Computers, Materials & Continua CMC, vol.61, no.2, pp.583-599, 2019

Comprehensive Analysis of Secure Data Aggregation Scheme for Industrial Wireless Sensor Network

Weidong Fang[1], Wuxiong Zhang[1, 2, *], QianQian Zhao[1, 2], Xiaohong Ji[3], Wei Chen[4] and Biruk Assefa[5]

Abstract: As an Industrial Wireless Sensor Network (IWSN) is usually deployed in a harsh or unattended environment, the privacy security of data aggregation is facing more and more challenges. Currently, the data aggregation protocols mainly focus on improving the efficiency of data transmitting and aggregating, alternately, the aim at enhancing the security of data. The performances of the secure data aggregation protocols are the trade-off of several metrics, which involves the transmission/fusion, the energy efficiency and the security in Wireless Sensor Network (WSN). Unfortunately, there is no paper in systematic analysis about the performance of the secure data aggregation protocols whether in IWSN or in WSN. In consideration of IWSN, we firstly review the security requirements and techniques in WSN data aggregation in this paper. Then, we give a holistic overview of the classical secure data aggregation protocols, which are divided into three categories: hop-by-hop encrypted data aggregation, end-to-end encrypted data aggregation and unencrypted secure data aggregation. Along this way, combining with the characteristics of industrial applications, we analyze the pros and cons of the existing security schemes in each category qualitatively, and realize that the security and the energy efficiency are suitable for IWSN. Finally, we make the conclusion about the techniques and approach in these categories, and highlight the future research directions of privacy preserving data aggregation in IWSN.

Keywords: Industrial wireless sensor network, wireless sensor network, cyber security, secure data aggregation protocol.

1 Introduction

The Industrial Wireless Sensor Network (IWSN) is more and more popular application systems which are employed with some resource-constrained with limited energy,

[1] Key Laboratory of Wireless Sensor Network & Communication, Shanghai Institute of Micro-system and Information Technology, Chinese Academy of Sciences, Shanghai, 201899, China.

[2] Shanghai Research Center for Wireless Communication, Shanghai, 201210, China.

[3] College of Physics and Electronic Information Engineering, Qinghai University for Nationalities, Xining, 810007, China.

[4] School of Computer Science and Technology, China University of Mining and Technology, Xuzhou, 221116, China.

[5] Infromation Communication Technology Department, Wollo University, Dessie Ethiopia, Ethiopia.

[*] Corresponding Author: Wuxiong Zhang. Email: wuxiong.zhang@wico.sh.

computation, memory and storage capacity. In practical application systems, the large number of sensor nodes is often randomly distributed in the predetermined area to collect the sensed parameter such as temperature, light, or smoke [Gilbert, Kaliaperumal and Rajsingh (2012); Zhao, Xu and Zhang (2016)]. Since the monitoring area of some adjacent nodes may overlap, the information sensed may have a lot of redundancy and may be highly correlated. If the sensed data are sent to the BS (Base Station) without any processing, the amount of data involved in the communication process will cause the sensor nodes to consume too much energy, and shorten the lifetime of the entire network [Feng, Jiang, Lim et al. (2013)]. On the other hand, the simultaneous transmission of sensed data from multiple nodes may increase the probability of data collisions and packet loss, which inevitably reduces the communication efficiency. Faced with this situation, it is apparent that data aggregation becomes an effective technique to aggregate these large redundant sensed data from multiple nodes into less redundant, high quality data in the intermediate nodes to reduce the number of transmitting data, and save the energy and bandwidth of nodes. Finally, it not only improves the accuracy and efficiency of data sensed, but also prolongs the network lifetime.

As the IWSN applications step into the measurement of the increasingly sensitive data associated with the daily life, the privacy protection of the sensed data has become a very important factor hindering the booming development of these applications. Although the information security can be enhanced through some traditional technologies, which involve the trust management pattern [Fang, Shi, Shan et al. (2015)], the encryption/decryption technologies [Zhao, Zhang, Li et al. (2016); Osama, Abdullah and Elankovan (2015)] and secure network coding [Fang, Shan, Jia et al. (2016)], unfortunately, these technologies are not congenitally suitable for being applied in IWSN, especially in some specific industrial scenarios (i.e., smart grid or meter [Wang, Lin, Liang et al. (2016)]). Currently, like WSN, the security issue is also facing enormous challenges and vulnerable to various types of attacks due to the inherent characteristic of IWSN and the features of the deployed environment. Therefore, the privacy preserving of data aggregation has aroused widespread concern in the industry and academia, and many different secure data aggregation protocols have been proposed, which achieves the data privacy through the encryption technology or not. However, in this paper, we do not construct any new privacy preserving data aggregation protocols, but just make an analytical summary in accordance with all the existing solutions. All we expect is that more and more secure and efficient mechanisms can be designed through this effort. The rest of the paper is organized as follows. In Section 2, we review the primary security requirements and security techniques employed in the existing secure data aggregation protocols. In Section 3, we classify the secure data aggregation protocols into three categories based on the implementation form of secure data aggregation. In Section 4, we analyze these security techniques of different protocols and compare the performance of the different protocols from two perspectives of security and performance. The conclusion is drawn in Section 5. Finally we give some future research directions about the secure data aggregation protocols in Section 6.

2 Related works

2.1 Data aggregation in IWSN

In IWSN, the researchers have been focusing on the lossless aggregation, which means that the content could usually not be aggregated into one single value which does not require additional overhead to the transmission. The relayed content must be formulated into its original form, opposed to aggregation functions [Dobslaw, Gidlund and Zhang (2015)]. The relevant content, however, did not often extend much more than a few bytes. Knowing the timestamp of the data packet creation, its origin, and the small sensor reading is usually sufficient for the IWSN application to operate the computation according to their requirements. Liu et al. [Liu, Yu and Zeng (2008)] proposed an algorithm which provided a suboptimal solution. In the proposed algorithm, the selection of aggregators, which is well-adapted to the network structure, was made under the constraint of minimal data collection cost.

