An Efficient Privacy-Friendly Multi-Hop Data Aggregation Scheme for Smart Grids

Prosanta Gope
National University of Singapore
21 Lower Kent Ridge Rd, Singapore 119077
Email: dcsprg@nus.edu.sg

Biplab Sikdar
National University of Singapore
21 Lower Kent Ridge Rd, Singapore 119077
Email: bsikdar@nus.edu.sg

Abstract—**In this article, we propose a privacy-friendly data aggregation scheme for smart grids. In our approach, smart meters transmit their usage reports through multi-hop communication, where an aggregation tree is constructed to minimize the aggregation overhead. Security and performance analyses show that our proposed scheme offers better privacy protection for electricity meter reading aggregation as well as computational efficiency, as compared to the existing solutions.**

Index Terms—**Privacy-friendly, Multi-hop, Data aggregation, Computational efficiency, Smart grids**

I. Introduction

Smart grids are envisioned to increase the efficiency of the current power grid, to cope up with volatile power prodution based on renewable resources for reducing the requirment of fossil-based energy resources and also to gaurantee the stability of power supply. To accomplish these goals, today's power grid is enhanced by information and communication technology (ICT) to increase the information flow and to enabled sophisticated power production and demand management strategies. However, the enhancement of power grid into a network of networks, not only makes them smarter but also more vulnerable to security threats [1]. In this context, one of the main challenges and major obstacle in the widespread deployment of smart grids is privacy, which is a primary concern from the customer's point of view. For instance, in [2-3], it is shown that complex usage patterns can be extracted from the high-resolution consumption information using simple off-the-shelf statistical tools, and the extracted information can be used to profile and monitor users for various purposes. Therefore, it becomes clear that it is of highest importance to keep power consumption data private and confidential. Furthermore, the computational resources at the consumer's side are usually very limited. Solutions for preserving user privacy should thus be computationally inexpensive.

In this paper, we propose an efficient and privacy-friendly data aggregation scheme for smart grids by using multi-hop communication. The proposed scheme only uses computationally inexpensive operations such as hash operations. Thus, the proposed scheme is well suited for the resource constrained devices in smart grids.

II. Related Work and Motivation

Several privacy-friendly data aggregation schemes have been proposed in the literature for addressing various privacy issues in smart grids (as mentioned above). For instance, Lu et al. introduced a privacy-friendly data aggregation protocol [4] by using the additive homomorphic crypto-system called Paillier encryption [5]. However, this results in high computational overhead on the entities like smart-meters. Liang et al. proposed a usage-based dynamic pricing scheme for smart grids by using a fully homomorphic technique [6]. However, Naehring et al. have shown that a fully homomorphic technique is difficult to implement with current computing resources [7]. Therefore, the scheme presented in [6] is regarded as an unrealistic one. Chia-Mu et al. introduced a ring signature based scheme to protect an individual's usage profile [8]. However, the computational cost of the proposed scheme increases with the size of the ring. Liu et al. [9] proposed an aggregation scheme based on blind signatures [9]. However, this scheme cannot protect the privacy of the consumers' usage data profile [10]. Zhang et al. proposed a self-certified signature scheme [10] and Sui et al. designed an incentive-based anonymous authentication scheme for smart grids [11]. These schemes are constructed with the assumption of an anonymity network, where the sources of usage reports are anonymous. Therefore, it is hard to identify the smart meter that produced a measurement. In [12], Kursawe et al. suggested a set of masking-based schemes for privacy in smart grids. In their schemes, the authors utilized the concept of Decisional Diffie-Hellman (DDH) group and Bilinear mapping for checking the correctness of the shared masking value, which are computationally expensive and ill-suited for resource constrained smart meters. Li et al. introduced a different technique for data aggregation in smart grids by using hop-by-hop communication [13]. However, the authors have not clarified how to construct the aggregation tree. More importantly, these schemes do not support message authentication. Therefore, a dishonest or fake smart-meter may falsify the data, which will cause inaccurate aggregation result. Recently, Knirsch et al. proposed a masking-based approach for data aggregation [14]. Their scheme utilizes the concept of homomorphic hashing for checking the correctness of the shared secrets. However, this construction has a couple of issues.

