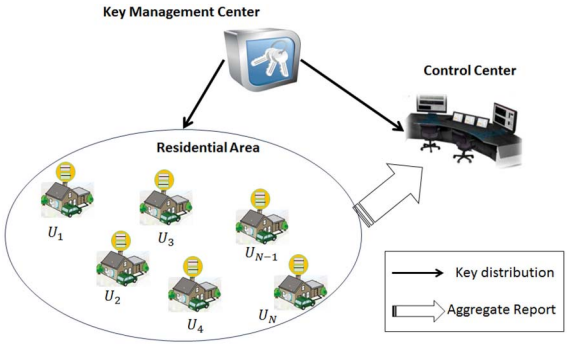


Fig. 2. System model in smart grid.

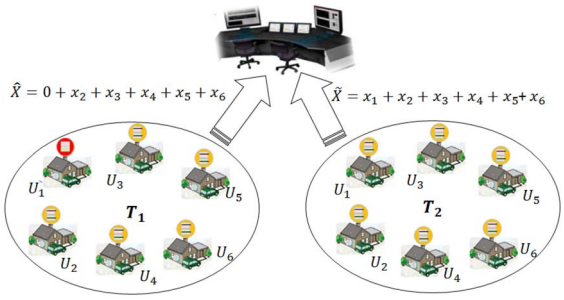


Fig. 3. Maliciously mining attack example.

in the database, should not affect the released result in a significant way.

The closely related works to this paper are [23] and [24]. Shi *et al.* [23] proposed the private stream aggregation (PSA) which combines data hiding from aggregator with differentially private aggregate computation. While this is an excellent attempt to achieve privacy-preserving aggregation, it is impractical in smart grid scenario because of user failure problem. To solve this problem, Chan *et al.* [24] employed the block aggregation (BA) which is binary tree framework on top of the PSA, to improve its performance, enabling the aggregation protocol with fault tolerance in expense of efficiency. Considering the error detection ability, we propose a novel aggregation protocol with error detection (APED) which utilizes the random meters pairing mechanism to achieve privacy-preserving aggregation with fault tolerance and error detection, in the conference version of this paper [25]. However, the efficiency of BA and APED is still an issue in term of computation cost. To combat this, and further consider malicious data mining attack, we propose an efficient privacy-preserving aggregation protocol with error detection, namely DG-APED, in this paper.

III. Models and Design Goals

In this section, we formalize the system model, attack model, and identify our design goals.

A. System Model

In our system model, we only focus on the data aggregation communications from the residential users to the CC in smart grid. Specifically, we consider a typical data aggregation model for a residential area, which includes a set of residential users $\mathbb{U} = \{U_1, U_2, \ldots, U_N\}$, a CC, and a key management center (KMC) [27], as shown in Fig. 2.

1) KMC: The KMC is a highly trusted entity, which is responsible for classifying the users and distributing key materials to residential users and the CC during the system initialization phase. After the key distribution is completed, the KMC will not involve in the subsequent aggregation.

2) CC: The CC that plays the role of aggregator is an untrusted entity, whose duty is to monitor the state of the smart grid in a real-time manner and aggregate the periodical electricity usage data from residential users to distribute the electricity efficiently.

3) Residential Users $\mathbb{U} = \{U_1, U_2, \ldots, U_N\}$: Each user $U_i \in \mathbb{U}$ is equipped with some smart appliances and a SM which periodically collects electricity usage data from smart appliances and reports these near real-time data to the CC. And each SMs reading is supposed in the range $\{0, 1, \ldots, \Delta\}$, the upper bound Δ is the power consumption in a short time, e.g., 15 mins or half an hour. According to [28], the SMs' lifetime is formulated as exponential distribution in this paper. Moreover, when a SM behaves malfunctioning sometimes, it requires the CC to remotely reset and recover it back to the normal status.

B. Attack Model

In the smart grid, the three most frequently launched attacks are: 1) external attack, where the adversary may compromise the privacy of residential users by eavesdropping the communication data from the residential users to the CC; 2) internal attack, where the attacker is usually the aggregator and could access or misuse the information of residential users to compromise the privacy of residential users; and 3) malicious data mining attack, where the attacker could maliciously mine the aggregation result by compromising some SMs to infer the residential users' data.

The external attack can be resisted by effective cryptographic algorithms. The aggregator in our model, i.e., the CC, is not a fully trusted entity, could misuse the aggregated information to compromise the privacy of users. However, this internal attack can be resisted by our proposed scheme, which will be described in detail in Section VI.

In the malicious data mining attack [29], the adversary may collude with a set of corrupted SMs to reveal the functioning SMs' data. Fig. 3 shows such an example. We assume that the electricity consumption aggregation is implemented on six users $(U_1, U_2, U_3, U_4, U_5, U_6)$ at two time slots T_1 and T_2. If the user U_1 is absent in T_1 but at home in T_2, the other users are compromised by the attacker and their electricity consumption remain relatively stable, then the attacker can mine the data of user U_1 by comparing two aggregation results, $\widetilde{X}$, $\widehat{X}$ in time slots T_2 and T_1, respectively, i.e., $x_1 = \widetilde{X} - \widehat{X}$.

C. Design Goals

Considering the above scenario, our design goal is to propose a privacy-preserving APED ability to identify these

malfunctioning SMs. Specifically, the following design goals should be achieved.

1) *Privacy Preservation:* A residential user's data is unaccessible to any other users or the CC, even when the user's SM is malfunctioning.
2) *Error-Detection:* The malfunctioning SMs in smart grid can be detected by the CC without privacy disclosure.
3) *Fault Tolerance:* The CC can still aggregate the data of functioning meters even when there are malfunctioning ones.
4) *Computation Efficiency:* The computation efficiency should be achieved in the proposed protocol to support hundreds and thousands residential users' data aggregation.

IV. G-PSA: Basic Privacy-Preserving Aggregation Protocol

In designing the aggregation protocol for smart grid, how to minimize the computational complexity and communication overhead while guaranteeing the strong privacy even when there are malfunctioning SMs is a major challenge that we have to face. Inspired by the PSA [23] which is one of the most promising aggregation protocols, we first propose a basic grouping-based privacy-preserving aggregation protocol, called G-PSA, which has fault tolerance ability and is more practical than PSA. As the basis of our proposed G-PSA, we first review the PSA before indulging in the G-PSA.

A. Private Stream Aggregation

The PSA flexibly leverages an aggregation encryption scheme that can allow the CC to decrypt the sum of all users' encrypted values, but learn nothing more. The PSA consists of the following steps [23].

1) Setup: During the system setup phase, the trusted KMC chooses a random generator g and $N+1$ private key sk_i for all devices including the users' SMs and the CC in the residential area. All the keys comply with the relation: $\sum_{i=0}^{N} sk_i = 0 \bmod p$.
2) NoisyEnc: After completing the Setup step, each user U_i calls NoisyEnc to randomize and encrypt its data x_i with noise r_i and private key sk_i : $c_i = g^{\widehat{x}_i} \cdot H(t)^{sk_i}$, where $\widehat{x}_i = x_i + r_i \bmod p$ and $H(t) : \{0,1\}^* \rightarrow \mathbb{G}$ is a cryptographic hash function which maps an integer $t \in \{0,1\}^*$ to a cyclic mathematical group $\mathbb{G}$ of prime order p.
3) AggrDec: After receiving all users' encrypted data c_i, $i = 1, 2, \ldots, N$, since $\prod_{i=0}^{N} H(t)^{sk_i} = 1$, the CC can compute an aggregate value $X(t) = H(t)^{sk_0} \prod_{i=1}^{N} c_i = g^{\sum_{i=1}^{N} \widehat{x}_i}$.

Since the aggregated sum is a small value in smart grid, the CC can obtain the aggregated sum $\sum_{i=1}^{n} \widehat{x}_i$ by the brute-force search. However, the PSA scheme will fail to aggregate in present of malfunctioning SM. To address this challenge, we propose our grouping-based PSA, G-PSA.

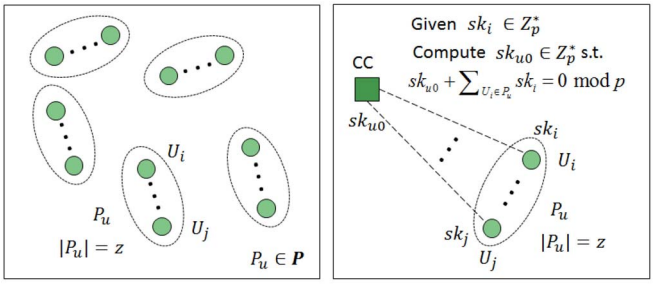


Fig. 4. KMC randomly divides N users into N/z groups and assigns key materials. (a) Random grouping $\mathbb{P}$. (b) Generate grouping keys.

B. Description of G-PSA

The G-PSA consists of three steps: 1) Setup; 2) Encrypt; and 3) Decrypt, and the details are described as follows.

