IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received June 12, 2017, accepted July 21, 2017, date of publication July 25, 2017, date of current version August 14, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2731819

Secure MAX/MIN Queries in Two-Tiered Wireless Sensor Networks

HUA DAI[1,2], MIN WANG[1], XUN YI[3], GENG YANG[1,2], AND JINGJING BAO[1]

[1]College of Computer Science and Technology, Nanjing University of Posts and Telecommunications, Nanjing 210023, China
[2]Jiangsu Key Laboratory of Big Data Security and Intelligent Processing, Nanjing 210023, China
[3]School of Science, RMIT University, Melbourne, VC 3001, Australia

Corresponding author: Hua Dai (daihua@njupt.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61300240, Grant 61402014, Grant 61572263, Grant 61502251, Grant 61472193, and Grant 61373138, in part by the Natural Science Foundation of Jiangsu Province under Grant BK20151511, Grant BK20161516, and Grant BK20130096, in part by the Project of Natural Science Research of Jiangsu University under Grant 14KJB520027 and Grant 15KJB520027, in part by the Postdoctoral Science Foundation of China under Grant 2015M581794, in part by the Postdoctoral Science Foundation of Jiangsu Province under Grant 1501023C, in part by NUPTSF under Grant NY214127, in part by the Natural Science Foundation of Anhui Province under Grant 1608085MF127, and in part by the Chuzhou University Scientific Research Fund under Grant 2015QD02.

ABSTRACT In wireless sensor networks, secure MAX/MIN query processing is a challenging issue, and it is useful in fields, where security is necessary. In this paper, we propose a secure MAX/MIN query processing method in two-tiered wireless sensor networks. To the best of our knowledge, it is the first work that can achieve data privacy protection and query result integrity verification simultaneously. Three schemes, naïve secure MAX/MIN query (NSMQ), complicated secure MAX/MIN query (CSMQ), and OSMQ, are designed to achieve secure MAX/MIN queries. In NSMQ, we present an intuitive and baseline solution that makes the master nodes return all the ciphertext as the query result. However, it may incur high query communication cost. To address this limitation, a CSMQ scheme is designed, which introduces the comparable factors (c-factors) based on 0–1 encoding verification to find the accurate encrypted query result from the stored ciphertext of the master nodes even when their real values are unknown. Then, a broadcasting method is introduced to generate minor-node-sets as the proofs for verifying the integrity of the query results. CSMQ can significantly reduce the query communication cost, but its in-cell communication cost is high because of the extra data submission and broadcasting. To balance the in-cell and query communication cost, OSMQ, as an optimized version of CSMQ, is proposed to address the minor-node-set compression and random c-factor selection. The proposed schemes are built upon symmetric encryption and hash-based message authentication coding primitives. OSMQ can prevent compromised master nodes from obtaining the plaintext of private data and force them to return integrity-satisfying query results to avoid being detected. Extensive theoretical and experimental studies have been conducted to demonstrate the efficacy and efficiency of the proposed schemes.

INDEX TERMS Tiered wireless sensor networks, MAX/MIN query, privacy preservation, integrity verification.

I. INTRODUCTION

The two-tiered wireless sensor network (TWSN) is a new architecture for sensor networks and is indispensable for prolonging network lifetime, as well as improving network scalability and stability [1], [2]. It can be utilized in a variety of critical applications, such as medical care, environment monitoring and national defense. As shown in Fig. 1, a TWSN is partitioned into multiple cells and has a base station, master nodes and sensor nodes in two tiers. The lower tier consists of a large number of resource-limited sensor nodes that are distributed in cells, while the higher tier consists of resource-rich master nodes, which are abundant in computation, storage and energy. Each cell has a master node. Sensor nodes collect data and submit it to the master node in the same cell for storage. The master nodes are responsible for processing ad hoc queries from the base station via an on-demand wireless link (e.g., satellite).

However, the master nodes usually attract attacks from adversaries in a hostile environment because of their critical role of storing the whole collected data of the network and

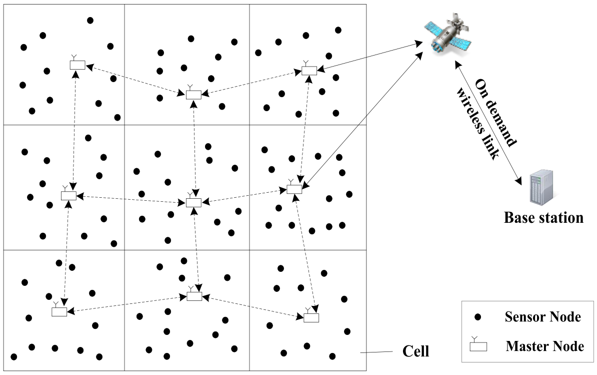


FIGURE 1. The architecture of a TWSN.

