Hybrid Secure Data Aggregation in Wireless Sensor Networks

Keyur Parmar(✉) and Devesh C. Jinwala

S.V. National Institute of Technology, Surat, India
{keyur.mtech,dcjinwala}@gmail.com

Abstract. Secure data aggregation aims at combining security and data aggregation together to meet the requirements of data-centric networks such as wireless sensor network. Secure data aggregation protocols provide either hop-by-hop security or end-to-end security. However, hop-by-hop secure data aggregation is vulnerable to attackers at intermediate nodes while end-to-end secure data aggregation increases the communication overhead. In this paper, we propose a hybrid secure data aggregation protocol to balance the trade-off between privacy and communication overhead. The proposed protocol uses the symmetric-key based privacy homomorphism to ensure the privacy of sensor readings at intermediate nodes. In addition, the proposed protocol efficiently deals with the key management issues that exist in the state-of-the-art symmetric-key based protocols. The proposed protocol also reduces the communication overhead as compared to the existing end-to-end secure data aggregation protocols. Comprehensive analysis and comparisons validate the viability of the proposed protocol in resource-constrained wireless sensor networks.

Keywords: Wireless sensor networks · Security · Secure data aggregation · Privacy homomorphism · Communication overhead

1 Introduction

Wireless sensor network (WSN), a collection of tiny and cost-effective sensor devices, has envisioned many applications such as battlefield surveillance, target tracking, environmental & health care monitoring and traffic regulation. [1]. These tiny sensor devices have very limited resources such as memory, processor, energy and bandwidth [1]. Amongst these resources, energy is the most limiting factor that has a profound impact on the WSNs' lifetime [6]. Therefore, the major objective of WSNs' protocols is to reduce the energy consumption. In addition, as communication operations in WSNs consume significantly more energy than computation operations [11], WSNs' protocols aim at reducing communication overhead. One of the techniques used for reducing the communication overhead is "In-network data aggregation" [6]. In-network data aggregation processes the raw sensor readings at intermediate nodes, and forwards the aggregated result towards the base station. Along with data aggregation, data security becomes

C. Berger and M.R. Mousavi (Eds.): CyPhy 2015, LNCS 9361, pp. 116–131, 2015.
DOI: 10.1007/978-3-319-25141-7_9

an important design parameter due to hostile and unattended deployments and unreliable communication channel [4,25].

The requirement to bind the security and data aggregation together leads the development of secure data aggregation protocols. Secure data aggregation protocols have been classified as either hop-by-hop secure data aggregation protocols or end-to-end secure data aggregation protocols. Hop-by-hop secure data aggregation protocols [12,18] assume that intermediate nodes are trustworthy. Hence, data forwarded by leaf nodes can be decrypted at intermediate nodes. Intermediate nodes perform the aggregation operations over raw sensor readings and encrypt the aggregated data before forwarding the result towards the next hop. Though viable, such hop-by-hop aggregation becomes problematic when intermediate nodes are not trustworthy. Malicious intermediate nodes can read and modify the sensor readings that eventually violate the privacy and confidentiality of sensor readings. In addition, hop-by-hop secure data aggregation also incurs extra computation overhead at intermediate nodes. Intermediate nodes have to decrypt the sensor readings, perform the aggregation, and re-encrypt the aggregated data before forwarding it to the parent nodes. Hence, with the aim to protect the privacy of sensor readings, and to reduce the computation overhead at intermediate nodes, Girao et al. [8] proposed the end-to-end secure data aggregation (also known as concealed data aggregation).

End-to-end secure data aggregation protocols [8,17,18,20] process the encrypted data at intermediate nodes. End-to-end secure data aggregation uses privacy homomorphism [23] to support encrypted data processing. End-to-end secure data aggregation can be classified in three categories: (1) Symmetric-key based end-to-end secure data aggregation [3,8] (2) Asymmetric-key based end-to-end secure data aggregation [17], and (3) Elliptic curve cryptography (ECC) based end-to-end secure data aggregation [7,19]. Amongst these protocols, asymmetric-key/ECC based protocols are not viable for resource-constrained sensor devices due to their high computation and communication overhead [12]. In addition, symmetric-key cryptosystems, such as SKIPJACK, with 80-bit key size can provide the same level of security as compared to the asymmetric-key based cryptosystems such as the RSA [24] with 1024-bit key size [10]. However, there exist numerous research articles [17,18,22] that use asymmetric-key/ECC based cryptosystems in WSNs. The only reason to pursue the costly asymmetric-key/ECC based cryptosystems is the key management issues related to the symmetric-key based cryptosystems.

Symmetric-key based cryptosystems use a shared secret key at both ends of the communication channel. Moreover, if data are encrypted with different pairwise keys in symmetric-key based protocols, such as Domingo-Ferrer's cryptosystem [5], aggregator nodes cannot perform the in-network processing. Hence in order to perform the in-network processing of encrypted data, the global shared secret key needs to be distributed throughout the network. In WSNs where the deployment of nodes is in hostile environments, such a global shared secret key mechanism has a devastating effect on the overall aggregated result. The only symmetric-key based cryptosystem that does not require the global shared secret

key across all nodes is Castelluccia et al.'s cryptosystem (The CMT cryptosystem) [2,3]. Although the CMT cryptosystem [2,3] mitigates the key management issues typically found in other symmetric-key based cryptosystems [8,21,26], it has an identity management issue. Each node in the CMT cryptosystem shares a unique secret key with the base station. Hence, if there exist non-responding nodes in the network, then the identities of non-responding nodes need to be forwarded towards the base station. The identity of a node is used to uniquely identify the node and to find the secret key it shares with the base station. As the identity-related information cannot be aggregated in the same way as the sensor readings, transmission of the identities of nodes increases the significant communication overhead.