1) Setup*(N,k):* Given the security parameter k, the KMC first outputs a cyclic group $\mathbb{G}$ with a generator g of the large prime order p where $|p| = k$. Let N be the number of residential users (N is a multiple of the size of residential users group, if not, we can easily add some dummy users to follow the rule). The KMC can randomly divide all the users $U_1, \ldots, U_N$ into small size groups, and denote them as group set $\mathbb{P}$. Each group $P_u \in \mathbb{P}, 1 \leq u \leq N/z$, includes z users $\{U_i\}_{U_i \in P_u}$, as shown in Fig. 4(a).

For each residential user $U_i \in \mathbb{U}$, the KMC assigns it a private key $sk_i \in \mathbb{Z}_p^*$. Then, for each group $P_u \in \mathbb{P}$, the KMC distributes a corresponding key $sk_{u0} \in \mathbb{Z}_p^*$ to the CC such that $\sum_{U_i \in P_u} sk_i + sk_{u0} = 0 \bmod p$, as shown in Fig. 4(b). In this way, the CC has N/z private keys corresponding to N/z groups. In addition, $sk_0 = \sum_{P_u \in \mathbb{P}} sk_{u0} \bmod p$ has to be assigned to the CC, then we can easily have the relationship $sk_0 + \sum_{i=1}^{N} sk_i = 0 \bmod p$.

2) Enc*(sk_i, $x_i(t)$):* At every time t, each residential user U_i calculates the ciphertext $c_i(t)$ on his/her individual data $x_i(t)$ with his/her private key sk_i as

$$c_i(t) = g^{x_i(t) \mod p} \cdot H(t)^{sk_i}. \tag{1}$$

3) Dec*(sk_d,{$c_i(t)$}):* Upon receiving ciphertexts $\{c_i(t)\}$, the CC decrypts the sum of corresponding residential users' data with the corresponding key sk_d.

When the CC tries to decrypt the sum of all residential users' data, we plug the corresponding key sk_0 and ciphertexts $\{c_1(t), c_2(t), \ldots, c_N(t)\}$ into decryption process which can be wrote as Dec$(sk_0, \{c_1(t), c_2(t), \ldots, c_N(t)\})$, then the CC can decrypt the sum of all users' data $\sum_{i=1}^{N} x_i(t)$ as

$$D(t) = \prod_{i=1}^{N} c_i(t) \cdot H(t)^{sk_0} = g^{\sum_{i=1}^{N} x_i(t)}. \tag{2}$$

Since the sum $\sum_{i=1}^{N} x_i(t)$ is a small value in smart grid scenario, it can be retrieved by the brute-force search.

When the CC tries to decrypt the sum of all residential users' data in a group P_u, we plug the corresponding key sk_{u0} and ciphertexts $\{c_i(t)\}_{U_i \in P_u}$ into decryption process which can be written as Dec$(sk_{u0}, \{c_i(t)\}_{U_i \in P_u})$, then the CC can decrypt

the sum of each group $\sum_{U_i \in P_u} x_i$ as

$$d_u(t) = \prod_{U_i \in P_u} c_i(t) \cdot H(t)^{sk_{u0}} = g^{\sum_{U_i \in P_u} x_i}. \quad (3)$$

Similarly, the corresponding sum $\sum_{U_i \in P_u} x_i$ for the user group P_u can be retrieved by the brute-force search.

C. Correctness Analysis

The CC can correctly get the sum of the SMs' data if all SMs are functioning.

For the decryption of all users, since $\sum_{i=0}^{N} sk_i = 0$

$$\begin{aligned} D(t) &= \prod_{i=1}^{N} c_i(t) \cdot H(t)^{sk_0} \\ &= g^{x_1(t)} H(t)^{sk_1} \cdot g^{x_2(t)} H(t)^{sk_2} \cdots g^{x_N(t)} H(t)^{sk_N} \cdot H(t)^{sk_0} \\ &= g^{x_1(t)+x_2(t)+\cdots+x_N(t)} \cdot H(t)^{sk_1+sk_2+\cdots+sk_N+sk_0} \\ &= g^{\sum_{i=1}^{N} x_i(t)}. \end{aligned}$$

Similarly, for the decryption of users in a group P_u, since $\sum_{U_i \in P_u} sk_i + sk_{u0} = 0$

$$\begin{aligned} d_u(t) &= \prod_{U_i \in P_u} c_i(t) \cdot H(t)^{sk_{u0}} \\ &= \prod_{U_i \in P_u} g^{x_i(t)} H(t)^{sk_i} \cdot H(t)^{sk_{u0}} \\ &= g^{\sum_{U_i \in P_u} x_i(t)} \cdot H(t)^{\sum_{U_i \in P_u} sk_i + sk_{u0}} \\ &= g^{\sum_{U_i \in P_u} x_i(t)}. \end{aligned}$$

When there are malfunctioning SMs, the CC will not consider the groups which include at least malfunctioning SMs, and only aggregate the sum of functioning groups by summing up their values

$$\sum_{i=1}^{N} x_i(t) \approx \sum_{P_u \in \mathbb{P}} \sum_{U_i \in P_u} x_i(t). \quad (4)$$

Note that, the aggregated result $\sum_{i=1}^{N} x_i(t)$ in (4) neglects the small data of functioning SMs in malfunctioning groups.

The G-PSA ensures that the CC learns nothing about the individual user's data but the overall aggregation result even when there are malfunctioning SMs. However, it is not efficient since malfunctioning SMs could be grouped with many functioning SMs with high probability, and is vulnerable to malicious data mining attack as depicted in Fig. 3. To improve the performance and the security of aggregation protocol, we further propose an advanced privacy-preserving APED based on diverse grouping mechanism, called DG-APED in the following section.

V. DG-APED: Advanced Privacy-Preserving Aggregation Protocol

The DG-APED, which takes the SMs' lifetime as exponential distribution and combines differential privacy technique with the G-PSA, is a more efficient aggregation protocol in smart grid scenario. In this section, we review the building blocks, namely the notion of exponential distribution and differential privacy firstly, then describe the intuition and the detailed design of DG-APED.

A. Preliminaries

1) Exponential Distribution: From the statistics point of view, the exponential distribution is usually adopted to describe the lifetime of electronic products and system [28]. Therefore, we formulate the distribution of SMs lifetime as exponential distribution.

Definition 1 (Exponential Distribution): For the exponential distribution with expectation μ, which can be denoted as $\exp(\mu)$, the probability density function is $f(x) = (1/\mu)e^{-(1/\mu)x}$, $x \geq 0$, and the corresponding cumulative distribution function is $F(x) = 1 - e^{-(1/\mu)\cdot x}$, $x \geq 0$.

The probability that a SM is functioning at time t can be denoted as

$$\rho_0(t, t_0) = P(T > t - t_0) = 1 - F(t - t_0) = e^{-\frac{1}{\mu}\cdot(t-t_0)} \quad (5)$$

in which t_0 is the initial time when the SM starts up. Thus, the probability that a SM is malfunctioning at time t can be denoted as

$$\rho_1(t, t_0) = 1 - \rho_0(t - t_0) = 1 - e^{-\frac{1}{\mu}\cdot(t-t_0)}. \quad (6)$$

In our protocol, $\rho_0(t, t_0)$ and $\rho_1(t, t_0)$, which can be written as $\rho_0(\Delta t_i)$ and $\rho_1(\Delta t_i)$, respectively, for U_i with $\Delta t_i = t - t_0$, are known to the KMC, and the KMC will classify all the users based on their probabilities in order to improve the efficiency of aggregation. Note that the probability depends on the operation time Δt. That is, the SMs can be classified into types $\mathrm{SM}_1, \ldots, \mathrm{SM}_L$ based on their operation time Δt.

2) Differential Privacy: The differential privacy is proposed by Dwork [21] to ensure that the users participating aggregation do not suffer increased risk of privacy disclosing. In general, the differential privacy is achieved by adding a random noise to perturb the outputs that will be indistinguishable for similar inputs.

Definition 2 [(ϵ, δ)-Differential Privacy]: For sets D_1 and D_2 differing on at most one element, and all $S \subseteq \mathrm{Range}(K)$, there is a randomized function K that can give δ-approximate ϵ-indistinguishability differential privacy

$$Pr[K(D_1) \in S] \leq \exp(\epsilon) \cdot Pr[K(D_2) \in S] + \delta.$$

To achieve differential privacy, the noise is binomial distribution in this paper, which can be regarded as an approximation for the gamma distribution [22].

Definition 3 (Unbiased Binomial Distribution): Binomial distribution, also called n-Bernoulli experiment, is usually denoted as $B(n, p)$. When $p = 1/2$, the distribution $B(n, 1/2)$ is an unbiased binomial distribution. The mass function for the binomial distribution $B(n, p)$ is $\binom{n}{k}(p)^k(1-p)^{n-k}$. When $p = 1/2$, the mass function at k is $\binom{n}{k}(1/2)^n$ and the expectation is $n/2$.