As we all know, wirelessHART is a wireless sensor networking technology based on the Highway Addressable Remote Transducer protocol (HART). Developed as a multi-vendor, interoperable wireless standard, wirelessHART was defined in the requirements of process industry field device networks [Wes (2007)]. Neander et al. [Neander, Lennvall and Gidlund (2011)] gave the direct appliance of data aggregation with wirelessHART for reducing energy-consumption. One of the free reserved bits of the control byte in the network layer was used to signal whether the packet was aggregated or not, and to make the approach backwards compatible. The other noteworthy contributions considering, adaptive slot-assignment [Barnawi (2012)], and interference models [Xu, Li and Song (2013)] had also been made. However, the wirelessHART was being deployed for control applications, the scalable topology evaluation algorithms, such as SchedEx [Dobslaw, Gidlund and Zhang (2016)], were seldom involved. Currently, the deployment of IWSN indeed required more longer and costly test-series. In the IWSN deployment process, the concern was not just network capacity, as well as the end-to-end latency. Data aggregation is a promising direction to achieve the reduced end-to-end latency [Dobslaw, Gidlund and Zhang (2015)]. Li et al. [Li, Zhang, Jia et al. (2015)] proposed a data aggregation mechanism for energy-efficient and real-time WirelessHART communication. Especially, a greedy-based heuristic was applied during the super-frame construction phase to assign package transmissions whose data could be aggregated at intermediate routing nodes into nearby time slots to improve the energy saving and to prolong the overall lifetime of the network [Wang, Zhang, Xiong et al. (2016)].

In addition, we believe that the data aggregation could improve the Quality of Service (QoS), and achieve the trade-off between end-to-end reliability and latency in IWSN. Some researchers have begun to focus on this aspect, such as Kreibich et al. [Kreibich, Neuzil and Smid (2014)]. Unfortunately the data aggregation techniques in IWSN were few researched. In view of the present situation, we have to admit that it is a very promising field with the advancement of industry 4.0.

2.2 Primary security requirements

Due to the hostile deployment and the inherent characteristics of resource-constrained sensor nodes, the achievement of security has been a challenging task for IWSN. The

primary security requirements of the privacy preserving data aggregation protocols in WSN are data confidentiality, data integrity, data freshness, data availability, data authentication, etc. [Kumar and Dutta (2015)]. As a matter of course, these security requirements are also suitable for IWSN. Before reviewing the specific protocols, we first give the introduction about the definitions of these safety requirements.

Data confidentiality: It ensures that unauthorized parties or individuals cannot obtain secret information about sensed data. The perceptual data can only be obtained by the designated users.

Data integrity: Although data confidentiality ensures that only the legitimate participants can obtain the plaintext data through their own master key information, it cannot ensure that the content of sensed data has not been altered by a malicious adversary. Besides, even without considering the existence of malicious attackers, due to the nature of wireless medium, the transmitted data might be destroyed or lost.

Data freshness: It ensures that the sensed data are updated over time and the previously used data cannot be replayed.

Data availability: It ensures that the network must be running normally, and the data received is identifiable and accessible.

Data Authentication: Because the malicious adversary can also inject the false data into the network, and capture the legal nodes to launch the Sybil attack, it is quite necessary to ensure that the communicating node is the one that it claims to be.

2.3 Security techniques

Up to now, some secure data aggregation schemes have been proposed to meet these requirements in above subsection. In this subsection, we briefly review some of the security mechanisms, which are widely used in secure data aggregation in WSN.

2.3.1 Homomorphism encryption

Homomorphism encryption is a new and special kind of encryption scheme that allows the direct computation of the encrypted data [Ertaul, Yang and Saldamli (2015); Chandra, Choksi and Wadhwani (2013)]. Depending on the form of operations, all homomorphic encryption schemes can be divided into additive homomorphisms and multiplicative homomorphisms. Let Q and R denote two rings, + and × denote additive and multiplicative respectively, and K represents the key space. E: K×Q→R represents an encryption algorithm. Similarly, D: K×R→Q represents a decryption algorithm. If a, b∈Q, k∈K, and additive homomorphism encryption algorithm and multiplicative homomorphism encryption algorithm are shown as follows.

$$a+b=D_k\left(E_k\left(a\right)+E_k\left(b\right)\right)_i \tag{1}$$

$$a\times b=D_k\left(E_k\left(a\right)\times E_k\left(b\right)\right)_i \tag{2}$$

2.3.2 Digital signature

During the process of digital signature, the *Tx* encrypts the hash value of the message

with a private key, and Rx decrypts the message with a public key [Panoat and Richard (2012)]. Because the private key cannot be calculated from the public key, so the public key does not leak any information about the private key. The digital signatures technology not only provide non-repudiation like handwritten signatures but also achieve the data integrity with the help of the hash value of the corresponding message.

2.3.3 MAC (Message Authentication Code)

MAC is a value, which is obtained from the key and the message digest to verify data integrity. Before transmitting the data, *Tx* first calculates the MAC of the message through a hash function shared with Rx. Then the data with MAC will be transmitted to Rx. When Rx received the data, it can recover data and MAC by session key, and then calculates the MAC of the received data. When the new MAC and the received MAC was compared, if they are equal, the message authentication is successful.

2.3.4 Digital watermarking

Digital watermarking is a kind of information hiding technology which can embed special information into the original data [Harjito, Potdar and Bambang (2012)]. Before the sensed data are transmitted, Tx calculates the digital watermarking of the sensed data via a special provision shared with Rx. Then, data packets are transmitted to Rx. Rx received and carried it out. Only the data containing the correct watermark are considered reliable, and will be processed and forward, otherwise, will be discarded.