In this paper, we propose a hybrid secure data aggregation protocol to deal with the challenges typically found in existing hop-by-hop as well as end-to-end secure data aggregation protocols. The proposed protocol uses the symmetric-key based privacy homomorphism to ensure the privacy of sensor readings nearer to the base station. In addition, the proposed protocol attempts to balance the trade-off between privacy and communication overhead. Finally, we compare the proposed protocol with existing hop-by-hop and end-to-end secure data aggregation protocols and validate the viability of the proposed protocol in resource-constrained WSNs.

The rest of the paper is organized as follows. In Sect. 2, we discuss the relevant literature. In Sect. 3, we provide a brief overview of the symmetric-key based CMT cryptosystem. Section 4 presents the proposed protocol for hybrid secure data aggregation. We analyze the resource overhead of the proposed protocol in Sect. 5. In Sect. 6, we analyze the security strength of the proposed protocol. Section 7 concludes the paper by emphasizing our contributions.

2 Related Work

Although security and data aggregation are vital design parameters for WSNs' protocols, their objectives are contradictory. Data aggregation protocols aim at reducing the communication traffic while security protocols increase the communication traffic. The need to provide security and data aggregation together has initiated secure data aggregation. Initial secure data aggregation protocols [12,18] aim at providing security in a hop-by-hop manner. Hu et al. [12] proposed a secure data aggregation protocol that ensures hop-by-hop security and en route data aggregation. Although, there have been numerous solutions [12,18] that provides hop-by-hop security, all of them consider the trustworthy intermediate nodes. Hence, intermediate aggregator nodes that contain a large volume of information gathered from their child nodes become the prime target for attackers.

Girao et al. [8] proposed a concealed data aggregation to protect the sensor readings at intermediate nodes. Authors used the privacy homomorphism introduced by Rivest et al. [23] to perform the encrypted data processing. Privacy

homomorphism can be classified as either symmetric-key based [3,21], asymmetric based key [5,24] or ECC based [13] privacy homomorphism. Asymmetric-key based and ECC based privacy homomorphism have much higher resource consumption than the symmetric-key based privacy homomorphism [9,15]. Although, Gura et al. [9] and Malan et al. [15] argue in favor of asymmetric-key based and ECC based cryptosystems, their results clearly show that the resource overhead of their protocols are significantly higher than the protocols based on symmetric-key based cryptosystems [3,8].

The reason to pursue asymmetric-key based or ECC based privacy homomorphism in sensor networks is due to the key management issue of symmetric-key based techniques, such as [8,21,26], or the identity management issue of symmetric-key based techniques such as [2,3,21]. Mlaih et al. [16] proposed the protocol to combine the hop-by-hop and end-to-end secure data aggregation to provide the flexibility during aggregation and optimal data privacy at intermediate nodes. However, their proposed protocol requires the identity transfer similar to the one required by the CMT cryptosystem. Hence, the communication cost incurred by identity transfer remains as high as the CMT cryptosystem.

3 Preliminaries

In this section, we briefly discuss Castelluccia et al.'s [2,3] stream cipher based additively homomorphic cryptosystem (The CMT cryptosystem). As the CMT cryptosystem supports additive homomorphism, it can be used to compute the mean, variance and standard deviation over encrypted data without performing any decryption. The majority of sensor network's applications require computing an optimum value, such as the sum, minimum, maximum, variance, movement detection, etc., over data [2,26]. Such operations can be easily computed with the help of the additively homomorphic CMT cryptosystem.

The CMT Cryptosystem

Encryption $\mathcal{E}$:

1. Represent a plaintext m as an integer $m \in [0, M-1]$, where M is the modulus.
2. Let k be a pseudo random number such that $k \in [0, M-1]$.
3. Compute the ciphertext, $c = \mathcal{E}_k(m) = m + k \mod M$.

Decryption $\mathcal{D}$:

1. Decrypt the ciphertext, $\mathcal{D}_K(c) = c - k \mod M = m$.

Ciphertexts Aggregation $\mathcal{A}$:

1. Given $c_1 = \mathcal{E}_{k_1}(m_1)$ and $c_2 = \mathcal{E}_{k_2}(m_2)$.
2. Compute an aggregated ciphertext, $C = c_1 + c_2 \mod M = \mathcal{E}_K(m_1 + m_2)$ where $K = k_1 + k_2 \mod M$.

To ensure the correctness, modulus M should be sufficiently larger than the sum of individual messages such as $M > \sum_{i=1}^{n}(m_i)$. If modulus M is smaller than the sum of aggregated messages, the correctness does not hold.

4 The Proposed Hybrid Secure Data Aggregation

The major advantage of using end-to-end secure data aggregation over hop-by-hop secure data aggregation is due to the privacy protection of sensor readings at intermediate aggregator nodes. However, end-to-end secure data aggregation protocols, such as the CMT cryptosystem, increase the communication overhead. As shown by Castelluccia et al. [2,3], the requirement of transmitting node identities (responding or non-responding, whichever is less) to the base station increases the communication overhead that in turn depletes the energy of sensor nodes. Due to the identity transfer of non-responding nodes, the communication overhead of the CMT cryptosystem remains nearly same as any non-aggregation based protocols when there exist a high number of non-responding nodes. In no-aggregation based approaches, data are not aggregated at intermediate nodes while in the CMT cryptosystem, the identities of nodes are not aggregated.