processing queries from the base station. If one of them is compromised, serious threats could be incurred. For instance, adversaries may manipulate compromised master nodes to eavesdrop and/or falsify the information of patients in a health-care monitoring network, which breaches the privacy of patients and/or even disturbs the medical decisions of patients. It is a challenge to design a secure query processing method for master nodes in a hostile environment because they must obtain information regarding the collected data for query processing, which may defeat the primary objectives of privacy preservation and integrity verification.

Data query is one of the important operations for event monitoring or data management in a TWSN. In recent years, secure range queries [3]–[12] and top-*k* queries [13]–[21] have been well addressed. However, secure MAX/MIN queries, which are made to obtain the maximum or minimum data securely in the interesting area and time slot, have not received much attention yet. The existing works regarding secure MAX/MIN queries in a TWSN, such as [22]–[24], focus on the privacy-preserving issues but never consider the integrity issues caused by data falsification attacks. To the best of our knowledge, there is still no literature that describes a secure MAX/MIN query method with privacy preservation and integrity verification simultaneously.

In this paper, we focus on processing secure MAX/MIN queries while preserving the privacy of data and verifying the integrity of the query results. In particular, we propose three schemes to achieve secure MAX/MIN query processing. The first one is a naïve secure MAX/MIN query scheme (NSMQ) that makes the master nodes return all the ciphertext contributed by the queried sensor nodes as the query result to the base station. However, NSMQ may incur a heavy query communication cost. We then propose a complicated secure MAX/MIN query scheme (CSMQ), in which the comparable factors (c-factors) based on 0-1 encoding verification are introduced to find the encrypted query result from the ciphertext stored in the master nodes even without knowing their real values, and then a broadcasting method is proposed to generate the minor-node-sets as the proofs for verifying the integrity of the query result. CSMQ can significantly reduce the query communication cost compared with NSMQ, but its in-cell communication cost is greater because of the extra data submission and broadcast. To balance the in-cell and query communication cost, we introduce minor-node-set compression and a random c-factor selection to optimize CSMQ, namely, OSMQ. The efficacy and efficiency of our schemes are validated by thorough theoretical and experimental study. With these in place, compromised master nodes will be prevented from obtaining the plaintext of private data and forced to return integrity-satisfying query results to avoid being detected.

The main contributions of this paper can be summarized as follows:

1) We propose a secure MAX/MIN query processing method that can not only protect the privacy of data but also verify the integrity of the query result in the TWSN. To the best of our knowledge, this is the first work that solves privacy and integrity issues simultaneously for secure MAX/MIN query processing.

2) We propose three schemes, NSMQ, CSMQ and OSMQ, where NSMQ is the most in-cell communication cost-saving scheme, CSMQ is the most query communication cost-saving scheme and OSMQ is a balanced scheme for both types of communication costs.

3) We conduct comprehensive simulations to evaluate the performance and communication costs of the proposed schemes.

The remainder of this paper is organized as follows. Related work is discussed in Section II, and the network model, query model, threat model and the problem statement are described in Section III. Three schemes, NSMQ, CSMQ and OSMQ, are presented in Section IV, V and VI, respectively. We evaluate the performance of our approach in Section VII and then conclude this paper in Section VIII.

II. RELATED WORK

In TWSNs, the security issues are the hot research spots in the data query area. In recent years, secure range queries [3]–[12] and the top-*k* query [13]–[21] have been broadly investigated. However, there has been limited research on MAX/MIN queries. To the best of our knowledge, there are only three works [22]–[24] that have studied or proposed solutions for secure MAX/MIN queries in TWSNs.