2.3.5 Data-slicing

In the aggregation tree, the leaf nodes randomly select a set of nodes in the h hop range to be the h hop neighbor nodes Si (K=Si). In large scale WSN, h can be equal to 1. Leaf nodes also belong to the set Si. When the data v_i sensed by node i, it sliced into k blocks and transmitted them on different paths. Leaf nodes not only transmit data slices, but also receive other data slices. Finally the leaf nodes aggregate their own data and other received data slices, and then send it to the upper node.

3 Secure Data Aggregation

In this section, we classify the existing secure data aggregation protocols into three categories: hop-by-hop encrypted data aggregation, end-to-end encrypted data aggregation and unencrypted secure data aggregation. The superficial difference is that, the aggregation node can decrypt the transmitted data and operate the aggregation algorithm based on it in the category of hop-by-hop encrypted data aggregation; the aggregation node does not have the access privilege to the sensed data and make the aggregation computation on the cipher text in the category of end-to-end encrypted data aggregation; no cryptographic operation exists in the category of unencrypted secure data aggregation. In the following, we will give a detailed description of the typical protocols in each category.

3.1 Hop-by-hop encrypted data aggregation

Yang et al. [Yang, Wang, Zhu et al. (2006)] proposed a secure data aggregation protocol

(SDAP) based on the divide and conquer and commit-and-attest specification. The protocol was made up of network construction phase, data aggregation phase and verification and attestation phase. Yoon et al. [Yoon, Jang, Kim et al. (2010)] proposed a new sensitive data aggregation (NSDA) scheme for protecting integrity. In order to preserve and hide the sensed data from multiple nodes, the protocol exploited complex number which a mathematical expression used arithmetic operations during the process of data aggregating and transmitting. All nodes made use of two different keys, one was pair wise secret key shared with a MD (Master Device) to be a trusted node of the network, and the other was a symmetric key shared with each sensor node lying on their aggregation tree in order to achieve secure transmissions.

In the following, Blaβ et al. [Blaβ and Zitterbart (2006)] presented a secure data aggregation (HCDA) protocol based on Hilbert-curve technique [Kim, Lee and Yoon (2013)] and seed exchanges among sensor nodes. Firstly, every node determined its sibling nodes, parent node, and child nodes by flooding scheme. In order to implement the network balance, HCDA specified the maximum number of child nodes for a node. Then, in order to exchange seed with other sibling nodes, each node generated a random seed data. The seed was used for hiding the original sensed data. The original sensed data can be changed by extracting some part of a seed value that sent to other nodes. Some part of the seed value was also added from another node. Finally, the original sensed data can be hidden.

Sanli et al. [Sanli, Ozdemir and Cam (2004)] also proposed a Secure Reference-based Data Aggregation (SRDA). Sensor data packets were compared with reference values at the beginning. Then, the difference was transmitted to a cluster head in the form of ciphertext. The data aggregation algorithm can be achieved by reducing the amount of the transmission information between the sensor nodes with the cluster head, and the energy loss is reduced enormously in the process of data transmission at the same time. Coincidentally, Rahayu et al. [Rahayu, Lee and Lee (2014)] proposed an Energy-efficient and Secure Pattern-based Data Aggregation (ESPDA). At first, the sensor nodes send their pattern codes based on the secret pattern seed, then the pattern codes are compared to the cluster head. Only the same ones are selected and they transmit the encrypted data to the cluster head. Employing the technique of Nonblocking OVSF Block-Hopping (NOVSF-BH), the security and the efficiency of data aggregation in WSN are significantly improved.

At the same time, considering the trades-offs among communication, application, security and computation, Li et al. [Li and Gong (2010)] proposed a secure date aggregation protocol of supporting additive aggregation with weights (SAAW). In this scheme, a weight parameter is introduced, which makes base station or aggregators know the weights of messages contributors. And the base station has a different key with nodes based on homomorphic hashing technique. A distributed compressed sensing-based privacy-preserving data aggregation (DCSPDA) scheme was proposed by Wu et al. [Wu, Yang and Wang (2016)]. This approach contains three phases, in the first phase of which the original date are measured; directly individually transmitted to the sink and the sink modified the positions of sparse coefficients to hide true sensor data in the second phase of which; in the third phase of which data are reported and aggregating. Parmar et al. [Parmar and Kadhiwala (2016)] proposed a protocol called zero configuration data aggregation (ZCDA) that attempts to achieve confidentiality, authentication, and integrity

together in one protocol. This protocol is divided into four subsections which include bootstrapping, aggregation tree construction, key establishment, and aggregation. Guo et al. [Guo, Zhang and Ma (2016)] introduced two versions of multi-functional secure data aggregation (MFSDA) called MFSDA-I and MFSDA-II respectively, where the former can obtain accurate results, and the later has low network traffic. To provide order-preserving and value-preserving, mapping step and coding step are used. Active nodes can be counted directly from aggregation data. Zhang et al. [Zhang, Han and Cai (2017)] proposed a ring-based privacy-preserving aggregation scheme (RiPPAS). This scheme utilizes the pseudonym mechanism for anonymous communication and utilizes the homomorphic encryption to add noise to the data, it also can handle different kinds of aggregation queries accompanying with the privacy preserving of data aggregation.

3.2 End-to-end encrypted data aggregation

Ozdemir [Ozdemir (2007)] proposed an end-to-end secure data aggregation protocol (CDAP) based on asymmetric privacy homomorphism. The node encrypted the sensed data with the public key of BS. Due to the symmetric key algorithm, although the compromised aggregating node may reveal the secrecy to neighboring nodes, it was only local. The aggregating node decrypted and aggregated all the received data. Then the aggregated data were encrypted with the public key of the BS and transmitted to the BS. Before BS obtains the aggregated data, the aggregating nodes hierarchically aggregate the encrypted data. The data were hidden from nodes in the path to the BS based on the privacy characteristic of homomorphism encryption. At the BS, the final aggregated data was decrypted by the private key of BS.