In this section, we present the proposed hybrid secure data aggregation protocol. The proposed protocol reduces the communication overhead and protects the privacy of sensor readings at aggregator nodes nearer to the base station. In Table 1, we describe the notations used in the proposed protocol.

Table 1. Notations used in the proposed protocol

Symbol	Description
i	A sensor node (leaf node) ID
j	An intermediate node (Aggregator)
$k_{i,j}$	A pair-wise secret key between a node i and its parent node j
m	The distance between a node i and a node j (in hops)
$k_{m'}$	A shared secret key between an intermediate node at m^{th} hop and the base station
S_i	A plaintext value sensed by a sensor node i
c_i	A ciphertext generated by a node i
S_j	An aggregated plaintext value generated by an aggregator node j
c_j	A ciphertext generated by an aggregator node j
$S_{m'}$	An aggregated plaintext value generated by an aggregator node at m^{th}-hop
$c_{m'}$	A ciphertext generated by an aggregator node at m^{th}-hop
$c_{m''}$	An aggregated ciphertext
h	A height of the data aggregation tree for a node at level m
p_m	An identity of a non-responding node, stored at its parent node at $(m+h)^{\text{th}}$-hop
$p_{m''}$	A product of non-responding nodes' identities
C	An aggregated ciphertext received at the base station
S	An aggregated plaintext retrieved at the base station

The Proposed Protocol for Hybrid Secure Data Aggregation

Upto m^{th} - hop

Encryption:

1. Each leaf node i, encrypts a sensor reading S_i using a shared secret key $k_{i,j}$ of a node i and its parent node j.
2. Compute $c_i = Enc_{k_{i,j}}(S_i)$.

Decryption:

1. Each parent node j, decrypts a ciphertext c_i using a shared secret key $k_{i,j}$ of a node j and its child node i.
2. Compute $S_i = Dec_{k_{i,j}}(c_i)$.

Plaintext Aggregation:

1. Each parent node j, aggregates the decrypted data S_i for all $i \in [1..n]$
2. Compute $S_j = \sum_{i=1}^{n} S_i$.
3. Node j re-encrypts the aggregated sensor readings S_j, using a shared secret key $k_{j,m'}$ of a node j and its parent node m'.
4. Compute $c_j = Enc_{k_{j,m'}}(S_j)$.

At m^{th} hop

Decryption:

1. Each node m' at m^{th} hop decrypts the ciphertext c_j using a shared secret key $k_{j,m'}$ of a node m' and its child node j.
2. Compute $S_j = Dec_{k_{j,m'}}(c_j)$.

Plaintext Aggregation:

1. Each node m' aggregates the decrypted data.
2. Compute $S_{m'} = \sum_{j=1}^{n} S_j$.

Encryption:

1. Each node m', encrypts an aggregated data $S_{m'}$ using a shared secret key $k_{m'}$.
2. $c_{m'} = Enc_{k_{m'}}(S_{m'})$.

m^{th} - hop onwards

Ciphertext Aggregation:

1. Each node m'' at $(m+h)$th hop, where $h \in [1...x]$, aggregates the encrypted data $c_{m'}$. Here, x is a height of a data aggregation tree.
2. $c_{m''} = \sum_{m'=1}^{n}(c_{m'})$.

Key Management:

1. Each node m'' at $(m+h)$th hop, where $h \in [1...x]$, computes the product of its child nodes' identities p_m that do not responded during the aggregation process.
2. $p_{m''} = \prod_{m=1}^{n}(p_m)$.

At the Base Station

Key Management:

1. The base station uniquely identifies the non-responding/responding nodes using the received product of primes $p_{m''}$.
2. $p_{m''} = \prod_{m=1}^{n}(p_m)$.
3. The base station uses these primes to uniquely identify the nodes at m^{th} hop and their respective keys $k_{m'}$.

Decryption:

1. The base station removes the shared secret keys $k_{m'}$ where $m' \in [1..n]$, of nodes at m^{th} hops, who responded during the aggregation process.
2. Compute $Dec(C) = c_{m''} - \sum_{m'=1}^{n} k_{m'} = \sum_{i=1}^{n} S_i$. Here, $n \in [1....x]$ represents the sensor nodes at m^{th} hops, that provided the sensor readings S_i.

5 Overhead Analysis

In this section, we comparatively evaluate the performance of the proposed protocol with hop-by-hop secure data aggregation and end-to-end secure data aggregation (The CMT cryptosystem) scenarios. For ease of calculation, we assume a ternary tree-based data aggregation topology in which the packets are forwarded from leaf nodes towards the base station. However, the proposed protocol can be seamlessly adopted for other network topologies such as, a cluster-based network topology or a hybrid network topology. In addition, we consider a network model similar to the one found in Castelluccia et al. [2] to calculate the bandwidth consumption.

5.1 Network Model

Let us assume a balanced k-ary tree with a sink node and multitude of sensor nodes. In Fig. 1, we present a ternary tree where $k = 3$. In addition, for the ease of calculation, we assume that leaf nodes are sensor nodes and remaining intermediate nodes are forwarders. We also assume that the sensor reading, S_i, of node i, (e.g. Temperature ranges between 0 and 127 Fahrenheit) is 7-bit long.