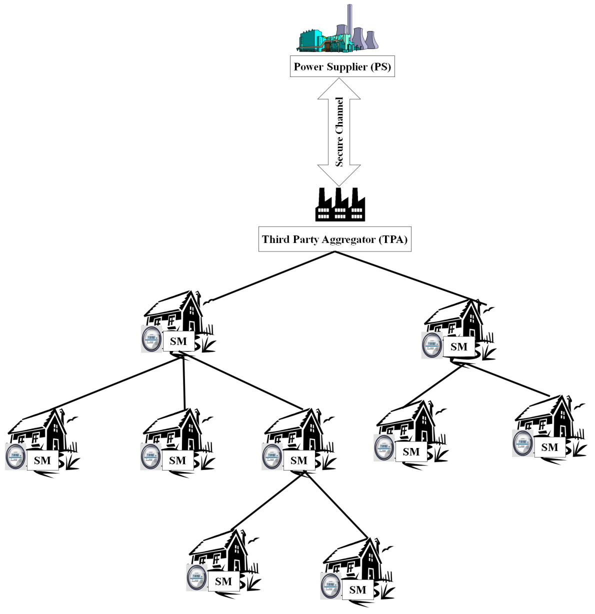

Figure 1. System Model for Multi-Hop Data Aggregation Scheme

First, it is complicated to implement and computationally expensive to compute. Second, it cannot ensure security of the hashed data, and an attacker can compute the original message block by taking the logarithm of the hash for that block. In addition, it can be shown that the data aggregation scheme presented in [14] is vulnerable to collusion attacks. In this case, when the aggregator (DC in [14]) colludes with a smart meter SM_2, then the aggregator can know the usage data of another smart meter SM_1, which is a serious privacy issue. Lastly, Mohammed et al. have proposed a multi-hop data aggregation scheme in [15], where, during data aggregation, each smart meter has to select n proxies and add masking values to their meter readings. Proxies remove the masking values added to the meters' readings to obtain an aggregated reading. However, this scheme is difficult to implement in practice, cannot ensure the integrity protection of the usage report, and does not provide sender authentication, which may cause an inaccurate aggregated result.

A. *Problem Statement and Motivation*

Conceive that there is a region/neighborhood with several apartment blocks or houses. Each apartment block or house has a set of one or more units and they are individually equipped with a smart meter. For maintaining proper balance between power generation and power consumption, the grid needs to know the aggregated electricity usage data for the entire region on a regular basis. Now, for the correctness of the aggregated usage report, the aggregator has to verify the legitimacy of each individual smart meter and the integrity of their readings. However, this will result in a very large burden on the aggregator, especially when the aggregator needs to handle a large number of smart meters. To address this issue, Li et al. [13] and Mohammed et al. [15] proposed two possible frameworks where smart meters transmit the usage reports in a hop-by-hop way. However, they have several weak points. For instance, the solutions proposed in [13] and [15] are malleable, where smart meters are not authenticated during data aggregation. Besides, source of the usage data and the identity of the smart meters are not verified. Consequently, a dishonest or fake smart meter may falsify the data, which will cause an inaccurate aggregated result. Furthermore, in many of the existing works [16-19], it is assumed that there is a secure channel between the third-party aggregator and smart meters, which is a strong assumption.

This paper seeks to address these issues by proposing an efficient and privacy-friendly data aggregation scheme for smart grids. In our approach, smart meters transmit usage reports through multi-hop communication. Even though some existing approaches can accomplish similar security features, our scheme has lower computational cost as shown by our performance analysis and experiments. Smart meters do not have to perform any computationally expensive operations (such as inefficient Paillier encryptions) during the data aggregation process. Hence, the proposed data aggregation scheme is suitable for the resource constrained devices in smart grids.

III. System and Adversary Model, and Security Goals

In this section, we first briefly describe the network architecture of privacy-friendly data aggregation for smart grids and also present the underlying adversary model. This section concludes with the security goals of the proposed scheme.

A. *System Model*

In our system model, we adopt the concept of data aggregation trees [11], which supports the usage data to be transmitted by hop-by-hop communication, as shown in Fig. 1. The system model consists of three major entities: the power supplier (PS), the third-party aggregator (TPA) employed by the power supplier, and a set of smart meters (SMs). In the system model, the PS is responsible for arranging and supplying electricity to a set of home-area-networks (HANs). Each HAN is equipped with a SM. The TPA periodically aggregates the electricity consumption of a group of HANs in a locality and helps the owners of the HANs to adjust their consumption according to the current loading conditions (e.g. through demand side management) and also inform the current demand conditions to the PS in order to help with supply-demand management. The consumption information allows the PS to optimally control its dispatchable generation, and conduct short and long term trades in the energy market. In this way, the TPA plays a crucial role in maintaining balance between power production and demand. In our system model, we assume that the TPA and the PS communicate through a secure channel. Each HAN is composed of a SM, which is assumed to be resistant against tampering. The SMs form a tree topology and report consumers' energy usage to the TPA through multi-hop communication.