Regarding the MAX/MIN query, a privacy-preserving MAX/MIN query processing method is proposed in [22] that is based on prefix membership verification (PMV) [25], [26]. The basic idea is to use PMV, HMAC [27] and symmetric encryption to generate the encrypted data items, together with the corresponding comparable codes, in sensor nodes; then these data are submitted to the master nodes. When a query is started, the master nodes determine the qualified ciphertext, embed the query result using the secure comparing rules of PMV, and then submit it to the base station for decryption. Because extra codes that are generated by PMV and HMAC need to be transmitted from sensor nodes to master nodes, the energy consumption is costly. To reduce the network communication cost, the literature [23] uses 0-1 encoding verification [28] instead of the PMV used in [22] to achieve an energy-efficient privacy-preserving MAX/MIN query (EMQP). The 0-1 encoding verification mechanism can also compare data items without knowing their real values, but the corresponding generated codes are much fewer than that of PMV, thereby saving more communication cost than PMV. Based on EMQP, a random secure comparator selection optimization is introduced to achieve a more efficient privacy-preserving MAX/MIN query (RSCS-PMQ) [24]. Such a random selection mechanism cuts off almost half of the comparable codes submitted from sensor nodes to master nodes in EMQP. The random selected codes are still able to determine a candidate set (the mean quantity of its elements is approximately 2) of encrypted data items, embedding the query result by the proposed MaxRSC algorithm in RSCS-PMQ. To summarize, we can see that the existing works for TWSNs all focus on privacy preservation, but they cannot check the integrity of the obtained query result when a master node is compromised by an adversary and could falsify its stored data from sensor nodes.

Although a top-k query can be customized into a MAX/MIN query when $k = 1$ is settled, it is wasteful in energy consumption. The reason is that each sensor node should submit all the data collected in every epoch since a top-k query to obtain the highest/lowest k data items, where k is variable. However, in the MAX/MIN query, only the maximum or minimum data needs to be submitted. It is obvious that taking a top-k query as a MAX/MIN query will result in significant unnecessary data communications. As a result, we can see that the existing secure top-k query methods [13]–[21] are not adaptable for secure MAX/MIN queries. In addition, the existing secure range query methods [3]–[12] cannot solve the secure MAX/MIN query issues because a range query is totally different from a MAX/MIN query.

In addition, there are a few works regarding privacy-preserving MAX/MIN queries in traditional multi-hop wireless sensor networks, such as CDAM [29], KIPDA [30], PMMA [31] and SDAMv [32]. Although they can solve the privacy-preserving problem in MAX/MIN query processing, they cannot be used to achieve a secure MAX/MIN query with privacy preservation and integrity verification in the TWSN because they are designed for different sensor network architectures, and only for the purpose of preserving data privacy.

III. MODELS AND PROBLEM STATEMENT

A. NETWORK AND QUERY MODELS

We take the TWSN model widely adopted by the literature [3]–[24]. Its architecture is shown in Fig. 1, where the network is partitioned into multiple cells. Each cell has a master node and several ordinary sensor nodes. Master nodes are powerful and have abundant resources in energy, computation, and storage, while sensor nodes are cheap devices constrained in resources. We assume that time is divided into *slots*. At the end of each slot, each sensor node submits its collected data to its affiliated master node. Once a query is started, master nodes process it over their stored data and send the results to the base station via an on-demand wireless link (e.g., satellite) that is often costly and relatively low-rate.

A MAX/MIN query is a request to obtain the maximum or minimum data from an interesting area and time. For simplicity, we define a basic MAX/MIN query as a four-element tuple:

$$Q_t = (MAX/MIN, t, C, \Gamma_t) \tag{1}$$

where MAX/MIN indicates the query type, t is the queried slot number, C is the ID of an interesting cell, and Γ_t is the set of IDs of the queried sensor nodes in C.

Based on the query definition, we can see that any complicated MAX/MIN query containing multiple time slots, cells and/or queried sensor nodes can be easily decomposed into multiple basic ones. Therefore, we mainly focus on the basic MAX/MIN query Q_t in this paper. Without loss of generality, we assume that the queried cell is C, which consists of a master node M and n sensor nodes $S = \{s_1, s_2, \ldots, s_n\}$ whose IDs compose the set $\Gamma = \{1, 2, \ldots, n\}$. For simplicity, the "basic" is omitted in the subsequent discussion, and only the MAX queries are investigated, because the MIN queries are processed similarly.

B. THREAT MODEL

In this paper, we use the same threat model as [3]–[9], which assumes that the adversary tries to carry out attacks in the following two ways. First, the adversary could obtain sensitive data items from the sensor network, which violates *data privacy*. In fact, private information leakage is a critical threat in many applications. Second, the adversary could return forged or incomplete replies as the query result to the base station without being detected, which breaches the query result*integrity*. Here, the integrity of the query result includes the following three aspects: (1) All data items returned from the master nodes are originally submitted by the sensor nodes and remain unmodified. (2) No qualified data are omitted from the query result by the master nodes. (3) The returned data are collected by sensor nodes during the time of interest.