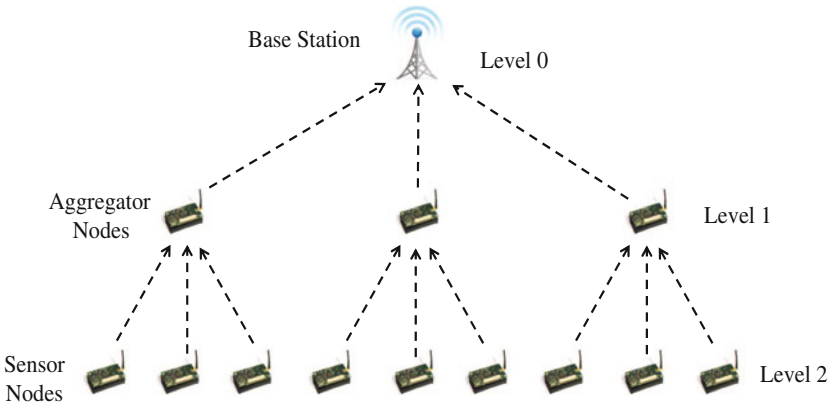

Fig. 1. Ternary tree-based data aggregation topology

We analyze the bandwidth consumption of nodes at various levels in the hierarchy by computing number of bits transmitted by the nodes. We consider a packet format used in TinyOS [14], an operating system for embedded devices, where the packet header (HDR) is of 56-bit and the maximum supported payload is 232 bits. We compare the proposed protocol with hop-by-hop secure data aggregation and end-to-end secure data aggregation scenarios. In addition, three scenarios considered for evaluation are as follows: (1) All nodes reply to their parent nodes (2) 10 % nodes are exhausted/compromised and do not reply (3) 30 % nodes are exhausted/compromised and do not reply. Here, we assume that non-responding nodes are distributed uniformly across the network.

5.2 Communication Overhead

In hop-by-hop secure data aggregation (SDA), a total number of bits transmitted by a node vary depending on the node's position in the hierarchy. As symmetric-key cryptosystems have negligible message expansion compared to asymmetric-key cryptosystems, a total number of bits transmitted by leaf nodes, remains the same as raw sensor readings, $log_2(t)$. Here, the total number of bits transmitted by leaf nodes in no aggregation based scenarios and hop-by-hop secure data aggregation scenarios remains nearly same. However, aggregator nodes in hop-by-hop secure data aggregation require to transmit a few more bits compared to the leaf nodes as they receive more data compared to the leaf nodes.

As shown in Fig. 2, the total number of bits transmitted by a leaf node in hop-by-hop encryption is HDR + $log_2(t)$, where t is the range of all possible sensor measurements. As shown above, temperature sensor requires 7-bit, $log_2(127)$, to represent 127 different temperature values. Hence, $56 + 7 = 63$-bit data are transmitted by each leaf node. Moreover, each intermediate node has to transmit $log_2(n' \cdot t)$ bits, where n' represents the aggregation of child nodes' sensor readings. Hop-by-hop encryption does not incur any additional communication overhead when there exist 10 % or 30 % non-responding nodes (NRN) in the network. In addition, the communication overhead reduces in these scenarios due to the less number of packets coming for aggregation.

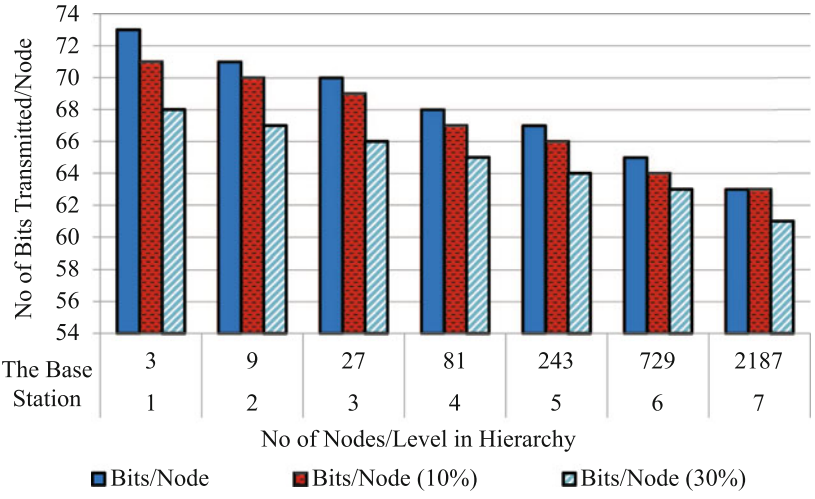

Fig. 2. Communication overhead of hop-by-hop SDA

If there aren't any non-responding nodes, the CMT cryptosystem (end-to-end secure data aggregation) performs nearly as good as hop-by-hop secure data aggregation protocols. In the CMT cryptosystem, a total number of bits transmitted by a node (leaf/intermediate node) remains constant due to en route

aggregation. The ciphertext size in the CMT cryptosystem depends on the modulus M. The total number of bits transmitted by the CMT cryptosystem is calculated as HDR $+log_2(n)+log_2(t)$, where n represents the total number of nodes in the network and t represents the range of sensor readings. Here, each sensor node has to transmit 56 + 12 + 7 = 75-bit. However, due to non-responding nodes' identity transfer, communication overhead increases drastically.