Fact 1: Let $X_1, \ldots, X_m$ be independent identically distributed, and $X_i \sim B(n_i, p)$, $i = 1, 2, \ldots, m$. Then, we have

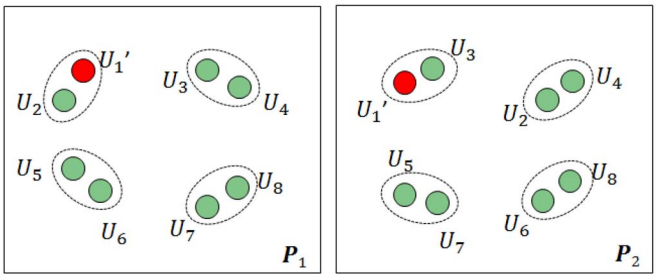


Fig. 5. Error detection example of DG-APED. (a) Aggregation in group set $\mathbb{P}_1$. (b) Detection in group set $\mathbb{P}_2$.

$X = X_1 + X_2 + \cdots + X_m \sim B(n, p)$, in which $n = n_1 + n_2 + \cdots + n_m$.

Theorem 1: Suppose each meter's reading is in the range $\{0, 1, \ldots, \Delta\}$, u and v are two integers and have the relationship $|u-v| \leq \Delta$, r is a random variable with unbiased binomial distribution $B(n, 1/2)$. Then $Pr[u + r = k] \leq e^{\epsilon} \cdot Pr[v + r = k] + \delta$ for any integer k, in which $\epsilon > 0$ and $\delta > 0$. The parameter n is chosen to be at least $\hbar = 64\Delta^2 \log(2/\delta)/\epsilon^2$. The proof of Theorem 1 is shown in Appendix A.

B. Basic Idea of DG-APED

Now, we give the basic idea of our protocol. Similar to the G-PSA, the CC aggregates all users' data directly when there is no malfunctioning SMs, otherwise, the CC aggregates in groups. To improve efficiency and provide error detection ability, the KMC classifies all the SMs into different types according to the distribution of their lifetimes. Then, the KMC divides all SMs into small group repeatedly in random for each SM type. The probability that malfunctioning SMs group with functioning SMs can be decreased, and consequently, the computation efficiency can be improved significantly.

In our model, the attacker has no way to access the private keys of any users, even the malfunctioning users, thus the groups consist malfunctioning SMs cannot be aggregated by the CC. To detect the malfunctioning SMs, the CC will go through all the groups that consist the suspected-malfunctioning SMs.

We give an illustrative example shown in Fig. 5 to explain our idea. Assume the user U_1 is malfunctioning, then group $\{U_1, U_2\}$ in group set $\mathbb{P}_1$ cannot be decrypted successfully. The CC will denote U_1 and U_2 as "suspected-malfunctioning SM." To detect out the malfunctioning SM, the CC will try to decrypt the groups $\{U_1, U_3\}$ and $\{U_2, U_4\}$ in group set $\mathbb{P}_2$. The CC will find neither of groups consisting user U_1 can be decrypted successfully, then U_1 can be detected out. For U_2, the group $\{U_2, U_4\}$ can be decrypted successfully, then U_2 can be vindicated as a functioning SM.

However, the practical smart grid scenario is much more complicated, as there are usually more than one malfunctioning SMs and the detection may not work out only in one group set. These issues are addressed in the proposed aggregation DG-APED protocol.

C. Detailed Design of DG-APED

Now, we present our proposed DG-APED in detail, which is an extension of G-PSA.

1) System Initialization: In this phase, the trusted KMC initializes the whole system.

Similar to that in G-PSA, the KMC generates the parameters $(\mathbb{G}, g, g_1, g_2, p)$ and a cryptographic hash function $H(t) = g_1^{h_1(t)} \cdot g_2^{h_2(t)}$, in which g, g_1, g_2 are random generators and $h_1(t), h_2(t)$ are also two cryptographic hash functions $\{0,1\}^* \rightarrow \mathbb{G}$. Then, the KMC runs the following steps to distribute key materials to residential users $\mathbb{U} = \{U_1, U_2, \ldots, U_N\}$ and the CC.

Step 1: For each $U_i \in \mathbb{U}$, the KMC chooses a random number $sk_i \in \mathbb{Z}_p^*$ and assigns sk_i as the private key to U_i.

Step 2: Different from the G-PSA, the KMC classifies the SMs into different types SM_l, $1 \leq l \leq L$ according to their functioning probability at time t, namely, $\rho_0(\Delta t)$.

Step 3: For SM_l, the KMC runs the random user grouping w_l rounds. For the vth user grouping, where $1 \leq v \leq w_l$, the KMC performs the following operations.

1) Similar to that in Fig. 4(a), the KMC randomly divides N_l users into N_l/z_l groups for SM_l-type users, and denotes the random grouping set as $\mathbb{P}_v^l$.
2) For each SM group $P_u^{v,l} \in \mathbb{P}_v^l$, the KMC computes the corresponding key $sk_{u0}^{v,l} \in \mathbb{Z}_p^*$ such that $sk_{u0}^{v,l} + \sum_{U_i \in P_u^{v,l}} sk_i = 0 \bmod p$, and assigns $g_1^{sk_{u0}^{v,l}}, g_2^{sk_{u0}^{v,l}}$ to the CC.
3) For each SM type SM_l, the KMC calculates the corresponding key $sk_{l0}^v = \sum_{P_u^{v,l} \in \mathbb{P}_v^l} sk_{u0}^{v,l} \bmod p$, and assign $g_1^{sk_{l0}^v}, g_2^{sk_{l0}^v}$ to the CC.
4) At last, the KMC calculates the global key $sk_0^v = \sum_{l=1}^{L} sk_{l0}^v \bmod p$, and assigns $g_1^{sk_0^v}, g_2^{sk_0^v}$ to the CC.

For any grouping set $\mathbb{P}_v$, it is easy to verify that $sk_{l0}^v = sk_{l0} = -\sum_{U_i \in \mathrm{SM}_l} sk_i \bmod p$ and $sk_0^v = sk_0 = -\sum_{U_i \in \mathbb{U}} sk_i \bmod p$.

Remark 1: Any group P_u should be distinct in all random groups sets $\mathbb{P}_1, \ldots, \mathbb{P}_{w_l}$.

According to the combinatorics design theory, the requirement in Remark 1 can be easily achieved for small $w_l \ll N_l$, e.g., $N_l = 1000$, $w_l = 10$.

2) Data Encryption and Reporting: At every time t, each user $U_i \in \mathbb{U}$ perturbs his/her sensed data $x_i(t)$ with generated noise $r_i \sim B(\hbar/z_{\min}, 1/2)$, $z_{\min} = \min\{z_1, z_2, \ldots, z_L\}$, $\hbar = 64\Delta^2 \log(2/\delta)\epsilon^2$. Then user U_i utilizes his/her private key sk_i to encrypt the randomized data $\widehat{x}_i(t) = x_i(t) + r_i \bmod p$ as

$$\begin{aligned} c_i(t) &= \mathsf{Encrypt}(sk_i, \widehat{x}_i(t)) = g^{\widehat{x}_i(t)} \cdot H(t)^{sk_i} \\ &= g^{\widehat{x}_i(t)} \cdot g_1^{h_1(t) \cdot sk_i} \cdot g_2^{h_2(t) \cdot sk_i} \end{aligned} \tag{7}$$

and then reports $U_i || c_i(t) || t$ to the CC.

3) Aggregation With Error Detection: Upon receiving all users' ciphertexts $c_i(t)$, for $i = 1, \ldots, N$ at time t, the CC uses the parameters $g_1^{sk_0}$, $g_2^{sk_0}$, $h_1(t)$, and $h_2(t)$ to try to recover the value of $\sum_{i=1}^{N} \widehat{x}_i(t)$. If all SMs work normally, we have the

aggregated result for all users by the brute-force search

$$\begin{aligned} D(t) &= \left(\prod_{i=1}^{N} c_i(t)\right) \cdot \left(g_1^{sk_0}\right)^{h_1(t)} \cdot \left(g_2^{sk_0}\right)^{h_2(t)} \\ &= \left(\prod_{i=1}^{N} g^{\widehat{x}_i(t)} \cdot g_1^{h_1(t)\cdot sk_i} g_2^{h_2(t)\cdot sk_i}\right) \cdot g_1^{h_1(t)\cdot sk_0} g_2^{h_2(t)\cdot sk_0} \\ &= g^{\sum_{i=1}^{N} \widehat{x}_i(t)} \cdot g_1^{h_1(t)\cdot \sum_{i=0}^{N} sk_i} \cdot g_2^{h_2(t)\cdot \sum_{i=0}^{N} sk_i} \\ &= g^{\sum_{i=1}^{N} \widehat{x}_i(t)}. \end{aligned} \tag{8}$$

$\mathsf{Decrypt}(sk_0, c_1(t), \ldots, c_N(t))$ is used to denote this process.