Boubiche et al. [Boubiche, Boubiche, Homero et al. (2016)] also put forward a secure data aggregation protocol based digital watermarking (SDAW). The watermark was integrated first in a fixed space to improve the security level. The packet was completed with the pure received data. In this way, the watermark would hardly be intercepted by an attacker. Based on the research above, when the sensed data are transmitted from a cluster member to the aggregation node, the homomorphic encryption technique is the most commonly used technology to guarantee integrity. However, the key generation and distribution mechanisms involve additional computation costs and consume more of energy. To address this problem the mechanism also used a lightweight fragile watermarking technique without encryption to insure the authentication and the integrity of the sensed data while saving the energy. The links between the sensor nodes with the aggregation nodes and the links between the aggregation nodes with the base station are secured by using the watermarking mechanism.

Omar et al. [Omar, Merad, Sidi et al. (2015)] proposed a secure data aggregation protocol for WSN based on one stateful public key encryption and homomorphic encryption (SASPKC). Before the deployment of the network, the BS generated its pair public-private and kept the private key. Each sensor was loaded with a secret key shared only with the BS, and also loaded with a large number M and a HMAC. Then, each node transmitted its state to BS in order to generate the sub-keys needed during the aggregation phase, and the CH (Cluster Head) acted as a data forwarder nor a data aggregator. In the phase of data aggregation, the data value captured by a key was encoded before encryption, then

ciphertext generated its MAC. CH sent the encrypted data and MACs to the BS. When the BS received all data packets, it invokes the decryption and verification processes.

In addition, Patel et al. [Patel and Raja (2015)] proposed a homomorphic aggregation system based on a public key encryption (PKE) scheme to protect the security of the sensed data. The proposed protocol could guarantee end-to-end confidentiality and privacy for hierarchical-based network. Parmar et al. [Parmar and Jinwala (2016)] put forward an integrity assured concealed data aggregation protocol (IACDAP). IACDAP used one symmetric key-based homomorphic primitives to provide end-to-end privacy and end-to-end integrity of reverse multicast traffic for tree-based network. In this protocol two different homomorphic primitives namely, homomorphic encryption and homomorphic message authentication code were used. This protocol had some viability and efficiency of resource-constrained devices. Moreover, it could defend some well-known attacks on concealed data aggregation. Pavithra et al. [Pavithra and Shruthika (2015)] also proposed one data aggregation approach on the sensed data by deploying the sensor nodes. This approach could conveniently reduce the number of transactions and increases the lifetime of wireless sensor network. The whole process are divided into four steps: selecting aggregator with node registration module, generating distributing key, aggregating and encrypting data, decrypting data and operating the computation in BS.

As mentioned above, we find that the symmetric key based homomorphic primitives are used to improve the energy efficiency and prolong the network's lifetime in resource-constrained WSN. However, the homomorphism encryption is inherently malleable, which makes sensor readings vulnerable against active attackers. As aggregator nodes do not require any secret information to aggregate data packets, any malicious node can inject fake data packets to falsify genuinely aggregated data. Hence, it is a formidable challenge to realize conflicting requirements, such as end-to-end privacy and end-to-end integrity, while performing data aggregation. Parmar et al. [Parmar and Jinwala (2016)] proposed the first protocol to achieve the data aggregation with end-to-end security for WSN (MRCDA). The proposed protocol uses an EC-ElGamal cryptosystem [Koblitz (1987)] based MAC for layer-wise integrity protection against outsider adversaries, and uses a homomorphic MAC algorithm [Agrawal and Boneh (2009)] for protecting the network against active insider adversaries. Additionally, the proposal also showed the resilience against the well-known cryptographic attacks such as the known plaintext/ciphertext attacks, malleability, node capture attacks, replay attacks and denial of service attacks. The comparison of the communication overhead with respect to the existing protocols exhibits the viability and efficiency of the proposed protocol on resource-constrained devices.

Considering the problem of active attack as well, Elhoseny et al. [Elhoseny, Elminir, Riad et al. (2016)] also proposed a novel encryption schema based on ECC [Zhou, Yang and He (2014)] and homomorphic encryption to secure data aggregation in WSN with dynamic clustering nature. Targeting the optimization of the lifespan of the whole network, genetic algorithm is applied to search for the most suitable sensor nodes as the cluster heads to transmit the messages to base station. The encryption key at each node is about 176 bits and is generated by the combining ECC key, identification number, and the distance to its cluster head (CH). The homomorphic encryption is also employed to allow CH to aggregate the encrypted data of its cluster members without decrypting them and

created the final message that will be sent to the base station. Thus, it prevents the attacker from knowing anything even if the CH is compromised, because CH is not responsible to encrypt messages. Compared with other methods, the experimental results demonstrated that the proposal greatly improves the network performance in consideration of lifetime, communication overhead, memory requirements, and energy consumption. Specially, it prevents a passive attack, CH compromised attack, and brute force attack.

Although most of the proposals indeed have satisfied the requirements for the privacy preserving of data aggregation at some extent, two common problems remain unsolved: high computation complexity and large communication overhead. Meanwhile, besides lightweight-oriented considerations, the integrity authentication of sensitive information can effectively identify whether or not the recipient of the message has been tampered by malicious nodes or some errors occurred during transition. Jiang et al. [Jiang, Chen, Zhu et al. (2016)] proposed a lightweight and integrity-protecting oriented data aggregation scheme for Wireless Sensor Network (LIPDA) which has lightweight, secure and easy operability to preserve data privacy and integrity during data aggregation in wireless sensor network. Before giving the concrete data aggregation scheme, one distance-based formation scheme of network topology is presented to balance the energy consumption of cluster heads. The complex number is encrypted by one additive homomorphic encryption method, which can realize the data aggregation without any decryption. Also, the reliability of data is ensured by using integrity verification method based on the complex operation. The theoretical analysis and simulation results show that the proposed scheme LIPDA can meet the requirement of privacy protection.