B. Adversary Model

In our adversary model we consider the PS to be a trusted organization (e.g., owned by the government, such as Singapore Power in Singapore and National Grid in United Kingdom). On the other hand, the TPA is owned by a private company whose main responsibility is to assist the PS. Therefore, in our system model we consider the TPA as an honest-but-curious entity, who may want to know the consumption data of each HAN and subsequently try to sell the usage information to another company (e.g., for marketing materials for home appliances). On the other hand, we assume that various elements inside the core network may also act as adversaries and be interested in private details of the power consumption of each HAN. A compromised network and its various elements (e.g. router or switch) can alter or fabricate a meter's consumption data. Hence, any communication through the network may not be secure. Also, any SM may be the adversary and be interested to know the consumption data of another SM from a different HAN. An outside attacker may also try to impersonate a legitimate entity such as a SM or the TPA to send data under its name. For instance, a dishonest or fake SM may falsify the data for causing inaccurate aggregation result. In addition, the outside attacker may eavesdrop on the network transmission media for obtaining the power consumption data and may also try to alter or re-transmit them.

C. Security Goals

- **Message Authentication:** In general, usage reports from each SM pass through the insecure wired and wireless links of the communication network. Therefore, before aggregating any data, the aggregator needs to validate whether the report has been received from a legitimate source or not. This will prevent inaccurate aggregation results.
- **Usage Data Privacy:** The secrecy of the end-to-end communication is vital. For example, if an adversary can know the power consumption data from a HAN, then he/she can determine its occupancy. This information can be used by robbers to determine the best day or time to break into a home. Therefore, the electricity consumption data is required to be kept secret from any third party for protecting the privacy of the customer. In this regard, if an outsider or an inside adversary like other SMs from different HANs or the TPA obtains the messages with electricity consumption information, then he/she should not be able to comprehend the encrypted message.
- **Usage Data Integrity:** To avoid any inaccurate data aggregation result, the TPA must verify the integrity of the data received from each SM of a HAN. On the other hand, during data aggregation, the TPA also needs to check the integrity of the relevant information received from the PS.

IV. Proposed Scheme

In this section, we propose our efficient and privacy-friendly data aggregation scheme for smart-grids, which consists of three phases: *initialization*, *multi-hop data aggregation*, and *secure billing*. In our tree model, each parent may have n children. The responsibility of the parent node is to accumulate and validate the usage reports of its child nodes and subsequently aggregate these usage data along with its own reading and send the aggregated result to its parent. In order to simplify the description, we assume that each SM (e.g. SM_i) only has one child (SM_{i-1}). However, it can be easily extended to other tree constructions. The proposed scheme requires the formation of a topology that constructs a tree topology with the TPA at the root and the SMs as children. Algorithms for tree construction have been widely investigated in literature, specially in the context of wireless sensor networks [29-30]. These algorithms may also be used for constructing the topology for the proposed scheme. The tree construction problem is in fact simpler in our case since the devices are static and do not have a power constraint (i.e., they are not battery powered). For example, in case of individual houses on a street with a SM in each house, the algorithms from [29-30] may be directly applied without the energy constraints (or setting the energy conditions at each SM to be the same). Similarly, in apartment blocks with multiple floors and one or more apartments per floor, the first SM (or a randomly chosen SM) from a floor can act as the parent node for all the SMs in the floor immediately above it. Since existing techniques can be readily used for creating the hop-by-hop topology, we do not focus on it in this paper.

In the initialization phase of the proposed scheme, the interrelated SMs establish their common secret keys. After that, during the data aggregation process, the SMs send their usage reports on a regular basis (e.g. every 15-30 minutes) using the multi-hop data aggregation process.

A. Initialization

Assume that there are *n* HANs in a locality that obtain their electrical power from the PS. During meter installation of a home HAN_i, the PS randomly generates a shadow identity SID_i and a secret key k_i and assigns them to the SM of HAN_i. Here, we also assume that each SM is equipped with a tamper-resistant black box [17]. The black box contains a key pair (PK, SK). Any other party has access to the public key PK. However, the secret key SK is stored within the black box and is never disclosed or changed. To ensure secure communication with its neighboring SMs, each smart meter SM_i executes a key establishment protocol (e.g., the protocol proposed in [22]) with its PK and SK. Consequently, the keys $kh_{i-1,i}$ and $kh_{i,i+1}$ are shared between SM_i and SM_{i-1}, and SM_i and SM_{i+1}, respectively. The smart meters can also update the shared keys with their neighbors by executing the key establishment protocol [22]. Now, the TPA generates the key pair (SK-TPA, PK-TPA) and publishes PK-TPA to others. Similarly, the PS generates the key pair (SK-PS, PK-PS) and publishes PK-PS to others. The TPA and the PS use their secret keys (SK-TPA, SK-PS) to generate signatures. Anyone who knows their public keys (PK-TPA, PK-PS) can verify the signature.