However, when there are malfunctioning SMs, the CC will fail to obtain the aggregated value $\sum_{i=1}^{N} \widehat{x}_i(t)$, because the incorrect private keys of malfunctioning SMs make $sk_0 + sk_1 + sk_2 + \cdots + sk_N \neq 0 \bmod p$. Then the CC tries to decrypt the received information in different types

$$X_l(t) = \sum_{U_i \in \mathrm{SM}_l} \widehat{x}_i(t) = \mathsf{Decrypt}\big(sk_{l0}, \{c_i(t)\}_{U_i \in \mathrm{SM}_l}\big). \tag{9}$$

The decryption of the aggregated statistics requires taking a discrete logarithm by brute-force method, which involves enumerating the plaintext space and trying each possible plaintext. Thus, the decryption can be done in at most $N\Delta$ tries using the brute-force approach, where N is the number of SMs and Δ is the upper bound of SMs reading. If the CC cannot find the plaintext after enumerating all the plaintexts, it will consider that there are malfunctioning SMs. Note that, the private key is unavailable for malfunctioning SM, without corresponding private key the information of malfunctioning SM will be decrypted into meaningless mess code or unreasonable value in a high probability, which will be discussed in following section.

When there are malfunctioning SMs, at least one of $X_l(t)$ cannot be decrypted successfully. Because the incorrect private keys of malfunctioning SMs make $sk_{l'0} + \sum_{U_i \in \mathrm{SM}_{l'}} sk_i \neq 0 \bmod p$, and hence $\sum_{U_i \in \mathrm{SM}_{l'}} \widehat{x}_i(t)$ cannot be separated from (9). In this case, for the successfully decrypted type SM_l, the CC could aggregate their values; for the type $\mathrm{SM}_{l'}$ which cannot be decrypted successfully, the error detection algorithm depicted in Algorithm 1 should be launched.

The $\mathsf{Decrypt}(sk_{u0}, \{c_i(t)\}_{U_i \in P_u})$ denotes the decryption of each group. The detail is depicted

$$\begin{aligned} d_u(t) &= \left(\prod_{U_i \in P_u} c_i(t)\right) \cdot \left(g_1^{sk_{u0}}\right)^{h_1(t)} \cdot \left(g_2^{sk_{u0}}\right)^{h_2(t)} \\ &= \left(\prod_{U_i \in P_u} g^{x_i(t)} g_1^{h_1(t)\cdot sk_i} g_2^{h_2(t)\cdot sk_i}\right) \cdot g_1^{h_1(t)\cdot sk_{u0}} g_2^{h_2(t)\cdot sk_{u0}} \\ &= g^{\sum_{U_i \in P_u} x_i(t)} \cdot g_1^{h_1(t)\sum_{U_i \in P_u} sk_i + sk_{u0}} \\ &\quad \cdot g_2^{h_2(t)\sum_{U_i \in P_u} sk_i + sk_{u0}} \\ &= g^{\sum_{U_i \in P_u} x_i(t)}. \end{aligned} \tag{10}$$

After the execution of Algorithm 1, the CC can efficiently detect the malfunctioning meters. In addition, the CC can also fault-tolerantly recover the aggregation result $\sum \widehat{x}_i(t)$ for $N - \sum_{l=1}^{L} z_l \cdot m_l$ functioning SMs. According to Fact 1, $\sum_{i=1}^{N} r_i \sim B(N\hbar/z_{\min}, 1/2)$, the average of added noise is $N\hbar/2z_{\min}$, thus the releasing output is $\sum \widehat{x}_i(t) - N\hbar/2z_{\min}$.

Algorithm 1 Error Detection

procedure ERRORDETECTIONALGORITHM
  If $\mathsf{Decrypt}(sk_{l0}, \{c_i(t)\}_{U_i \in \mathrm{SM}_l})$ fails, the CC applies this algorithm for error detection in random group set $\mathbb{P}_v$, $1 \le v \le w_l$
  According to the random group set $\mathbb{P}_1$, for each group $P_u \in \mathbb{P}_1$, decrypt $\sum_{U_i \in P_u} \widehat{x}_i(t) = \mathsf{Decrypt}(sk_{u0}, \{c_i(t)\}_{U_i \in P_u})$. If the result of Decrypt is recovered, $U_i \in P_u$ work normally; otherwise, at least one of $U_i \in P_u$ is malfunctioning
  After N_l/z_l Decrypt operations, the CC can find out all malfunctioning groups (e.g., m_l groups) that failed in the decryption, where $m_l < N_l/z_l$
  [Fault Tolerance] For the rest $(N_l/z_l - m_l)$ correct decryption groups, the CC sums up their results as the nearest aggregation value for SM type SM_l
  [Error Detection] For each user $U_{i'}$ in the malfunctioning groups, the following steps are performed:
  for $v = 2, 3 \ldots w_l$ **do**
    According to random group $\mathbb{P}_v$, the CC tries to decrypt $U_{i'}$ in a new group $P_{u'} = (U_{i'}, \{U_i\}) \in \mathbb{P}_v$
    if the Decrypt on group $P_{u'}$ can be decrypted successfully **then**
      denote the $U_{i'}$ as a functioning SM
      break
    end if
    if the group $P_{u'}$ fails in the Decrypt decryption **then**
      do the next round of decryption for $U_{i'}$ based on another random group set
    end if
  end for
  if no Decrypt decryption on $U_{i'}$ is correct for the random groups $\mathbb{P}_2, \ldots, \mathbb{P}_{w_l}$ **then**
    denote $U_{i'}$ as a malfunctioning SM
  end if
end procedure

D. Dynamic Join and Leave

For each SM type, first, we consider that the number of SM after the newly joined SM is exactly multiple of group size, that is, $N_l = k \cdot z_l$, $k \in Z^*$.

1) When a new SM U_j joins, it is supposed to contact the trusted KMC, classified into corresponding SM type and assigned a private key sk_j for encryption. Meanwhile, the CC will be assigned corresponding decryption materials. However, the existing SMs need not to be notified.
2) When an existing SM U_i leaves, its private key will be erased permanently to guarantee the security of aggregation. The vacancy generated by the left SM will be replaced with a dummy SM which will encrypt meaningless information with a new private key. However, to reduce system redundancy, sometimes we

may consider repeating the system initialization phase when too many SMs have left.

If N_l is not multiple of z_l, we will round it up to the nearest multiple of z_l, and the dummy SM will "submit" meaningless information to the CC to balance out the privacy-preserving aggregation. To enable that users can join or leave dynamically, the KMC must be available for the newly joined SMs, but do not have to be online for the periodic aggregation phases.

VI. Security Analysis

In this section, the security of the proposed protocol will be discussed. Following the attack model discussed earlier, our analysis mainly focus on how the proposed DG-APED can achieve the confidentiality of the communications from the users to the CC, the individual residential user's report privacy preservation and the malfunctioning SMs detection with a high probability.

A. Confidentiality of the Data Aggregation Communication From the Users to the CC is Achieved

In the proposed DG-APED, each residential user U_i's data is encrypted as $c_i(t) = g^{\widehat{x}_i(t)} \cdot H(t)^{sk_i}$ by SMs before being reported to the CC. Without knowing the private key sk_i, it is infeasible for an attacker to access the user U_i' data $x_i(t)$.

On the other hand, because the relationships that $sk_{u0} + \sum_{U_i \in P_u} sk_i = 0 \bmod p$, $sk_{l0} + \sum_{U_i \in \mathrm{SM}_l} sk_i = 0 \bmod p$ and $sk_0 + \sum_{i=1}^{n} sk_i = 0 \bmod p$, the CC can always decrypt the correct aggregation values even when there are malfunctioning SMs.

As a result, the proposed DG-APED achieves the confidentiality, i.e., the information is secure against the external attacker's eavesdropping attack on data aggregation communications.

B. Individual Residential User's Report is Privacy-Preserving

In order to prevent the privacy of residential users from disclosing, the random grouping mechanism and differential privacy technique are applied to our proposed protocol.

On one hand, the CC is not allowed to recover each individual user's data, but the whole aggregate data. In the proposed DG-APED, the CC aggregates the sensed data from more than one SM and detects malfunctioning SMs by decrypting in different groups with different key sets. When there is no malfunctioning SM in the residential area, the sum of all residential users' data can be recovered from the reported ciphertexts $c_1(t), c_2(t), \ldots, c_N(t)$ with decryption keys $g_1^{sk_0}$ and $g_2^{sk_0}$, that is

$$\begin{aligned}
&\left(\prod_{i=1}^{N} c_i(t)\right) \cdot \left(g_1^{sk_0}\right)^{h_1(t)} \cdot \left(g_2^{sk_0}\right)^{h_2(t)} \\
&= \left(\prod_{i=1}^{N} g^{x_i(t)} \cdot g_1^{h_1(t) \cdot sk_i} \cdot g_2^{h_2(t) \cdot sk_i}\right) \cdot g_1^{sk_0 \cdot h_1(t)} \cdot g_2^{sk_0 \cdot h_2(t)} \\
&= g^{\sum_{i=1}^{N} \widehat{x}_i(t)} \cdot g_1^{h_1(t) \cdot \left(sk_0 + \sum_{i=1}^{N} sk_i\right)} g_2^{h_2(t) \cdot \left(sk_0 + \sum_{i=1}^{N} sk_i\right)} \\
&= g^{\sum_{i=1}^{N} \widehat{x}_i(t)} \cdot H(t)^{sk_0 + \sum_{i=1}^{N} sk_i}. \qquad (11)
\end{aligned}$$