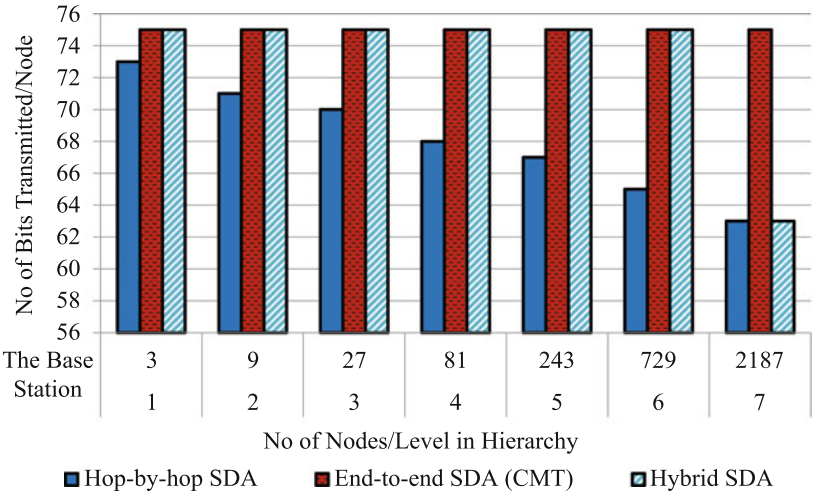

Fig. 3. Communication overhead of hybrid SDA ($m = 1$ and 0 % NRN)

As shown in Fig. 3, the proposed hybrid secure data aggregation protocol has the same communication overhead as a hop-by-hop secure data aggregation protocol between leaf nodes to m^{th} hop intermediate nodes. In addition, from m^{th} hop onwards, the communication overhead of the proposed protocol remains same as the CMT cryptosystem. Figure 4 comparatively evaluates the performance of the proposed protocol and presents the communication overhead when data are aggregated after 2^{nd} hop intermediate node, $m = 2$. In addition, when we compared the proposed hybrid secure data aggregation protocol with hop-by-hop secure data aggregation, it has negligible additional communication overhead. The proposed protocol does not require the extra computation overhead at each intermediate nodes as required by the hop-by-hop secure data aggregation protocols. In addition, unlike hop-by-hop secure data aggregation protocols, the proposed protocol ensures the privacy of sensor readings at intermediate nodes.

As shown in Fig. 5, when we choose $m = 1$ and if 10 % nodes are not responding to the aggregator nodes, the communication overhead of the proposed protocol is 2.2 times less compared to the end-to-end secure data aggregation protocol. In addition, the proposed protocol has 1.9 times more communication overhead compared to the hop-by-hop secure data aggregation protocol. However, the

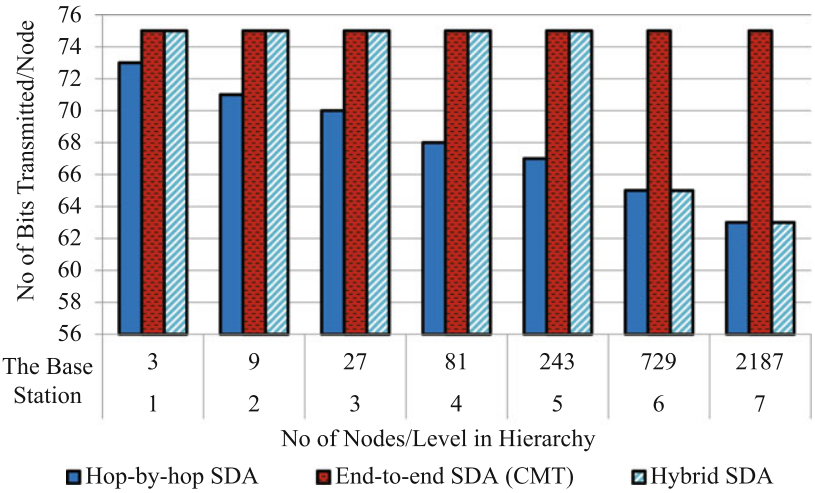

Fig. 4. Communication overhead of hybrid SDA ($m = 2$ and 0 % NRN)

increase in communication overhead is due to the much higher level of security at intermediate nodes in the network. The communication overhead of the proposed protocol remains same as the hop-by-hop secure data aggregation for the 1^{st} hop ($m = 1$), it reduces significantly compared to the CMT cryptosystem after 1^{st} hop onwards ($m = 1$). The communication overhead of the proposed protocol after m^{th} hop is HDR $+log_2(n)+log_2(t)+log_2(n'')$. Here, n'' represents the number of bits used to uniquely identify the child nodes of an intermediate node at $(m + 1)$th hop. Moreover, the proposed protocol ensures the privacy of sensor readings at intermediate nodes higher than 1^{st} hop intermediate nodes. In addition, for 30 % non-responding nodes (Fig. 6), the proposed protocol has nearly 2.9 times less communication overhead compared to the end-to-end secure data aggregation protocol.

As shown in Fig. 7, if we choose $m = 2$, the proposed protocol has 3.5 times less communication overhead compared to the end-to-end secure data aggregation protocol. In addition, the communication overhead of the proposed protocol is only 1.2 times more compared to the hop-by-hop secure data aggregation. For 30 % non-responding nodes at level $m = 2$ (Fig. 8), the proposed protocol has nearly 7 times less communication overhead compared to the CMT cryptosystem. In addition, the communication overhead of the proposed protocol is only 1.6 times more compared to the hop-by-hop secure data aggregation. The comparison of the proposed protocol with hop-by-hop secure data aggregation and end-to-end secure data aggregation protocol proves that the proposed hybrid secure data aggregation protocol reduces the significant communication overhead without affecting the privacy of a major part of the network.