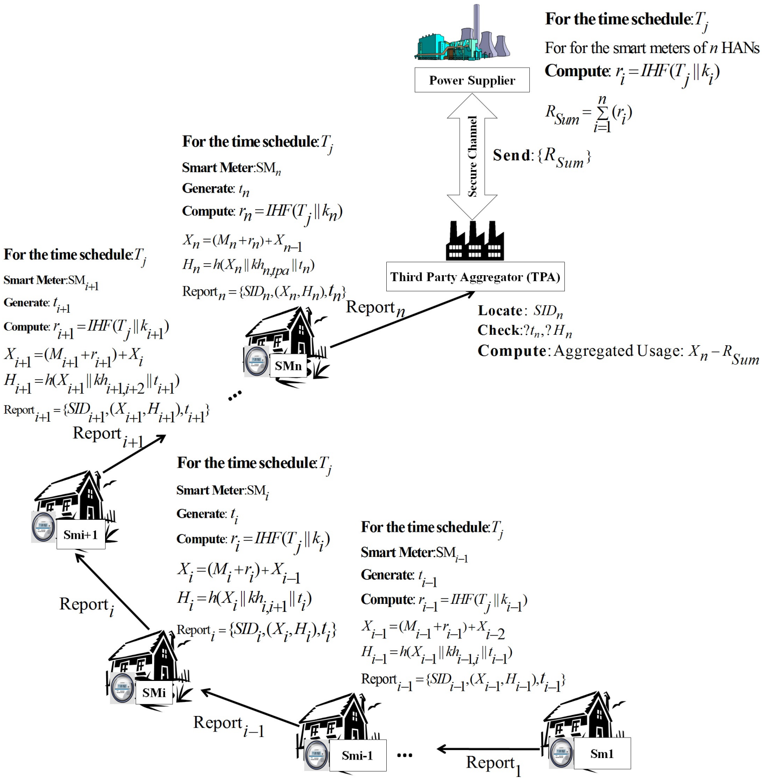
For the time schedule: T_j
For for the smart meters of n HANs
Compute: $r_i = IHF(T_j \| k_i)$
$R_{Sum} = \sum_{i=1}^{n}(r_i)$
Power Supplier
Secure Channel
Send: $\{R_{Sum}\}$
Third Party Aggregator (TPA)
Locate: SID_n
Check: $?t_n, ?H_n$
Compute: Aggregated Usage: $X_n - R_{Sum}$
For the time schedule: T_j
Smart Meter: SM_n
Generate: t_n
Compute: $r_n = IHF(T_j \| k_n)$
$X_n = (M_n + r_n) + X_{n-1}$
$H_n = h(X_n \| kh_{n,tpa} \| t_n)$
$\mathrm{Report}_n = \{SID_n, (X_n, H_n), t_n\}$
Report_n
SMn
For the time schedule: T_j
Smart Meter: SM_{i+1}
Generate: t_{i+1}
Compute: $r_{i+1} = IHF(T_j \| k_{i+1})$
$X_{i+1} = (M_{i+1} + r_{i+1}) + X_i$
$H_{i+1} = h(X_{i+1} \| kh_{i+1,i+2} \| t_{i+1})$
$\mathrm{Report}_{i+1} = \{SID_{i+1}, (X_{i+1}, H_{i+1}), t_{i+1}\}$
...
Report_{i+1}
Smi+1
For the time schedule: T_j
Smart Meter: SM_i
Generate: t_i
Compute: $r_i = IHF(T_j \| k_i)$
$X_i = (M_i + r_i) + X_{i-1}$
$H_i = h(X_i \| kh_{i,i+1} \| t_i)$
$\mathrm{Report}_i = \{SID_i, (X_i, H_i), t_i\}$
Report_i
SMi
For the time schedule: T_j
Smart Meter: SM_{i-1}
Generate: t_{i-1}
Compute: $r_{i-1} = IHF(T_j \| k_{i-1})$
$X_{i-1} = (M_{i-1} + r_{i-1}) + X_{i-2}$
$H_{i-1} = h(X_{i-1} \| kh_{i-1,i} \| t_{i-1})$
$\mathrm{Report}_{i-1} = \{SID_{i-1}, (X_{i-1}, H_{i-1}), t_{i-1}\}$
Report_{i-1}
Smi-1
...
Report_1
Sm1

Figure 2. Proposed multi-hop data aggregation scheme.