Since $sk_0 + sk_1 + sk_2 + \cdots + sk_N = 0 \bmod p$, the CC can only access the aggregate data $\sum_{i=1}^{N} \widehat{x}_i(t)$ from $g^{\sum_{i=1}^{N} \widehat{x}_i(t)}$. Except the aggregate data $\sum_{i=1}^{N} \widehat{x}_i(t)$, the CC is unable to learn anything more about each individual user' data. When there are malfunctioning SMs, the CC will fail to recover the aggregate data for all residential users, and employ the $\mathsf{Decrypt}(sk_{l0}, \{c_i(t)\}_{U_i \in \mathrm{SM}_l})$ for functioning SM type SM_l and $\mathsf{Decrypt}(sk_{u0}, \{c_i(t)\}_{U_i \in P_u})$ for each group in malfunctioning SM type $\mathrm{SM}_{l'}$. If all SMs in group P_u are functioning, we have

$$\begin{aligned}
&\left(\prod_{U_i \in P_u} c_i(t)\right) \cdot \left(g_1^{sk_{u0}}\right)^{h_1(t)} \cdot \left(g_2^{sk_{u0}}\right)^{h_2(t)} \\
&\qquad = g^{\sum_{U_i \in P_u} \widehat{x}_i(t)} \cdot H(t)^{sk_{u0} + \sum_{U_i \in P_u} sk_i}. \qquad (12)
\end{aligned}$$

Since $sk_{u0} + \sum_{U_i \in P_u} sk_i = 0 \bmod p$, the $\sum_{U_i \in P_u} \widehat{x}_i(t)$ can be retrieved from $g^{\sum_{U_i \in P_u} \widehat{x}_i(t)}$. From the recovered sum $\sum_{U_i \in P_j} \widehat{x}_i(t)$, it is still hard to get the individual data. If at least one of users $U_i \in P_u$ is malfunctioning, $\sum_{U_i \in P_u} \widehat{x}_i(t)$ cannot be recovered, let alone the values of individual users. Therefore, even the CC cannot get individual user data when some malfunctioning SMs exist.

On the other hand, the added noise r_i perturbs the output of each user, which achieves differential privacy to resist malicious data mining attack. Let $R = \sum_{i=1}^{N} r_i$, since $r_i \sim B(\hbar/z_{\min}, 1/2)$, we have $R \sim B(N\hbar/z_{\min}, 1/2)$ according to Fact 1. Thus the aggregated result $\sum_{i=1}^{N} \widehat{x}_i(t) = \sum_{i=1}^{N} x_i(t) + \sum_{i=1}^{N} r_i$ is differentially private based on Theorem 1. To measure the utility of our perturbation, we define the relative error $\gamma = (|\sum \widehat{x}_i(t) - N\hbar/z_{\min} - \sum x_i(t)|)/(\sum x_i(t) + 1)$, whose expectation is

$$\begin{aligned}
E(\gamma) &= \frac{1}{\sum x_i(t) + 1} E\left|\sum x_i(t) + \sum_{i=1}^{N} r_i - \frac{N\hbar}{z_{\min}} - \sum x_i(t)\right| \\
&= \frac{1}{\sum x_i(t) + 1} E\left|\sum_{i=1}^{N} r_i - N\hbar/z_{\min}\right|. \qquad (13)
\end{aligned}$$

In binomial distribution, since the average of $E(R)$ is $N\hbar/2z_{\min}$, we have $E(\gamma) = (1/\sum x_i(t) + 1)E|R - E(R)|$. Note that $E|R - E(R)|$ is the mean absolute deviation of the binomial distribution $B(n, p)$, denoted as $D_n(p)$

$$D_n(p) = 2(1-p)^{n - \lfloor np \rfloor} p^{\lfloor np \rfloor + 1} (\lfloor np \rfloor + 1) \binom{n}{\lfloor np \rfloor + 1}. \qquad (14)$$

As a result, $E(\gamma) = (D_n(p))/(\sum x_i(t) + 1)$. According to [30], $S_n(p)/\sqrt{2} \le D_n(p) \le S_n(p)$, where $S_n(p) = \sqrt{np(1-p)}$. According to the above analysis, the proposed DG-APED scheme can achieve $O(\sqrt{N/z_{\min}})$ error.

C. Malfunctioning SMs Can be Detected Out With High Probability

In our protocol, both encryption and decryption involve modular arithmetic, therefore each ciphertext or plaintext is in $[0, p-1)$. In presence of a malfunctioning SM U_i, the final aggregated ciphertext misses the parameter $H(t)^{sk_i}$, then the decrypted data will be distributed random in $[0, p-1)$.

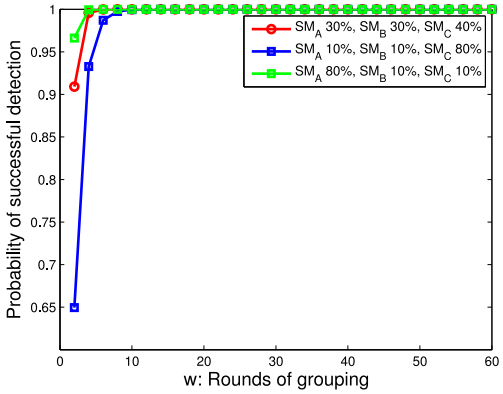


Fig. 6. Detection probability of functioning-suspected SMs in DG-APED.

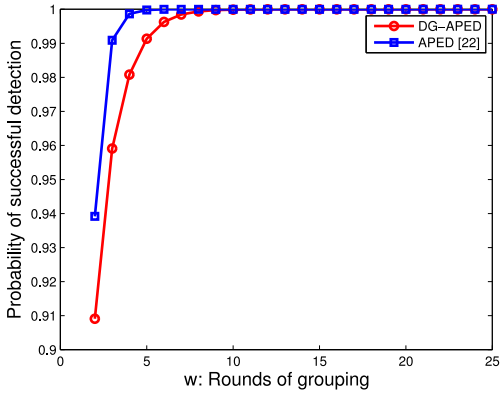


Fig. 7. Detection probability of functioning-suspected SMs in DG-APED and APED.

In this way, the probability of decrypted data in plaintext range $[0, N\Delta)$ is $N\Delta/2^p$, e.g., when $p = 512$-bit, the false positive probability is very slim.

Therefore, in presence of malfunctioning SMs, the CC will repeat w_l decryption rounds in SM_l according to Algorithm 1. After the first decryption round in $\mathbb{P}_1$, the CC can detect the malfunctioning groups easily, in which the SMs can be classified into two types: 1) the real-malfunctioning SMs and 2) the functioning-suspected SMs which are real-functioning SMs. Now, we analyze the detection probability from two aspects: 1) the detection probability of malfunctioning SMs and 2) the detection probability of functioning-suspected SMs.

1) For malfunctioning SM, after failing in $w_l - 1$ decryption rounds according to $\mathbb{P}_i$, for $i = 2, \ldots, w_l$, it can be detected out with probability $\Pr = 1$.
2) For functioning-suspected SM, it can be vindicated as functioning SM once being decrypted successfully.

In the best case, it can be detected after the second round by grouping with $z_l - 1$ functioning SMs in $\mathbb{P}_2$. However, in the worst case, the functioning-suspected SM is grouped with different malfunctioning SMs in $\mathbb{P}_2, \mathbb{P}_3, \ldots, \mathbb{P}_{w_l-1}$ and with $z_l - 1$ functioning SMs in $\mathbb{P}_{w_l}$, then it cannot be detected out until the last decryption.

When a functioning-suspected SM is always grouped with malfunctioning SMs in all group sets $\mathbb{P}_2, \mathbb{P}_3, \ldots, \mathbb{P}_{w_l}$, it will fail in detection. According to the proposed protocol in Section V, the probability that the SMs except the functioning-suspected SM in a group are all functioning is $\rho_0^{z_l-1}$, where ρ_0 denotes the probability that a SM is functioning and z_l denotes the number of SMs in a group. Otherwise, the probability that a SM group consists malfunctioning SMs is $\rho = 1 - \rho_0^{z_l-1}$, that is, the probability of a SM group cannot be decrypted.

If the CC fails in decryption based on all groups $\mathbb{P}_2, \mathbb{P}_3, \ldots, \mathbb{P}_{w_l}$, then the functioning-suspected SM is misjudged as malfunctioning SM with the misjudge probability ρ^{w_l-1}. Then, we can easily have the probability of successful detection for functioning-suspected SM as $\Pr = 1 - \rho^{w_l-1}$, which is depicted in Fig. 6, where the simulation conditions are as follows. The expected value of SMs lifetime is ten years, so the distribution of SMs lifetime is exp(10), and the SMs are classified according to their probability from $\rho_0(0)$ to $\rho_0(3)$, $\rho_0(3)$ to $\rho_0(6)$, and $\rho_0(6)$ to $\rho_0(10)$ in SM types SM_A, SM_B, and SM_C, respectively. For each type, the corresponding size of group is $z_A = 8, z_B = 4, z_C = 2$ and the corresponding rounds of grouping is $w_A = 3w, w_B = 2w, w_C = w$.