It is apparent that Elliptic Curve El-Gamal homomorphic encryption algorithm has been widely used to protect end-to-end data confidentiality. However, these works suffer from the expensive mapping function during decryption. If the aggregated results are huge, the base station has no way to gain the original data due to the hardness of the elliptic curve discrete logarithm problem. Therefore, these schemes are unsuitable for the large-scale WSN. Cui et al. [Cui, Shao, Zhong et al. (2017)] proposed a secure energy-saving data aggregation scheme designed for the large-scale WSN (DALS). It employed Okamoto-Uchiyama homomorphic encryption algorithm (OU homomorphic encryption algorithm) [Okamoto and Uchiyama (1998)] to protect end-to-end data confidentiality, employed MAC to achieve the in-network false data filtering, and utilized the homomorphic MAC algorithm (H-MAC scheme) [Agrawal and Boneh (2009)] to achieve end-to-end data integrity. Two popular IEEE 802.15.4 compliant wireless sensor network platforms, Tmote Sky and iMote 2, have been used to evaluate the efficiency and feasibility of the scheme. The results demonstrate that the proposal achieved better performance in reducing energy consumption.

3.3 Unencrypted secure data aggregation

Goat et al. [Groat, He and Forrest (2011)] proposed an unencrypted secure data aggregation protocol (KIPDA), which preserved the privacy of aggregated data through adding a set of camouflage values to the sensed data set. A message set was formed by combining the data sensed by the sensor nodes and the camouflage data, and designed a special selected principle on the position of real data. Sensor node transmitted sensed data to the aggregating node to aggregate data by nonlinear functions such as MAX/MIN.

Zhang et al. [Zhang, Wang and Feng (2008)] proposed a secure data aggregation protocol based on the perturbation histogram model (GP^2S). Sensor nodes mapped the data collected to the histogram to achieve the generalization of the data. The processed data were disturbed and transmitted to the upper node. BS received final aggregated data; removed the perturbation data; obtained the distribution histograms of all data, and finally achieved the approximate data aggregation result, such as Max/Min, sum, averages, etc. He et al. [He, Liu, Nguyen et al. (2007)] proposed a secure data aggregation called Sice-Mix-AggRegaTe (SMART) based on slicing technology. This scheme was used to protect the privacy for data aggregation while leading to a large number of the exchanged messages in the network. Liu et al. [Liu, Liu, Zhang et al. (2013)] proposed a secure data aggregation called High Energy-Efficient and Privacy-Preserving (HEEPP) based on modified slicing and assembling technology. By using the scheme, higher data accuracy was gotten, with lower energy consumption and less communication overhead. Li et al. [Li, Lin and Li (2011)] proposed an Energy-Efficient and High-Accuracy data aggregation (EEHA) to achieve the accurate data aggregation and the privacy protection with the high communication overhead and computational requirements in cluster/tree topology. The nodes were divided into leaf nodes and intermediate nodes in the scheme and the performance of data aggregation was improved. The Energy-Efficient and High-Accuracy data aggregation involves four steps: Aggregation tree construction, Slicing, Mixing and Aggregation.

The existing cluster-based private data aggregation techniques are the most energy-intensive due to the high message transmission complexity. Reliable data transmissions are also vital for resource constraint WSN. Manjula and Raja, as well as Liu et al. [Manjula and Raja (2018); Liu and Liu (2018)] proposed a reliability enabled private data aggregation technique that has the message transmission complexity of $\mathbf{O}(N)$. Every node in the cluster cleaves its data into n integers using simple modular arithmetic with suitable prime moduli and transmits it to the cluster heads (CHs) for intermediate aggregation. The CHs, in turn, forward the aggregate data to the base station where the final aggregated data is recovered using an elegant Chinese remainder theorem. The protocol exploits data privacy, communication overhead, and reliability metrics to gauge the performance of the proposed work. Numerical and simulation results demonstrate that the proposed solution outperforms the existing schemes having $\mathbf{O}(N^2)$ communication complexity.

Although the encrypted/decrypted techniques could provide and improve the security, the computational cost and energy consumption could be increased with sink nodes respectively. At present, the unencrypted secure data aggregation is a better technique than other techniques, which is a trade-off between the security and energy efficiency. While considering the application scenarios in WSN, we think the Hop-by-hop Encrypted Data Aggregation is more suitable for WSN due to its achievement of the trade-off between the security and the network layout.

4 Comparison and analysis

4.1 Security analysis

In this subsection, we analyze the different secure data aggregation protocols based on security technologies and aggregation functions, the result as shown in Tab. 1.

Table 1: Security analysis of secure data aggregation protocols

Classification	Protocol	Security technologies	Aggregation functions
Hop-by-hop encrypted	SDAP	Symmetric encryption, MAC	MAX/MIN
	NSDA	Symmetric encryption, Data hiding	SUM
	HCDA	Symmetric encryption, Hilbert-curve, seed exchange	SUM
	ESPDA	Nonblocking OVSF Block-Hopping	Additive
	SRDA	Reference values	MAX/MIN
	SAAW	Homomorphic hashing	SUM
	DCSPDA	Data hiding	SUM
	ZCDA	AES	Additive
	MFSDA	Homomorphic encryption and Non homomorphic encryption	Additive
	RiPPAS	pseudonym mechanism/ Homomorphic encryption	(MAX/MI)/ SUM
End-to-end encrypted	CDAP	Homomorphic encryption	Additive
	SDAW	Digital watermarking	SUM
	SASPKC	Stateful public key encryption and homomorphic encryption	SUM
	Patel	Public key encryption and homomorphic encryption	SUM
	IACDAP	Symmetric key encryption, homomorphic encryption and homomorphic MAC	SUM
	Pavithra	Public (encryption) key and the private (decryption) key	Additive
	MRCDA	EC-ElGamal cryptosystem based MAC and homomorphic MAC algorithm	Additive
	Elhoseny	Elliptic Curve Cryptography and homomorphic encryption	Additive
	LIPDA	Additive homomorphic encryption	Additive
	DALS	OU homomorphic encryption algorithm and the H-MAC scheme	Additive
Unencrypted	KIPDA	Data perturbation	MIN/MAX
	GP^2S	Data perturbation histogram	MIN/MAX, SUM,AVE
	SMART	Slicing and associative property of addition	Additive
	HEEPP	Modified slicing and assembling	SUM
	EEHA	Slicing and mixing	SUM
	SDAW	Watermarking technique	MIN/MAX, SUM,AVE
	Raja	Chinese remainder theorem	SUM