B. Multi-hop Data Aggregation

To maintain balance between power production and demand, the PS periodically (say, every 15 or 30 minutes) needs to know the electricity consumption of the group of *n* HANs. To do so, for the interval T_j, PS calculates $r_i = IHF(T_j||k_i)$ and $R_{Sum} = \sum_{i=1}^{n} r_i$, and sends R_{Sum} to the TPA, where IHF denotes the integer hash function [20-21]. Here, T_j includes both the date and the time interval. Thus, the value of r_i will be different for each time interval. Next, similar to the PS, for the interval T_j, each smart meter SM_i calculates $r_i = IHF(T_j||k_i)$ and subsequently generates a timestamp t_i and computes its blinded measurement $X_i = (M_i + r_i) + X_{i-1}$ and $H_i = h(X_i||kh_{i,i+1}||t_i)$, where X_{i-1} denotes the blinded measurement received from its neighbor SM_{i-1}. Next, SM_i composes $\mathrm{Report}_i = \{SID_i, X_i, H_i, t_i\}$ and sends it to its parent SM_{i+1}. Upon receipt of Report_i, smart meter SM_{i+1} first validates the time stamp t_i and also computes and validates H_i using the secret key $kh_{i,i+1}$. If the validation is successful, SM_{i+1} generates a timestamp t_{i+1} and calculates the blinded measurement $X_{i+1} = (M_{i+1} + r_{i+1}) + X_i$ and $H_{i+1} = h(X_{i+1}||kh_{i+1,i+2}||t_{i+1})$. Finally, smart meter SM_{i+1} composes $\mathrm{Report}_{i+1} = \{SID_{i+1}, X_{i+1}, H_{i+1}, t_{i+1}\}$ and sends it to its parent SM_{i+2}. Continuing in this way, upon receipt of usage report $\mathrm{Report}_n = \{SID_n, X_n, H_n, t_n\}$ from SM_n, the TPA first checks its validity. If it is valid, the TPA calculates $X_n - R_{Sum}$ to obtain the *aggregated usage* data of the n HANs.

Next, the TPA sends the aggregated usage data to the PS. If the PS finds a mismatch between the energy production and consumption, it takes the necessary steps to increase production. In addition, the PS may employ demand-side management in order to modulate consumer behavior. Towards this end, the PS first composes the instructions or information for demand-side management (denoted by Γ) and conveys it to the TPA for dissemination to the consumers. Next, the TPA generates a valid signature $f = \textbf{Sign}(h(\Gamma, t), \text{SK-TPA})$ and subsequently broadcasts the instructions and the signature (Γ, f, t) to all the SMs. When a SM receives the usage instructions, it first checks the timestamp t and checks $\textbf{Ver}(h(\Gamma, t), f, \text{PK-TPA})$. If they are valid, then the SM informs its owner to adjust their usage; otherwise, it just ignores the instruction and signature. Here, **Sign** and **Ver** denote the signing algorithm and verification algorithm of a secure public key signature scheme [27-28].

It should be noted that for the correctness of our protocol, all the smart meters need to participate during the data aggregation process. To avoid the failure report problem (i.e., the absence of reports when a smart meter fails), each smart meter needs to do ping tests with its neighbors at regular intervals. In case smart meter SM_i does not receive any response from its neighbor SM_{i-1}, then SM_i informs the PS with the help of the TPA through multi-hop communication. The PS then abstains from creating any r_i for that particular smart meter and initiates technical support steps to resolve the issue.

V. Discussion

In this section, we first analyze the proposed scheme with respect to the security goals listed in Section II. Subsequently, we demonstrate that our proposed solution incurs reasonable computational overhead that is acceptable even for the resource constrained entities in smart-grids.