From Fig. 6, we find out that the detection probability Pr is approaching to 1 with appropriate groups. The more the grouping rounds, the higher the successful detection probability for functioning-suspected SMs. In addition, the proportion of each SM type also has some degree of influence on the results. For the same grouping rounds, the smart grid with higher proportion of SM_A has higher detection probability of functioning SMs.

Based on above analysis, real-malfunctioning SM will be find out and functioning-suspected SM will be vindicated as functioning SM in a high probability. In a word, through our algorithm the false positive probability is 0 and the false negative probability is low.

We compare the proposed DG-APED with APED [25] under the same simulation conditions as in Fig. 7. The proportions of $\text{SM}_A, \text{SM}_B, \text{SM}_C$ evenly distributed with time are 30%, 30%, 40%, respectively. The simulation result depicted in Fig. 6 shows that although DG-APED uses a more complex grouping scheme, its detection probability can catch up with APED after about ten grouping rounds.

VII. Performance Evaluation

In this section, to evaluate the performance of DG-APED, we show its effectiveness based on utility analysis and demonstrate its efficiency by considering the computational complexity and communication overhead.

A. Utility Analysis

When there are malfunctioning SMs, the CC can only aggregate the data of functioning groups. In this way, the data of functioning-suspected SMs in the malfunctioning groups will be lost. Since the number of functioning-suspected SMs, denoted as N_θ, is much less than N which is the number of all SMs in smart grid, the lost data is negligible in comparison with the real-aggregation result. In Fig. 8, we simulate the accuracy of the proposed DG-APED for different malfunctioning SMs number m and SMs number N. As shown in the figure, the relative accuracy is above 90% in the worst case.

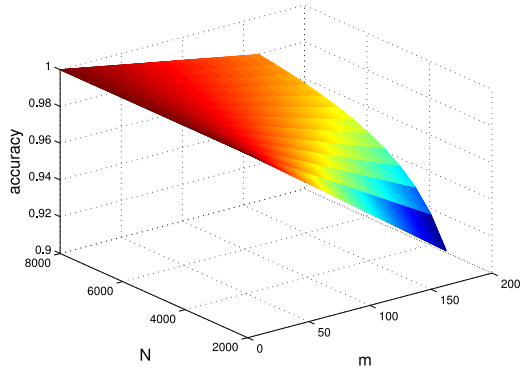


Fig. 8. Accuracy of aggregation.

That is, the proposed protocol can guarantee the validity of aggregation data.

B. Computational Complexity

For the proposed DG-APED, each user U_i only needs to store a key, i.e., sk_i and the CC has to store $2\sum_{l=1}^{L}(w_l \cdot (n_l/z_l)) + 2L + 2$ keys, i.e., $\{g_1^{sk_{u0}}, g_2^{sk_{u0}}\}_{P_u \in \mathbb{P}}$, $\{g_1^{sk_{l0}}, g_2^{sk_{l0}}\}_{l=1}^{L}$ and $g_1^{sk_0}, g_2^{sk_0}$. Thus, the number of stored keys in the system is proportional to number of SMs N_l, rounds of random groupings w_l, and inversely proportional to size of group z_l.

To evaluate the computational complexity of the error detection algorithm proposed in this paper, we denote that the number of malfunctioning SMs in the whole smart grid is m and in SM_l is m_l. For SM_l, we consider the worst case, where the CC detects out m_l malfunctioning groups from the first random group $\mathbb{P}_1$, i.e., each malfunctioning group contains a malfunctioning SM and $z_l - 1$ functioning-suspected SMs. Denoting one decryption processing as a computation cost unit, the computation cost for detecting malfunctioning groups in $\mathbb{P}_1$ is $o_1 = \sum_{l=1}^{L} N_l/z_l$. Then the CC detects malfunctioning SMs in random groups $\mathbb{P}_i$, for $i = 2, 3, \ldots, w_l$. The computation cost for detecting out a malfunctioning SM is $w_l - 1$, and we can easily have that the computation cost for m_l malfunctioning SMs is $m_l(w_l - 1)$. For the whole system, the computation cost for malfunctioning SM is $o_2 = \sum_{l=1}^{L} m_l(w_l - 1)$.

For a functioning-suspected SM, the CC has to perform $w_l - 1$ decryption rounds in the worst case, and one decryption round in the best case. Since a functioning-suspected SM can be detected out only by grouping with $z_l - 1$ functioning SM, the probability for the ith successful detection is a geometric distribution, $\rho^{i-1}(1-\rho)$, $i = 1, 2, \ldots, w_l - 1$, the mean value of which is

$$\overline{o}_l = \frac{1-\rho^{w_l-1}}{1-\rho} - (w_l - 1)\rho^{w_l-1}. \tag{15}$$

The correctness of computation cost is proofed in Appendix B.

Then, we can have the average computation cost to successfully detect out m_l functioning-suspected SMs: $m_l \cdot ((1-\rho^{w_l-1})/(1-\rho) - (w_l-1)\rho^{w_l-1})$. For the whole system, the average computation cost for detection of functioning-suspected SMs is $o_3 = \sum_{l=1}^{L} m_l((1-\rho^{w_l-1}/1-\rho) - (w_l-1)\rho^{w_l-1})$.

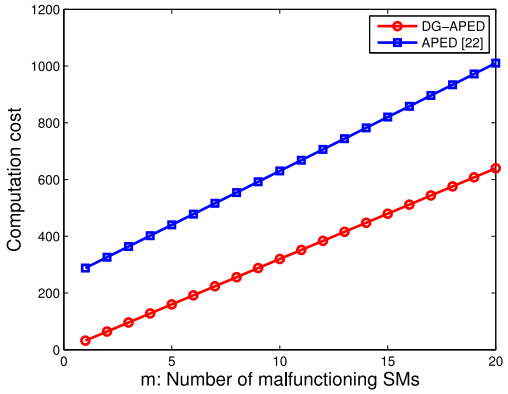


Fig. 9. Computation cost versus variable m ($N = 500, w = 20$).

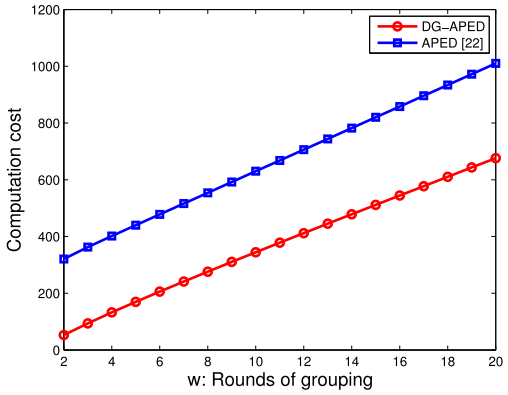


Fig. 10. Computation cost versus variable w ($N = 500, m = 20$).

According to the above analysis, we have the total computation cost for error detection in the worst case

$$o = \sum_{l=1}^{L}\left(\frac{N_l}{z_l} + m_l(w_l - 1) + m_l\left(\frac{1-\rho^{w_l-1}}{1-\rho} - (w_l-1)\rho^{w_l-1}\right)\right). \tag{16}$$

With the simulation condition in Section V, Figs. 9 and 10 compare the DG-APED with APED in terms of the computation cost versus the number of malfunctioning SMs $m = \sum_{l=1}^{L} m_l$ and the round of random groupings $w = \sum_{l=1}^{L} w_l$, respectively. As shown in Figs. 9 and 10, the DG-APED is much more efficient than APED in term of total computation cost.

Based on the above analysis, w and z are related to the detection probability and computation cost. Since z is a constant value in general, the value of w is critical to the performance of our proposed protocol. By comparing Fig. 7 with Fig. 10, it can be seen that the larger w, the higher detection probability and the higher computation cost. In conclusion, it is a trade-off between detection probability and computation cost in our proposed protocol.

C. Practical Performance

In the cryptographic construction of the proposed aggregation protocol, encryption involves a hash operation, two modular exponentiations, and one multiplication. The running time is dominated by the exponentiation, as the time for the

hash function and multiplication are much smaller in comparison with the time for an exponentiation. On a modern 64-bit desktop computer, it takes roughly 0.3 ms to compute a modular exponentiation using high-speed elliptic curves [23]. Therefore, each encryption takes about 0.6 ms on a modern computer. Each user only needs to perform one encryption, therefore, a user's computation overhead is 0.6 ms which is beneficial for users.

Decryption of the aggregated statistics requires taking a discrete logarithm. The brute-force method involves enumerating the plaintext space and taking 0.3 ms to try each possible plaintext. Thus, the decryption can be done in at most $N\Delta$ tries using the brute-force approach, where N is the number of users and Δ is the range of users' data. For example, when each participant's value is in the set $\{0, 1, \ldots, 5\}$, and when there are 1000 users, the decryption can be done in about 1.5 s. However, to decrypt the ciphertext by brute force, the CC can only perform at most $N\Delta$ multiplication instead of modular exponentiation, which can greatly decrease the decryption time. Besides, we can speed up decryption significantly by using Pollard's rho method [31] for discrete logarithm, which reduces the decryption overhead to about $\sqrt{N\Delta}$ group operations. Therefore, the aggregation can be guaranteed in real-time.