4.2 Performance analysis

We evaluate the performances of some classical protocols based on the following metrics qualitatively.

- Communication cost (CMC): The number of messages transmitted in the entire IWSN.
- Computation cost (CPC): The processing overhead of processor to achieve secure data aggregation.
- Data accuracy (DA): The aggregated data received at the BS divided by the sum of real original data give the description of the accuracy level.
- Delay (DLY): The time taken to get the sensed data from the source node to the BS.
- Data Integrity (DI): It guarantees that the aggregated data has not been altered by an adversary and other malicious nodes during the transmission of the sensed data.

Table 2: Performance comparison of different schemes

Classification	Protocol	CMC	CPC	DA	DLY	DI
Hop-by-hop encrypted	SDAP	M	H	H	H	N
	NSDA	L	M	H	L	N
	HCDA	L	M	H	L	Y
	ESPDA	L	H	M	H	Y
	SRDA	M	L	M	L	Y
	SAAW	H	H	L	M	Y
	DCSPDA	M	L	M	L	Y
	ZCDA	M	L	H	H	Y
	MFSDA-I	L	M	H	H	N
	MFSDA-II	L	M	L	H	N
	RiPPAS	L	L	M	L	Y
End-to-end encrypted	CDAP	M	M	M	L	Y
	SDAW	L	M	H	M	Y
	SASPKC	L	H	H	L	Y
	Patel	M	L	L	H	N
	IACDAP	L	M	M	L	N
	Pavithra	M	H	M	L	Y
	MRCDA	H	H	H	H	Y
	Elhoseny	M	H	H	L	N
	LIPDA	L	L	H	L	Y
	DALS	H	H	H	L	Y
Unencrypted	KIPDA	L	L	H	L	Y
	GP^2S	L	M	L	L	N
	SMART	L	H	L	M	Y
	HEEPP	M	M	M	L	N
	EEHA	M	H	H	L	Y
	SDAW	L	L	H	L	Y
	Raja	L	L	M	L	N

Legend: L=Low, M=Medium, H=High, Y=Yes, N=No.

5 Future directions

Data aggregation is crucial for IWSN, and its security performance requires more consideration. Some issues need to be studied as follow.

Most of the researches assume that the aggregate node is a well-resourced. It is feasible to a secure data aggregation protocol in the absence of the well-resourced node. Data aggregation not only can save energy and prolong network lifetime, but also bring some issues about security. Therefore, the trade-off between the security and energy efficiency is paramount to improve performance.

Most secure data aggregation protocols are only suitable for static IWSN. Therefore, it is necessary to design a protocol that is suitable for dynamic IWSN. Meanwhile, each secure data aggregation protocol has its suitable application scenarios, so it is a great challenge to design a secure data aggregation protocol suitable for more and even all the application systems.

Most secure data aggregation protocols are suitable for the uplink in WSN since no additional relay logic is required. However, for downlink traffic in IWSN, where the content may be sent to multiple destinations, the problem becomes more complex. In IWSN, these large packets have to be fragmented into smaller sub-packets according to their routes and destinations. Hence, a combination security of aggregation and de-aggregation would be required for arbitrary flow or sensor actuator network, which increases the solution space and makes the identification of a sufficient secure aggregation scheme a more complex problem.

6 Conclusions

Just as WSN, the secure data aggregation protocol in IWSN not only reduces data traffic and energy consumption, but also achieves data security in a certain degree. In this paper, we investigate many different secure data aggregation protocols. These protocols could be classified into the hop-by-hop encrypted data aggregation, the end-to-end encrypted data aggregation and the unencrypted secure data aggregation. Security technology and aggregation function of different protocols are analyzed. Besides, the performance characteristics of these protocols are given in the form of a distinct table based on performance metrics. Finally, we have discussed some issues about the requirements of the secure aggregation protocols to be studied in the future. We think that our works will help to design more effective and secure data aggregation protocols for IWSN.

Acknowledgement: This work is partially supported by the National Natural Science Foundation of China (61571004), the Shanghai Natural Science Foundation (No. 17ZR1429100), the National Science and Technology Major Project of China (No. 2018ZX03001017-004), and the Scientific Instrument Developing Project of the Chinese Academy of Sciences (No. YJKYYQ20170074).

References

Agrawal, S.; Boneh, D. (2009): Homomorphic MACs: MAC-based integrity for network coding. *International Conference on Applied Cryptography and Network Security*, pp. 292-305.

Barnawi, A. Y. (2012): Adaptive tdma slot assignment using request aggregation in wireless sensor network. *Procedia Computer Science*, vol. 10, pp. 78-85.

Blaβ, E. O.; Zitterbart, M. (2006): An efficient key establishment scheme for secure aggregating sensor networks. *ACM Symposium on Information, Computer and Communications Security*, pp. 303-310.