A. Security Analysis

- **Message Authentication:** In the proposed multi-hop data aggregation scheme, when smart meter SM_i receives the usage report from its neighbor SM_{i-1}, then SM_i first checks the timestamp t_{i-1}. If the timestamp is valid then SM_i computes $H^*_{i-1} = h(X_{i-1}||kh_{i-1,i}||t_{i-1})$ and verifies whether H^*_{i-1} is equal to H_{i-1} or not. If they are equal, then SM_{i-1} passes the authentication process. In this way, all the smart meters authenticate their neighbor before aggregation and finally when the TPA receives the usage report from its neighbor SM_n, the TPA computes $H^*_n = h(X_{i-1}||kh_{n,tpa}||t_n)$ and checks whether H^*_n is equal to H_n or not. If so, the TPA calculates $X_n - R_{Sum}$ and obtains the *aggregated usage* data of the n HANs. In this way, the proposed multi-hop data aggregation scheme ensures the message authentication property. Furthermore, in the proposed scheme, if an adversary tries to perform a replay attack, the receiving end can easily comprehend such activities by using the timestamps. Therefore, our proposed scheme is also secure against replay attacks.
- **Usage Data Confidentiality:** The amount of electricity usage in each HAN (e.g. HAN_i) is blinded with a random integer (r_i), which is generated from a secure Integer Hash Function. Hence, when the neighbor aggregator (HAN_{i+1}) receives the usage report Report_i, it can only see the blinded measurement of the HAN. Similarly, when the TPA calculates $X_n - R_{Sum}$, it can only know the summation of the usage data of a group of HANs. Since each element of $R_n = \{r_1, r_2, \cdots, r_n\}$ is random, even if two consecutive readings from a HAN are the same, an adversary (even the TPA) cannot comprehend that from the blinded measurements. Thus, the pattern of the electricity consumption is protected from detection by any eavesdropper.
- **Usage Data Integrity:** In the proposed scheme, when a smart meter SM_i receives the usage report from its neighbor SM_{i-1}, then it checks whether the message it has received is the same as that sent by SM_{i-1}. In this regard, SM_i computes $H^*_{i-1} = h(X_{i-1}||kh_{i-1,i}||t_{i-1})$ by using the shared secret key $kh_{i-1,i}$ and then verifies whether H^*_{i-1} is equal to H_{i-1} or not. This approach helps to detect any manipulation of the aggregated usage data.

B. Performance Evaluation

Now, we compare the aggregation time of the schemes, which includes the computation and the report transmission time. In [13], all SMs need to do homomorphic encryption, where the data encryption computation of the SMs can be calculated in parallel. We assume that the number of SMs is n and the communication time between any two neighbor SMs (denoted by T_c) is fixed and the height of the aggregation tree is no less than $\log_q n$, where q denotes the number of children of each node in the tree. Based on these assumptions, the multi-hop data aggregation scheme presented in [13] has an aggregation time of $T(q) = T_h + T_c \log_q n$, where T_h denotes the homomorphic computation time. In [15], a smart meter needs to decrypt all the reports received from its children. Therefore, the multi-hop data aggregation scheme presented in [15] has an aggregation time of $T(q) = qT_{dec} + T_c \log_q n$. On the other hand, the aggregation time in the proposed multi-hop data aggregation scheme is $T(q) = T_{dec} + (qT_a + T_c) \log_q n$, where T_{dec} denotes the AES decryption computation time in Step AG1, and T_a denotes the message authentication and hash operation time.

Next, for analyzing the performance of the proposed scheme with respect to [13] and [15], we conducted simulations of the cryptographic operations in all the schemes on an Ubuntu 12.04 virtual machine with an Intel Core i5-4300 dual-core 2.60 GHz CPU (operating as the TPA or the SP as per the scheme). To simulate a smart meter, we used a single core 798 MHz CPU and 256 MB of RAM, which is similar to the computational capability of real SMs [26]. The simulations used the JCE library [23] and the Pailler library libpaillier-0.8 [24] to evaluate the execution time of different cryptographic operations used in the proposed scheme, [13] and [15]. The form of the usage report of SM_i is $\{PID_i, X_i, H_i, t_i\}$ whose size is taken to be 576 bits. Based on the simulations, the mean values of T_h, T_a, and T_{dec} are 22.69 ms, 0.0167 ms, and 0.021 ms, respectively. Next, for communication cost we consider the data rate of WIMAX presented in [25]. We set the communication rate to 2 Mbps and the maximum number of children for each SM as 3. With these parameters, the proposed scheme has an aggregation time of 5.32 ms while the schemes presented in [13] and [15] have an aggregation time of 11.58 ms and 7.96 ms, respectively. The variation in the aggregation time as a function of the number of SMs for the proposed scheme, [13], and [15] is shown in Fig. 4. From Fig. 4, we can see that the aggregation time of the proposed scheme is lower as compared to [13] and [15]. Overall, we argue that the performance of the proposed hop-by-hop data aggregation scheme is better than that of [13] and [15], and hence it is more suitable for smart grid security.