In the detection process, the CC has to perform o decryptions, so the practical time for detection is $0.3 \cdot o \cdot N\Delta$ using the brute-force approach. Under the simulation condition in Section VII-B, the detection can be done in less than half an hour, which is available since real-time is not required in the detection process.

D. Communication Overhead

In the proposed DG-APED, each U_i requires no peer-to-peer communication that is important in smart grid scenario as it requires all users to be online and able to interact with each other. For each aggregation process, each SM U_i has to send its encrypted report $U_i||c_i(t)||t$ to the CC. Thus, it requires only a single round of communication from users to the CC. As a result, the total communication in smart grid is $O(N)$ for DG-APED, less than that of the BA $O(NlogN)$ in [24], and no more than that of the PSA proposed in [23] and APED in [25].

Moreover, the PSA may suffer from malfunctioning SMs, and the BA and APED are not efficient when the number of users is high. In comparison, the proposed protocol can handle user failure, detect malfunctioning SMs and aggregate data efficiently without privacy disclosing. Thus, the major problems remained in [23]–[25] have been addressed in this paper.

VIII. Conclusion

In this paper, we have proposed an efficient DG-APED for smart grid. We have also provided security analysis to demonstrate its security and privacy-preserving, and conducted performance evaluation to show its effectiveness and efficiency. In future, we are ready to conduct experiments to further verify the feasibility of our proposed protocol. In addition, other security properties will also be integrated to enhance the system security and reliability.

Appendix A
Proof of Theorem 1

This section presents the proof of Theorem 1.

Proof: We consider the extreme case when $|u-v| = \Delta$. We give the proof of $u - v = \Delta$, the proof of $v - u = \Delta$ is as same as $u - v = \Delta$. Then

$$\frac{Pr[u + r = k]}{Pr[v + r = k]} = \frac{Pr[r = k - u]}{Pr[r = k - u + \Delta]}.$$

Denoting $k - u$ with $n/2 + x$, we determine the range of x which makes

$$\frac{Pr[r = n/2 + x]}{Pr[r = n/2 + x + \Delta]} = \frac{\binom{n}{n/2+x}}{\binom{n}{n/2+x+\Delta}} < e^{\epsilon}.$$

Computing the above inequality, the inequality holds as long as $x < n\epsilon/8\Delta$.

Let δ limit the relative shift in the case $x > n\epsilon/8\Delta$. That is, $Pr[r > n/2 + \epsilon n/8\Delta] \leq \delta/2$. According to the Chernoff bound, we have $Pr[r > n/2 + \epsilon n/8\Delta] \leq \exp(-\epsilon^2 n/64\Delta^2)$. Thus, we get δ-approximate ϵ-indistinguishability as long as n is chosen to be at least $64\Delta^2 \log(2/\delta)/\epsilon^2$. Thus, the proof is completed. ■

Appendix B
Proof of (15)

This section presents the proof of (15).

Proof: Now, we prove the mean value for geometric distribution $\mathbf{G} : \rho^{i-1}(1-\rho), i = 1, 2, \ldots, w_l - 1$ is right. Based on the probability for the ith successful detection of the functioning-suspected SM: $\rho^{i-1}(1-\rho)$, we have the mean value

$$\overline{o}_l = 1 \cdot (1-\rho) + 2 \cdot \rho(1-\rho) + \cdots + (w_l - 1) \cdot \rho^{w_l - 2}(1-\rho).$$

Employ dislocation phase subtraction to calculate $\overline{r}$

$$\begin{aligned}
\overline{o}_l &= 1 \cdot (1-\rho) + 2 \cdot \rho(1-\rho) + 3 \cdot \rho^2(1-\rho) + \cdots \\
&\quad + (w_l - 1) \cdot \rho^{w_l - 2}(1-\rho) \\
\rho \cdot \overline{o}_l &= 1 \cdot \rho(1-\rho) + 2 \cdot \rho^2(1-\rho) + 3 \cdot \rho^3(1-\rho) + \cdots \\
&\quad + (w_l - 2) \cdot \rho^{w_l - 2}(1-\rho) + (w_l - 1) \cdot \rho^{w_l - 1}(1-\rho).
\end{aligned}$$

Calculating $\overline{o}_l - \rho \cdot \overline{o}_l$, we have

$$\begin{aligned}
(1-\rho) \cdot \overline{o}_l &= (1-\rho) + \rho(1-\rho) + \rho^2(1-\rho) + \cdots \\
&\quad + \rho^{w_l - 2}(1-\rho) - (w_l - 1) \cdot \rho^{w_l - 1}(1-\rho) \\
&= (1-\rho) \cdot \Big[1 + \rho + \rho^2 + \cdots + \rho^{w_l - 2} \\
&\qquad - (w_l - 1) \cdot \rho^{w_l - 1}\Big] \\
&= (1-\rho) \cdot \left[\frac{1 - \rho^{w_l - 1}}{1-\rho} - (w_l - 1)\rho^{w_l - 1}\right] \\
\Rightarrow \overline{o}_l &= \frac{1 - \rho^{w_l - 1}}{1-\rho} - (w_l - 1)\rho^{w_l - 1}.
\end{aligned}$$

Thus, the proof is completed. ■

REFERENCES

[1] X. Li *et al.*, "Securing smart grid: Cyber attacks, countermeasures, and challenges," *IEEE Commun. Mag.*, vol. 50, no. 8, pp. 38–45, Aug. 2012.

[2] A. R. Metke and R. L. Ekl, "Security technology for smart grid networks," *IEEE Trans. Smart Grid*, vol. 1, no. 1, pp. 99–107, Jun. 2010.

[3] B. Chai and Z. Yang, "Impacts of unreliable communication and modified regret matching based anti-jamming approach in smart microgrid," *Ad Hoc Netw.*, vol. 22, pp. 69–82, Nov. 2014.

[4] *NIST Framework and Roadmap for Smart Grid Interoperability Standards*, NIST Standard 1108, Jan. 2010.

[5] R. V. Gerwen, S. Jaarsma, and R. Wilhite, "Smart metering," *Distrib. Gener.*, vol. 9, pp. 1–9, Jul. 2006.

[6] G. T. Heydt, "The next generation of power distribution systems," *IEEE Trans. Smart Grid*, vol. 1, no. 3, pp. 225–235, Dec. 2010.

[7] F. Rahimi and A. Ipakchi, "Demand response as a market resource under the smart grid paradigm," *IEEE Trans. Smart Grid*, vol. 1, no. 1, pp. 82–88, Jun. 2010.

[8] R. Deng, Z. Yang, J. Chen, N. R. Asr, and M.-Y. Chow, "Residential energy consumption scheduling: A coupled-constraint game approach," *IEEE Trans. Smart Grid*, vol. 5, no. 3, pp. 1340–1350, May 2014.

[9] Z. Yang, L. Sun, M. Ke, Z. Shi, and J. Chen, "Optimal charging strategy for plug-in electric taxi with time-varying profits," *IEEE Trans. Smart Grid*, vol. 5, no. 6, pp. 2787–2797, Nov. 2014.

[10] R. Lu, X. Li, X. Liang, X. Shen, and X. Lin, "GRS: The green, reliability, and security of emerging machine to machine communications," *IEEE Commun. Mag.*, vol. 49, no. 4, pp. 28–35, Apr. 2011.

[11] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "Malicious data attacks on the smart grid," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 645–658, Dec. 2011.

[12] L. Hu and D. Evans, "Secure aggregation for wireless networks," in *Proc. IEEE Symp. Appl. Internet Workshops*, Orlando, FL, USA, 2003, pp. 384–391.

[13] B. Przydatek, D. Song, and A. Perrig, "SIA: Secure information aggregation in sensor networks," in *Proc. 1st Int. Conf. Embedded Netw. Sensor Syst.*, Los Angeles, CA, USA, 2003, pp. 255–265.

[14] A. Mahimkar and T. S. Rappaport, "SecureDAV: A secure data aggregation and verification protocol for sensor networks," in *Proc. IEEE GLOBECOM*, vol. 4. Dallas, TX, USA, 2004, pp. 2175–2179.

[15] R. Lu, X. Liang, X. Li, X. Lin, and X. Shen, "EPPA: An efficient and privacy-preserving aggregation scheme for secure smart grid communications," *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 9, pp. 1621–1631, Sep. 2012.

[16] R. Lu, X. Lin, Z. Shi, and X. Shen, "EATH: An efficient aggregate authentication protocol for smart grid communications," in *Proc. IEEE WCNC*, Shanghai, China, 2013, pp. 1819–1824.

[17] M. Wen *et al.*, "PaRQ: A privacy-preserving range query scheme over encrypted metering data for smart grid," *IEEE Trans. Emerg. Topics Comput.*, vol. 1, no. 1, pp. 178–191, Jun. 2013.