Boubiche, D. E.; Boubiche, S.; Homero, T. C.; Pathan, A. K.; Bilami, A. et al. (2016): SDAW: secure data aggregation watermarking-based scheme in homogeneous WSN. *Telecommunication Systems*, vol. 62, no. 2, pp. 277-288.

Chandra, A.; Choksi, C.; Wadhwani, M. (2013): A survey of the privacy homomorphism in wireless sensor network. *International Conference on Distributed Computing and Internet Technology*, pp. 46-50.

Cui, J.; Shao, L. L.; Zhong, H.; Xu, Y.; Liu, L. (2017): Data aggregation with end-to-end confidentiality and integrity for large-scale wireless sensor network. *Peer-to-Peer Networking and Applications*, vol. 11, no. 5, pp. 1022-1037.

Dobslaw, F.; Gidlund, M.; Zhang, T. (2015): Challenges for the use of data aggregation in industrial wireless sensor network. *IEEE International Conference on Automation Science and Engineering*, pp. 138-144.

Dobslaw, F.; Gidlund, M.; Zhang, T. (2016): End-to-end reliability-aware scheduling for wireless sensor network. *Transactions on Industrial Informatics*, vol. 99, no. 2, pp. 1-10.

Elhoseny, M.; Elminir, H.; Riad, A. E. D. M.; Yuan, X. H. (2016): A secure data routing schema for WSN using elliptic curve cryptography and homomorphic encryption. *Journal of King Saud University-Computer and Information Sciences*, vol. 28, no. 3, pp. 262-275.

Ertaul, L.; Yang, J.; Saldamli, G. (2015): Analyzing homomorphic encryption schemes in securing wireless sensor network (WSN). *International Journal of Computer Science and Network Security*, vol. 15, no. 5, pp. 1-11.

Fang, W.; Shan, L.; Jia, G.; Ji, X.; Chen, S. (2016): A low complexity secure network coding in wireless sensor network. *Journal of Internet Technology*, vol. 17, no. 5, pp. 905-913.

Fang, W.; Shi, Z.; Shan, L.; Li, F.; Wang, X. (2015): Trusted scheme for defending on-off attack based on BETA distribution. *Journal of System Simulation*, vol. 27, no. 11, pp. 2722-2728.

Feng, D.; Jiang, C.; Lim, G.; Cimini, L. J.; Feng, G. et al. (2013): A survey of energy-efficient wireless communications. *IEEE Communications Surveys & Tutorials*, vol. 15 no. 1, pp. 167-177.

Gilbert, E. P. K.; Kaliaperumal, B.; Rajsingh, E. B. (2012): Research issues in wireless sensor network applications: a survey. *International Journal of Information and Electronics Engineering*, vol. 2, no. 5, pp. 702-706.

Groat, M. M.; He, W.; Forrest, S. (2011): KIPDA: k-indistinguishable privacy-preserving data aggregation in wireless senor networks. *IEEE INFOCOM*, pp. 2024-2032.

Guo, K.; Zhang, P.; Ma, J. (2016): Secure and cost-effective distributed aggregation for mobile sensor networks. *Sensors*, vol. 16, no. 4, pp. 583.

Harjito, B.; Potdar, V.; Bambang, H. J. S. (2012): Watermarking technique for wireless sensor network: a state of the art. *IEEE Eighth International Conference on Semantics, Knowledge and Grids*, pp. 253-256.

He, W.; Liu, X.; Nguyen, H.; Nahrstedt, K.; Abdelzaher, T. (2007): PDA: privacy-preserving data aggregation in wireless sensor network. *IEEE 26th IEEE International Conference on Computer Communications*, pp. 2045-2053.

Jiang, S. S.; Chen, Q. Z.; Zhu, J. B.; Liang, X. L.; Zhao, X. M. (2016): Lightweight and integrity-protecting oriented data aggregation scheme for Wireless sensor network. *IET Information Security*, vol. 11, no. 2, pp. 82-88.

Kim, Y.; Lee, H.; Yoon, M. (2013): Hilbert-Curve based data aggregation scheme to enforce data privacy and data integrity for wireless sensor network. *International Journal of Distributed Sensor Network*s, vol. 2013, no. 4, pp. 1-14.

Koblitz, N. (1987): Elliptic curve cryptosystems. *Mathematics of Computation*, vol. 48, no. 177, pp. 203-209.

Kreibich, O.; Neuzil, J.; Smid, R. (2014): Quality-Based multiple-sensor fusion in an industrial wireless sensor network for MCM. *IEEE Transactions on Industrial Electronics*, vol. 61, no. 9, pp. 4903-4911,.

Kumar, M.; Dutta, K. (2015): A survey of security concerns in various data aggregation techniques in wireless sensor network. *IEEE International Conference on Computer Design*, pp. 1-15.

Li, F.; Zhang, Z.; Jia, Z.; Ju, L. (2015): Superframe scheduling for data aggregation in WirelessHART networks. *IEEE International Conference on High Performance Computing and Communications*, pp. 1540-1545.

Li, H. J.; Lin, K.; Li, K. Q. (2011): Energy-efficient and high-accuracy secure data aggregation in Wireless Sensor Network. *Computer Communications*, vol. 34, no. 4, pp. 591-597.

Li, Z.; Gong, G. (2010): Data aggregation integrity based on homomorphic primitives in sensor networks. *Proceedings of the 9th International Conference on Ad-hoc, Mobile and Wireless Networks*, pp. 149-162.

Liu, C. X.; Liu, Y.; Zhang, Z. J.; Cheng, Z. Y. (2013): High energy-efficient and privacy-preserving secure data aggregation for wireless sensor network. International *Journal of Communication Systems*, vol. 26, no. 3, pp. 380-394.

Liu, L.; Yu, H.; Zeng, P. (2008): An optimized aggregators selection problem for industrial wireless sensor network. *Seventh World Congress on Intelligent Control and Automation*, pp. 4057-4062.