VI. Conclusion

In this paper we developed a novel *multi-hop* data aggregation scheme for smart grids. The usage reports are aggregated according to the aggregation tree. Security analysis shows that the proposed scheme satisfies all the desired requirements. Computation and communication analyses show that the proposed scheme has better performance than existing multi-hop data aggregation schemes. Therefore, we argue that the proposed scheme is efficient, practical, and more suitable for

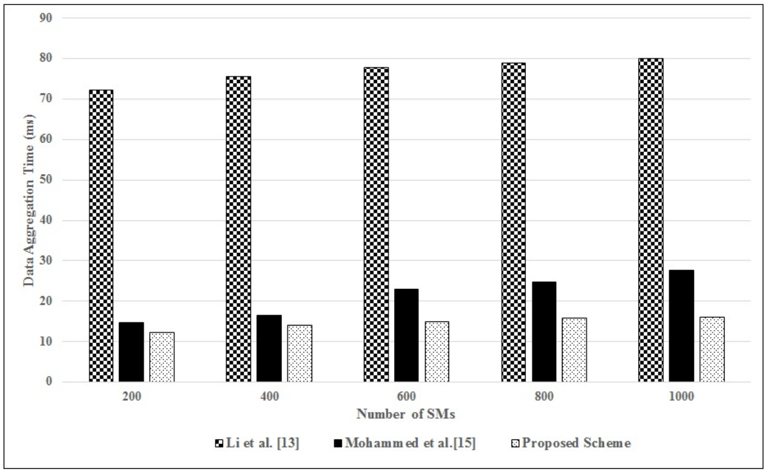

Figure 3. Variation of aggregation time in terms of number of SMs.

real-time requirements than other similar approaches for smart grid security.

Acknowledgment

This research was supported by the National Research Foundation, Prime Minister’s Office, Singapore under its Corporate Laboratory@University Scheme,

References

[1] Y. Li, X. Cheng and Y.Cao, “Smart Choice for the Smart Grid: Narrowband Internet of Things (NB-IoT),” *IEEE Internet of Things Journal*, DOI: 10.1109/JIOT.2017.2781251, 2017.

[2] Z. Guan et al., “Achieving Efficient and Secure Data Acquisition for Cloud-Supported Internet of Things in Smart Grid,” *IEEE Internet of Things Journal*, vol. 4(6), pp. 1934-1944, December 2017.

[3] U.S. NIST, “Guidelines for smart grid cyber security,” NIST IR-7628, Aug. 2010, available at: http://csrc.nist.gov/publications/PubsNISTIRs.html#NIST-IR-7628.

[4] R. Lu, X. Liang, X.L Li, and X. Shen, “Eppa: an efficient and privacy-preserving aggregation scheme for secure smart grid communications,” *IEEE Trans. Parallel Distrib. Syst.* vol. 23(9), pp. 1621–1631, 2012.

[5] P. Paillier, “Public-key cryptosystems based on composite degree residuosity classes,” *in Proc. EUROCRYPT,* pp. 223–228, Prague, Czech Republic, 1999.

[6] X. Liang, X. Li, R. Lu, X. Lin and X. Shen, “UDP: usage based dynamic pricing with privacy preservation for smart grid,” *IEEE Transactions on Smart Grid*, vol. 4(1), pp. 141-150, March 2013.

[7] M. Naehrig, K. Lauter and V. Vaikuntanathan, “Can homomorphic encryption be practical?,” *In Proc. the 3rd ACM Cloud Computing Security Workshop*, pp. 113-124, 2011.

[8] Y. Chia-Mu, C.-Y. Chen, S.-Y. Kuo, H.-C. Chao, “Privacy-preserving power request in smart grid networks,” *IEEE Syst. J.* vol. 8(2), pp. 441–449, 2014.

[9] X. Liu, Y. Zhang, B. Wang, H. Wang, “An anonymous data aggregation scheme for smart grid systems,”*Secur. Commun. Netw.* vol. 7(3), pp. 602–610, 2014.

[10] J. Zhang, L. Liu, Y. Cui, Z. Chen, “SP 2 DAS: self-certified PKC-based privacy-preserving data aggregation scheme in smart grid,”. *Int. J. Distrib. Sens. Netw*. 2013, 1–11.