The reason to pursue the hybrid secure data aggregation protocol is to reduce the communication overhead of symmetric-key based cryptosystem [3], or

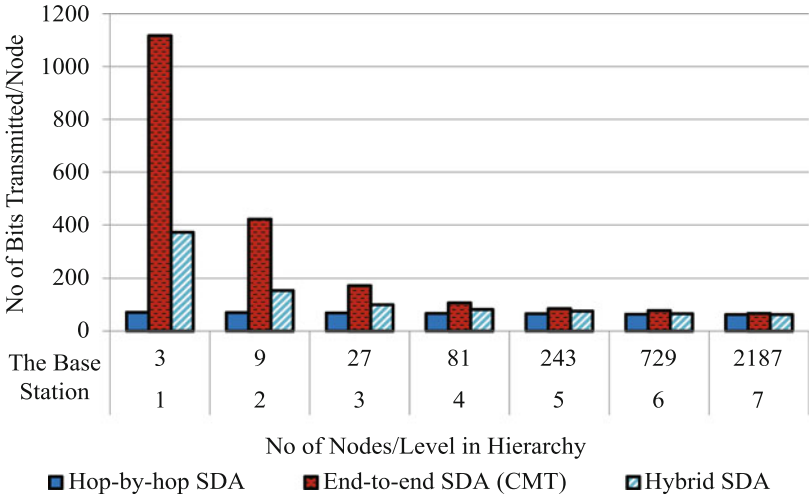

Fig. 5. Communication overhead of hybrid SDA ($m = 1$ and 10 % NRN)

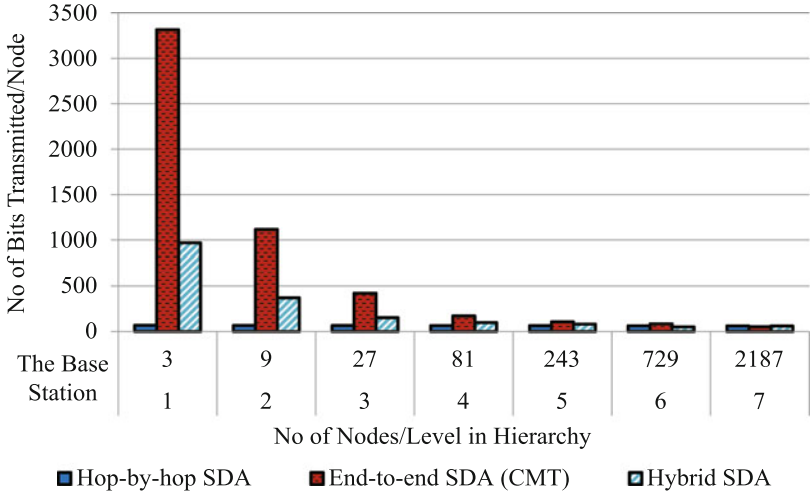

Fig. 6. Communication overhead of hybrid SDA ($m = 1$ and 30 % NRN)

asymmetric-key based or ECC based cryptosystems. In addition, without affecting the privacy of a network at a large, we can achieve a significant reduction in communication traffic. Hence, instead of having hop-by-hop or end-to-end secure data aggregation protocols, if we choose the hybrid secure data aggregation protocol, we can trade-off between communication overhead and privacy requirements for different applications, depending on the needs of applications.

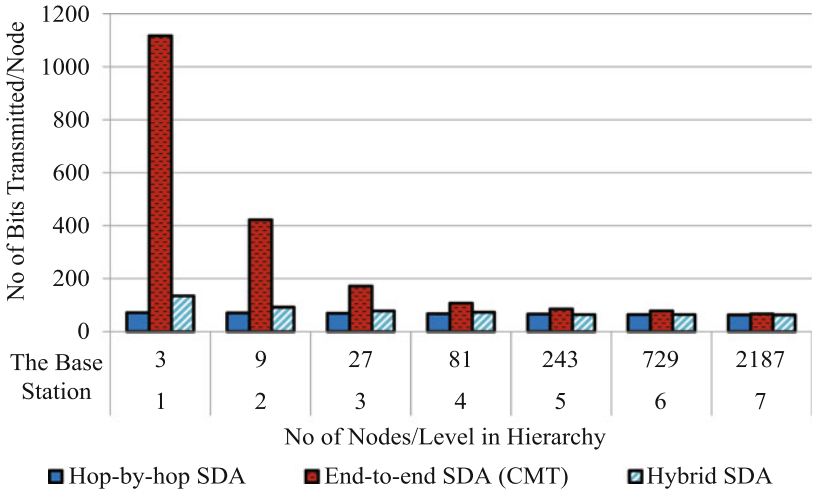

Fig. 7. Communication overhead of hybrid SDA ($m = 2$ and 10 % NRN)

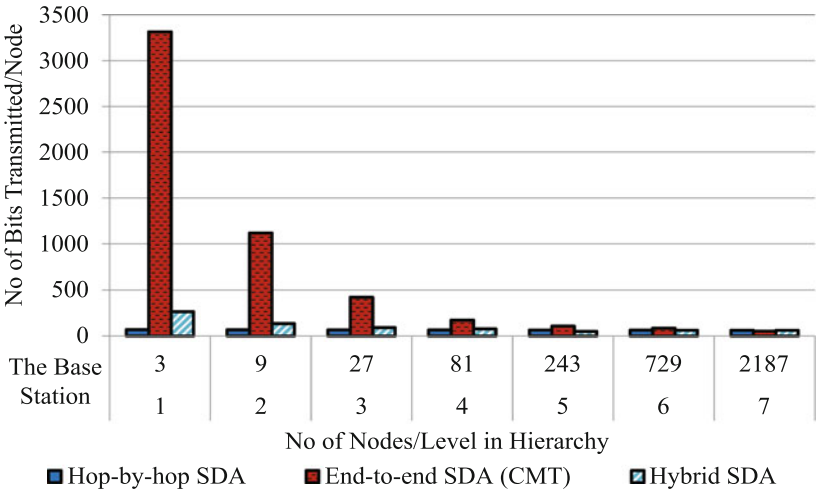

Fig. 8. Communication overhead of hybrid SDA ($m = 2$ and 30 % NRN)

Computation operations also consume sensor nodes' limited resources. However, energy consumption due to the CPU processing is negligible compared to the radio frequency operations. As shown by Hill et al. [11], transmitting a single bit over a meter range requires the same amount of energy as required by the 1000 CPU instructions. Hence, in this paper, we focus on the communication overhead only, and we neglect the computation overhead that is negligible [11] compared to the communication overhead.