[18] L. Chen, R. Lu, and Z. Cao, "PDAFT: A privacy-preserving data aggregation scheme with fault tolerance for smart grid communications," *Peer-to-Peer Netw. Appl.*, to be published.

[19] F. D. Garcia and B. Jacobs, "Privacy-friendly energy-metering via homomorphic encryption," in *Security and Trust Management*. Berlin, Germany: Springer-Verlag, 2011, pp. 226–238.

[20] K. Kursawe, G. Danezis, and M. Kohlweiss, "Privacy-friendly aggregation for the smart-grid," in *Privacy Enhancing Technologies*. Berlin, Germany: Springer-Verlag, 2011, pp. 175–191.

[21] C. Dwork, "Differential privacy," in *Automata, Languages and Programming*. Berlin, Germany: Springer-Verlag, 2006, pp. 1–12.

[22] C. Dwork, K. Kenthapadi, F. McSherry, I. Mironov, and M. Naor, "Our data, ourselves: Privacy via distributed noise generation," in *Advances in Cryptology—EUROCRYPT*. Berlin, Germany: Springer-Verlag, 2006, pp. 486–503.

[23] E. Shi, T.-H. H. Chan, E. G. Rieffel, R. Chow, and D. Song, "Privacy-preserving aggregation of time-series data," in *Proc. NDSS*, vol. 2. San Diego, CA, USA, 2011, pp. 1–17.

[24] T. Chan, E. Shi, and D. Song, "Privacy-preserving stream aggregation with fault tolerance," in *Financial Cryptography and Data Security*. Berlin, Germany: Springer-Verlag, 2012, pp. 200–214.

[25] R. Sun, Z. Shi, R. Lu, M. Lu, and X. Shen, "APED: An efficient aggregation protocol with error detection for smart grid communications," in *Proc. IEEE GLOBECOM*, Atlanta, GA, USA, 2013, pp. 1–6.

[26] H. Bao and R. Lu, "A new differentially private data aggregation with fault tolerance for smart grid communications," *IEEE Internet Things J.*, vol. 2, no. 3, pp. 248–258, Jun. 2015.

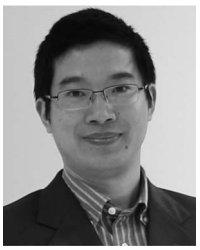
[27] Y. Yan, Y. Qian, H. Sharif, and D. Tipper, "A survey on smart grid communication infrastructures: Motivations, requirements and challenges," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 1, pp. 5–20, Mar. 2013.

[28] M. G. Pecht and F. Nash, "Predicting the reliability of electronic equipment," *Proc. IEEE*, vol. 82, no. 7, pp. 992–1004, Jul. 1994.

[29] W. Jia, H. Zhu, Z. Cao, X. Dong, and C. Xiao, "Human-factor-aware privacy-preserving aggregation in smart grid," *IEEE Syst. J.*, vol. 8, no. 2, pp. 598–607, Jun. 2014.

[30] D. Berend and A. Kontorovich, "A sharp estimate of the binomial mean absolute deviation with applications," *Stat. Prob. Lett.*, vol. 83, no. 4, pp. 1254–1259, 2013.

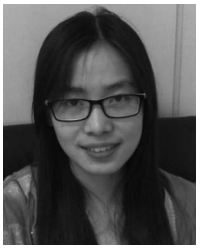
[31] J. M. Pollard, "Monte Carlo methods for index computation," *Math. Comput.*, vol. 32, no. 143, pp. 918–924, 1978.



Zhiguo Shi (M'10) received the B.S. and Ph.D. degrees in electronic engineering from Zhejiang University, Hangzhou, China, in 2001 and 2006, respectively.

Since 2006, he has been a Faculty Member with the Department of Information and Electronic Engineering, Zhejiang University, where he is currently a Full Professor. From 2011 to 2013, he was visiting the Broadband Communications Research Group, University of Waterloo, Waterloo, ON, Canada. His current research interests include signal and data processing, and smart grid communication and network.

Prof. Shi was a recipient of the Best Paper Award from the IEEE Wireless Communications and Networking Conference (WCNC) 2013, Shanghai, China; the IEEE/CIC International Conference on Communications in China 2013, Xi'an, China; and the IEEE Wireless Communications and Signal Processing 2012, Huangshan, China, and the Scientific and Technological Award of Zhejiang Province, China, in 2012. He served as an Editor for the *KSII Transactions on Internet and Information Systems* and *IET Communications* in 2003 and 2004, respectively.



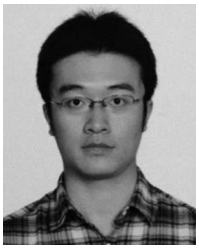
Ruixue Sun received the B.Sc. degree in communication engineering from Xidian University, Xi'an, China, in 2012. She is currently pursuing the Master's degree with the Department of Information and Electronic Engineering, Zhejiang University, Hangzhou, China.

Her current research interests include security and privacy in millimeter wave communication and smart grid.

Rongxing Lu (S'09–M'11) received the Ph.D. degree in computer science from Shanghai Jiao Tong University, Shanghai, China, in 2006, and the Ph.D. degree in electrical and computer engineering from the University of Waterloo, Waterloo, ON, Canada, in 2012.

Since 2013, he has been an Assistant Professor with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. His current research interests include computer, network and communication security, applied cryptography, security and privacy analysis for vehicular network, eHealthcare system, and smart grid communications.

Dr. Lu was a recipient of the Canada Governor General Gold Medal and the IEEE Communications Society Asia Pacific Outstanding Young Researcher Award in 2013.



Le Chen received the B.Sc. and Ph.D. degrees in computer science from Shanghai Jiaotong University, Shanghai, China, in 2008 and 2015, respectively.

From 2013 to 2015, he was a Visiting Scholar with the INFINITUS Laboratory, School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. His current research interests include secure data aggregation, wireless network security, and applied cryptography.

Jiming Chen (M'08–SM'11) received the B.Sc. and Ph.D. degrees in control science and engineering from Zhejiang University, Hangzhou, China, in 2000 and 2005, respectively.

He was a Visiting Researcher with INRIA, Rocquencourt, France, in 2006; the National University of Singapore, Singapore, in 2007; and the University of Waterloo, Waterloo, ON, Canada, from 2008 to 2010. He is currently a Full Professor with the Department of Control Science and Engineering, the Coordinator of the Group of Networked Sensing and Control, State Key Laboratory of Industrial Control Technology, and the Vice Director of the Institute of Industrial Process Control, Zhejiang University. His current research interests include sensor networks and networked control.

Dr. Chen currently serves as an Associate Editor for several international journals, including the IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEM, IEEE NETWORK, and the IEEE TRANSACTIONS ON CONTROL OF NETWORK SYSTEMS. He was a Guest Editor of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL.

Xuemin (Sherman) Shen (M'97–SM'02–F'09) received the B.Sc. degree from Dalian Maritime University, Dalian, China, in 1982, and the M.Sc. and Ph.D. degrees from Rutgers University, Newark, NJ, USA, in 1987 and 1990, respectively, all in electrical engineering.

He was a Professor and the University Research Chair with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada, where he was the Associate Chair for Graduate Studies from 2004 to 2008. His current research interests include resource management in interconnected wireless/wired networks, wireless network security, wireless body area networks, vehicular *ad hoc*, and sensor networks. He has co-authored/edited six books, and published over 600 papers and book chapters in wireless communications and networks and control and filtering.

Dr. Shen was a recipient of the Distinguished Performance Award from the Faculty of Engineering, University of Waterloo, in 2002 and 2007; the Premiers Research Excellence Award from the Province of Ontario in 2003; the Outstanding Performance Award from the University of Waterloo, in 2004, 2007, and 2010; and the Excellent Graduate Supervision Award in 2006. He has served as the Technical Program Committee Chair for the IEEE VTC 2010; the Symposia Chair for the IEEE ICC 2010; the Tutorial Chair for the IEEE VTC 2011 and the IEEE ICC 2008; the Technical Program Committee Chair for the IEEE GLOBECOM 2007 and the IEEE INFOCOM 2014; the General Co-Chair for Chinacom 2007, QShine 2006, and ACM MobiHoc 2015; and the Chair for the IEEE Communications Society Technical Committee on Wireless Communications and Peer-to-Peer (P2P) Communications and Networking. He also serves/served as an Editor-in-Chief for IEEE NETWORK, *Peer-to-Peer Networking and Application*, and *IET Communications*; a Founding Area Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS; an Associate Editor for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, *Computer Networks*, and *ACM/Wireless Networks*; and a Guest Editor for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, IEEE WIRELESS COMMUNICATIONS, IEEE COMMUNICATIONS MAGAZINE, and *ACM Mobile Networks and Applications*. He is a Distinguished Lecturer of the IEEE Vehicular Technology and Communications Society; a registered Professional Engineer of Ontario, Canada; an Engineering Institute of Canada Fellow, and a Canadian Academy of Engineering Fellow.