[11] Z. Sui, A. Alyousef, H. de Meer, “IAA: incentive-based anonymous authentication scheme in smart grids,” *In: Tiropanis, T.*, Vakali, A., Sartori, L., Burnap, P. (eds.) Internet Science. LNCS, vol. 9089, pp. 133–144. Springer, Heidelberg (2015)

[12] K. Kursawe, G. Danezis, and M. Kohlweiss, “Privacy-friendly aggregation for the smart grid,”*in Proc. Privacy Enhanced Technology Symposium*, pp. 175–191, 2011.

[13] F. Li, B. Luo, P. Liu, “Secure information aggregation for smart grids using homomorphic encryption,”*in First IEEE International Conference on Smart Grid Communications (SmartGridComm)*, Gaithersburg, USA, 4–6 October, pp. 327–332. IEEE (2010).

[14] F. Knirsch et al. “Error-resilient Masking Approaches for Privacy Preserving Data Aggregation,” *IEEE Trans. Smart Grid,* DOI 10.1109/TSG.2016.2630803, 2016.

[15] H. Mohammed et al. ""Efficient Privacy-Preserving Data Collection Scheme for Smart Grid AMI Networks"," *in Proc. IEEE GLOBECOM*, Washington DC, USA, 2016.

[16] X. Dong, J. Zhou, K. Alharbi, X. Lin and Z. Cao, "An ElGamal-based efficient and privacy-preserving data aggregation scheme for smart grid," *2014 IEEE Global Communications Conference*, Austin, TX, 2014, pp. 4720-4725. doi: 10.1109/GLOCOM.2014.7037553.

[17] C. Castelluccia, A.C.-F. Chan, E. Mykletun and G. Tsudik, “Efficient and provably secure aggregatioin of encrypted data in wireless sensor networks, ” *ACM Trans. Sen. Netw.* vol. 5(20): 1-36, 2009.

[18] Nico Saputro, Kemal Akkaya, “Performance evaluation of Smart Grid data aggregation via homomorphic encryption,” *IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 2945-2950, 2012, ISSN 1525-3511.

[19] J. Won, Chris Y. T. Ma, David K. Y. Yau, Nageswara S. V. Rao, “Privacy-Assured Aggregation Protocol for Smart Metering: A Proactive Fault-Tolerant Approach,” *IEEE/ACM Transactions on Networking*, vol. 24, pp. 1661-1674, 2016, ISSN 1063-6692.

[20] http://web.archive.org/web/20071223173210 /http://www.concentric.net/ Ttwang/tech/inthash.htm

[21] http://burtleburtle.net/bob/hash/integer.html

[22] Diffie, W., Hellman, M.E.“New directions in cryptography,” *IEEE Trans. Inf. Theory* 22(6), 644–654 (1976). IEEE.

[23] Oracle Technology Network. Java Cryptography Architecture (JCA). [Online]. Available: http://docs.oracle.com/javase/6/docs/technotes/guides/crypto/CrypoSpec.html, accessed Apr. 20, 2017.

[24] libpaillier-0.8. Tech. rep. http://hms.isi.jhu.edu/acsc/libpaillier/ (accessed on 16 April 2017).

[25] V. C. Gungor, D. Sahin, T. Kocak, S. Ergut, S., C. Buccella, C. Cecati, G.P. Hancke, “Smart grid technologies: communication technologies and standards,” *IEEE Trans. Industr. Inf.* vol. 7(4), 529–539 (2011)

[26] Atmel’s family of smart power meters. http://www.atmel.com/products/smart-energy/power-metering/ (accessed on 28 May 2017).

[27] R. L. Rivest, A. Shamir, L. Adleman, “A Method for Obtaining Digital Signatures and Public-Key Cryptosystems,”*Comm. ACM,* vol. 21(2) pp. 120-126, 1978.

[28] G. J. Popek, C. S. Kline “ncryption Protocols Public-Key Algorithms and Digital Signatures in Computer Networks,”*in Foundations of Secure Computation,* pp. 133–153, 1999.

[29] Y. Wu, S. Fahmy and N. Shroff, “On the construction of a maximum-lifetime data gathering tree in sensor networks: NP-completeness and ap-

proximation algorithm," *Proc. IEEE INFOCOM,* pp. 356-360, Phoenix, AZ, April 2008.

[30] H. Tan, I. Korpeoglu and I. Stojmenovi, "Computing localized power-efficient data aggregation trees for sensor networks," *IEEE Transactions on Parallel and Distributed Systems,* vol. 22, no. 3, pp. 489-500, March 2011.