6 Security Analysis

In this section, we discuss the resilience of the proposed protocol against well-known cryptographic attacks [22]. In addition, we analyze the security of the proposed protocol against active and passive adversaries.

Ciphertext Analysis. In the ciphertext analysis, an adversary eavesdrops the ciphertexts and analyzes them to infer knowledge about the corresponding plaintexts or the key(s). In the proposed protocol, two different symmetric-key based cryptosystems have been used to secure the communication. Symmetric-key based cryptosystems remain secure against ciphertext analysis as long as the key being shared remains secret. Here, the shared secret key needs to be updated periodically to thwart the ciphertext analysis attacks.

Known-Plaintext Attack. In a known-plaintext attack, an adversary has plaintext-ciphertext pairs and the objective is to recover the complete or partial information related to the secret key or the plaintext(s). As WSNs are deployed in hostile and unattended environments, for an adversary to capture the plaintext-ciphertext pairs becomes relatively easy. In the proposed protocol, we use two different symmetric-key based cryptosystems. The proposed protocol can use symmetric-key based cryptosystem, such as AES, DES, or Triple DES, for communication between the leaf nodes to the m^{th} hop nodes. Any node before m^{th} hop shares a unique secret key with its parent node. Hence, if an adversary gets the hold of the node's secret key, it cannot decrypt the ciphertexts produced by other sensor nodes. From m^{th} hop onwards, we use additively homomorphic symmetric-key based CMT cryptosystem. In the CMT cryptosystem, each node shares a unique pairwise key with the base station. As the CMT cryptosystem does not have a limitation to share a global shared secret key in order to perform en route aggregation, the proposed protocol remains secure against known-plaintext attacks.

Forge Packets. If an adversary can generate a valid ciphertext with a specific content, then it does not have to modify the existing ciphertexts. An attacker can easily insert the ciphertext into the network without being detected. Any asymmetric-key based cryptosystem, where the public key is used to generate the ciphertext, is vulnerable to this attack. However, in the proposed cryptosystem, we use symmetric-key based cryptosystems. Hence, an adversary must have to compromise a sensor node and extracts the shared secret key in order to generate a valid ciphertext.

Denial of Service Attacks. Amongst the various types of denial of service attacks, sensor networks are more prone to the attacks where the scarce energy is a target. In this attack, an adversary's goal is to keep sensor nodes busy doing activities that deplete sensor nodes' precious energy. In addition, there cannot

be any cryptographic solution for such attacks as the cryptographic solution also consumes the energy. In the proposed protocol, the communication overhead is significantly less compared to the end-to-end secure data aggregation. In addition, although it requires a little bit more communication overhead compared to the hop-by-hop secure data aggregation, the privacy preservation at intermediate nodes compensates that extra communication overhead. As the radio frequency operations has the highest impact on the energy consumption [11], the reduced communication overhead can significantly improve the performance against denial of service attacks.

7 Conclusions

Although data aggregation and security remain essential design parameters for secure data aggregation protocols, both of them have conflicting requirements. Data aggregation protocols lessen the communication overhead in order to reduce the energy consumption while the security protocols add extra communication overhead in order to ensure the security of sensor readings. Amongst secure data aggregation protocols, hop-by-hop secure data aggregation protocols ensure lesser communication overhead while end-to-end secure data aggregation protocols ensure the privacy of sensor readings at intermediate nodes. Hence, with the intent to reduce the communication overhead and to ensure the privacy of sensor readings at intermediate nodes, we proposed a hybrid secure data aggregation protocol. The proposed protocol balances the trade-off between privacy and communication overhead. It protects the privacy of sensor readings nearer to the base station where it is required the most. In addition, the proposed protocol ensures lesser communication overhead that eventually increases the lifespan of WSNs. As per our knowledge, the proposed protocol is the first that achieves advantages of both hop-by-hop secure data aggregation and end-to-end secure data aggregation. As future work, we intend to formalize the security analysis. In addition, we plan to implement the proposed protocol for measuring the impact of computation and communication operations on the energy of sensor devices.

Acknowledgments. This research was a part of the project "A Secure Data Aggregation System and An Intrusion Detection System for Wireless Sensor Networks". It was supported by the Department of Electronics and Information Technology, Ministry of Communications and Information Technology, Government of India.