Liu, X.; Liu, Q. (2018) A dual-spline approach to load error repair in a hems sensor network. *Computers, Materials & Continua*, vol. 57, no. 2, pp. 179-194.

Manjula, R.; Raja, D. (2018): Efficient aggregation technique for data privacy in wireless sensor network. *IET Networks*, vol. 7, no. 5, pp. 287-293.

Neander, J.; Lennvall, T.; Gidlund, M. (2011): Prolonging wireless HART network lifetime using packet aggregation. *IEEE International Symposium on Industrial Electronics*, pp. 1230-1236.

Okamoto T.; Uchiyama, S. (1998): A new public-key cryptosystem as secure as factoring. *International Conference on the Theory and Applications of Cryptographic Techniques*, pp. 308-318.

Omar, R.; Merad, B.; Sidi, M. S.; Mohammed, F. (2015): A novel secure aggregation scheme for wireless sensor network using stateful public key cryptography. *Ad Hoc Networks*, vol. 32, no. C, pp. 98-113.

Osama, A. K.; Abdullah, M. Z.; Elankovan, A. S. (2015): ImgFS: a transparent cryptography for stored images using a file system in user space. *Frontiers of Information Technology & Electronic Engineering*, vol. 16, no. 1, pp. 28-42.

Ozdemir, S. (2007): Concealed data aggregation in heterogeneous sensor networks using privacy homomorphism. *International Conference on Pervasive Services*, pp. 165-168.

Panoat, C.; Richard, E. N. (2012): Multi-Resolution elliptic curve digital signature. *IEEE Conference on Local Computer Networks*, pp. 93-101.

Parmar, K.; Jinwala, D. C. (2016): *Malleability Resilient Concealed Data Aggregation in Wireless Sensor Network*. Kluwer Academic Publishers, Germany.

Parmar, K.; Jinwala, D. C. (2016): Symmetric-Key based homomorphic primitives for end-to-end secure data aggregation in wireless sensor network. *Journal of Information Security*, vol. 6, pp. 38-50.

Parmar, P.; Kadhiwala, B. (2016): Secure data aggregation protocol using AES in wireless sensor network. *Emerging Research in Computing, Information, Communication and Applications*, pp. 421-432.

Patel, K. J.; Raja, N. M. (2015): Secure end to end data aggregation using public key encryption in wireless sensor network. *International Journal of Computer Applications*, vol. 122, no. 6, pp. 27-32.

Pavithra, C. N.; Shruthika, C. A. (2015): The technique for securing end-to-end data aggregation in wireless sensor network. *International Journal of Advanced Engineering Research and Technology*, vol. 3, pp. 73-77.

Rahayu, T. M.; Lee, S. G.; Lee, H. J. (2014): Security analysis of secure data aggregation protocols in Wireless Sensor Network. *International Conference on Advanced Communication Technology*, pp. 471-474.

Sanli, H. O.; Ozdemir, S.; Cam, H. (2004): SRDA: secure reference-based data aggregation protocol for wireless sensor network. *IEEE 60th Vehicular Technology Conference*, pp. 4650-4654.

Wes, I. (2007): WirelessHart ready for prime time.

https://www.automationworld.com/dcs/wirelesshart-ready-prime-time.

Wang, H.; Zhang, P.; Xiong, L.; Liu, X.; Hu, C. (2016): A secure and high-performance multi-controller architecture for software-defined networking. *Frontiers of Information Technology & Electronic Engineering*, vol. 17, no. 7, pp. 634-646.

Wang, Y.; Lin, Z.; Liang, X.; Xu, W.; Yang, Q. et al. (2016): On modeling of electrical cyber-physical systems considering cyber security. *Frontiers of Information Technology & Electronic Engineering*, vol. 17, no. 5, pp. 465-478.

Wu, D.; Yang, B.; Wang, H. (2016): Privacy-preserving multimedia big data aggregation in large-scale wireless sensor network. *ACM Transactions on Multimedia Computing Communications & Applications*, vol. 12, no. 4s, pp. 60.

Xu, X.; Li, X.; Song, M. (2013): Efficient aggregation scheduling in multihop wireless sensor network with SINR constraints. *IEEE Transactions on Mobile Computing*, vol. 12, no. 12, pp. 2518-2528.

Yang, Y.; Wang, X.; Zhu, S.; Cao, G. (2006): SDAP: A secure hop-by- hop data aggregation protocol for sensor networks. *7th ACM international symposium on Mobile Ad Hoc Networking and Computing*, pp. 356-367.

Yoon, M.; Jang, M.; Kim, H.; Chang, J. (2010): A new sensitive data aggregation scheme for protecting integrity in wireless sensor network. *IEEE Computer Society International Conference on Computer and Information Technology*, pp. 2463-2470.

Zhang, K.; Han, Q.; Cai, Z. (2017): Rippas: a ring-based privacy-preserving aggregation scheme in wireless sensor network. *Sensors*, vol. 17, no. 2, pp. 300.

Zhang, W. S.; Wang, C.; Feng, T. M. (2008): GP^{2}S: generic privacy-preservation solutions for approximate aggregation of sensor data. *IEEE International Conference on Pervasive Computing and Communications*, pp.179-184.

Zhao, Y.; Xu, H.; Zhang, Q. (2016): An analysis in metal barcode label design for reference. *Frontiers of Information Technology & Electronic Engineering*, vol. 17, no. 2, pp. 173-184.

Zhao, Y.; Zhang, W.; Li, D.; Huang, Z.; Li, M. et al. (2016): Pegasus: a distributed and load-balancing fingerprint identification system. *Frontiers of Information Technology & Electronic Engineering*, vol. 17, no. 8, pp. 766-780.

Zhou, Q.; Yang, G.; He, L. (2014): A secure enhanced data aggregation based on ECC in wireless sensor network. *Sensors*, vol. 14, no. 4, pp. 6701-6721.