References

1. Akyildiz, I.F., Su, W., Sankarasubramaniam, Y., Cayirci, E.: Wireless sensor networks: a survey. Comput. Netw. Int. J. Comput. Telecommun. Netw. **38**(4), 393–422 (2002)
2. Castelluccia, C., Chan, A.C.F., Mykletun, E., Tsudik, G.: Efficient and provably secure aggregation of encrypted data in wireless sensor networks. ACM Trans. Sens. Netw. (TOSN) **5**(3), 20:1–20:36 (2009)

3. Castelluccia, C., Mykletun, E., Tsudik, G.: Efficient aggregation of encrypted data in wireless sensor networks. In: Proceedings of the 2nd Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services, MOBIQUITOUS 2005, pp. 109–117. IEEE, Washington, D.C., July 2005
4. Chan, H., Perrig, A.: Security and privacy in sensor networks. Computer **36**(10), 103–105 (2003)
5. Domingo-Ferrer, J.: A provably secure additive and multiplicative privacy homomorphism. In: Chan, A.H., Gligor, V.D. (eds.) ISC 2002. LNCS, vol. 2433, pp. 471–483. Springer, Heidelberg (2002)
6. Fasolo, E., Rossi, M., Widmer, J., Zorzi, M.: In-network aggregation techniques for wireless sensor networks: a survey. Wirel. Commun. **14**(2), 70–87 (2007)
7. Girao, J., Westhoff, D., Mykletun, E., Araki, T.: TinyPEDS: tiny persistent encrypted data storage in asynchronous wireless sensor networks. Ad Hoc Netw. **5**(7), 1073–1089 (2007)
8. Girao, J., Westhoff, D., Schneider, M.: CDA: concealed data aggregation for reverse multicast traffic in wireless sensor networks. In: Proceedings of the 40th International Conference on Communications, ICC 2005, pp. 3044–3049. IEEE, Seoul, May 2005
9. Gura, N., Patel, A., Wander, A., Eberle, H., Shantz, S.C.: Comparing elliptic curve cryptography and RSA on 8-bit CPUs. In: Joye, M., Quisquater, J.-J. (eds.) CHES 2004. LNCS, vol. 3156, pp. 119–132. Springer, Heidelberg (2004)
10. Hankerson, D., Menezes, A.J., Vanstone, S.: Guide to Elliptic Curve Cryptography, 1st edn. Springer, Secaucus (2003)
11. Hill, J., Szewczyk, R., Woo, A., Hollar, S., Culler, D., Pister, K.: System architecture directions for networked sensors. ACM SIGPLAN Not. **35**(11), 93–104 (2000)
12. Hu, L., Evans, D.: Secure aggregation for wireless networks. In: Proceedings of the Symposium on Applications and the Internet Workshops, SAINT 2003, pp. 384–391. IEEE, Washington, D.C., January 2003
13. Koblitz, N.: Elliptic curve cryptosystems. Math. Comput. **48**(177), 203–209 (1987)
14. Levis, P., Madden, S., Polastre, J., Szewczyk, R., Whitehouse, K., Woo, A., Gay, D., Hill, J., Welsh, M., Brewer, E., Culler, D.: TinyOS: an operating system for sensor networks. In: Weber, W., Rabaey, J.M., Aarts, E. (eds.) Ambient Intelligence, pp. 115–148. Springer, Heidelberg (2005)
15. Malan, D.J., Welsh, M., Smith, M.D.: A public-key infrastructure for key distribution in TinyOS based on elliptic curve cryptography. In: Proceedings of the 1st Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks, SECON 2004, pp. 71–80. IEEE, Santa Clara, October 2004
16. Mlaih, E., Aly, S.A.: Secure hop-by-hop aggregation of end-to-end concealed data in wireless sensor networks. In: Proceedings of the 2nd IEEE Workshop on Mission Critical Networking in Conjunction with Infocom 2008, MCN 2008, pp. 1–6. IEEE, Phoenix, April 2008
17. Mykletun, E., Girao, J., Westhoff, D.: Public key based cryptoschemes for data concealment in wireless sensor networks. In: Proceedings of the IEEE International Conference on Communications, ICC 2006, pp. 2288–2295. IEEE, Istanbul, June 2006
18. Ozdemir, S., Xiao, Y.: Secure data aggregation in wireless sensor networks: a comprehensive overview. Comput. Netw. Int. J. Comput. Telecommun. Netw. **53**(12), 2022–2037 (2009)

19. Parmar, K., Jinwala, D.C.: Malleability resilient concealed data aggregation. In: Kermarrec, Y. (ed.) EUNICE 2014. LNCS, vol. 8846, pp. 160–172. Springer, Heidelberg (2014)
20. Parmar, K., Jinwala, D.C.: Symmetric-key based homomorphic primitives for end-to-end secure data aggregation in wireless sensor networks. J. Inf. Secur. **6**(1), 38–50 (2015)
21. Peter, S., Piotrowski, K., Langendoerfer, P.: On concealed data aggregation for WSNs. In: Proceedings of the 4th IEEE Consumer Communications Networking Conference, CCNC 2007, pp. 192–196. IEEE, Las Vegas, January 2007
22. Peter, S., Westhoff, D., Castelluccia, C.: A survey on the encryption of convergecast traffic with in-network processing. IEEE Trans. Dependable Secure Comput. **7**(1), 20–34 (2010)
23. Rivest, R.L., Adleman, L., Dertouzos, M.L.: On data banks and privacy homomorphisms. Found. Secure Comput. **4**(11), 169–180 (1978)
24. Rivest, R.L., Shamir, A., Adleman, L.: A method for obtaining digital signatures and public-key cryptosystems. Commun. ACM **21**(2), 120–126 (1978)
25. Wang, Y., Attebury, G., Ramamurthy, B.: A survey of security issues in wireless sensor networks. IEEE Commun. Surv. Tutor. **8**(2), 2–23 (2006)
26. Westhoff, D., Girao, J., Acharya, M.: Concealed data aggregation for reverse multicast traffic in sensor networks: encryption, key distribution, and routing adaptation. IEEE Trans. Mob. Comput. **5**(10), 1417–1431 (2006)