Because the master nodes not only store all the data items collected by the sensor nodes but also take charge of processing queries, they are the most likely to attract attacks from an adversary in a hostile environment. Once a master node is compromised, not only can an adversary easily obtain the sensitive data stored in it, but can also provide forged or incomplete responses to the base station as query results, which may mislead users' decisions. Meanwhile, the sensor nodes may also be compromised; they may leak their collected data, or may submit forged data that are extremely difficult to detect without tamper-proof hardware to their affiliated master nodes. However, the data from one sensor node is minor in relation to the whole network, and the non-compromised sensor nodes are always the majority; otherwise the network will be useless. Therefore, we mainly focus on the scenario where a master node is compromised and investigate countermeasures against the compromised master node. The aim of our work is to provide data privacy preservation and integrity verification for the MAX/MIN queries.

C. PROBLEM STATEMENT

We assume that each sensor $s_i \in S$ collects N data items during each time slot, denoted by $D_i = \{d_{i,1}, d_{i,2}, \ldots, d_{i,N}\}$. To obtain a unique maximum or minimum from multiple data items, it is necessary to determine the bigger or smaller one when comparing the values of any two data items. Moreover, even if their values are equal, they are still comparable if we take the node ID and the time of data collection into consideration. Therefore, we assume that all the data items collected in cell C during t are mutually different, which indicates that a unique query result R_t exists for any MAX query that satisfies

$$\forall d_{i,j}(d_{i,j} \in D_i \wedge i \in \Gamma_t) \rightarrow R_t > d_{i,j}. \tag{2}$$

The problem we solve in this paper is how to provide privacy preservation and integrity verification for MAX/MIN queries when confronted with a compromised master node. In detail, the goal of the former is to prevent the collected data items from being exposed to master nodes, while the latter is to enable the base station to verify the integrity of the query results.

Moreover, to evaluate the performance of our schemes, we introduce the following metrics:

1) Φ_{ic} is the in-cell communication cost and is measured by the total messages in bits transmitted by data submission of sensor nodes in a cell per time slot. The in-cell communication cost affects the lifetime of the network directly because of the energy limitation of the sensor nodes.

2) Φ_{qc} is the query communication cost and is measured by the total messages in bits transmitted between master nodes and the base station via an on-demand wireless link for processing a secure MAX/MIN query. Because the wireless link, such as a satellite link, is often costly and relatively low-rate, the query communication cost will directly affect the cost and efficiency of query processing.

We will perform the performance evaluations on the above three metrics in Section VII.

D. NOTATIONS

In this paper, the notations presented in Table 1 are used.

IV. NAIVE SECURE MAX/MIN QUERY SCHEME

In this section, we first propose a naive secure MAX/MIN query scheme, denoted as NSMQ, which is obviously efficient in Φ_{ic}. We use symmetrical encryption, e.g., DES or AES, to protect the security of the data, where k_i is a unique secret key shared with each sensor s_i and the base station.

TABLE 1. Notations in this paper.

Para.	Notations
k_i	The unique secret key shared with each sensor s_i and the base station
d_i	The maximum of data items collected by s_i in time slot t
c_i	The ciphertext of the concatenation
$\tau(c_i)$	The embedded time slot number in c_i
$\nu(c_i)$	The embedded collected data item in c_i
$\zeta(c_i)$	The minor-node-set in c_i
Ω_i	A set of the IDs of sensor nodes whose maximum is smaller than s_i
$CF^0(d_i)$	The 0-encoding comparable factors of d_i
$CF^1(d_i)$	The1-encoding comparable factors of d_i
R_t	The unique query result
t	The queried time slot number
C	The queried cell ID
Γ_t	The set of IDs of the queried sensor nodes in cell C
$bm(\Omega_i)$	An n-bit bitmap array map for Ω_i

A. DATA SUBMISSION

In data submission, each sensor transmits its collected data to the affiliated master node $\mathcal{M}$. Practically, after collecting N data items $D_i = \{d_{i,1}, d_{i,2}, \ldots, d_{i,N}\}$ in a time slot, each sensor $s_i \in S$ performs the following steps:

1) Obtain the maximum of D_i, which is denoted as $d_i = max(D_i)$.

2) Concatenate t and d_i, and encrypt the concatenation with k_i. After that, the ciphertext $(t \parallel d_i)_{k_i}$ is generated.

3) Submit the following message to $\mathcal{M}$.

$$s_i \rightarrow \mathcal{M} :< i, t, (t \parallel d_i)_{k_i} >$$

The above steps show that the maximum of data items collected by s_i in each time slot will be stored in $\mathcal{M}$ under encryption where the key is only shared with the base station. Therefore, it is computationally infeasible for $\mathcal{M}$ to obtain the plaintexts of the data received from any sensor node, and the time slot number t embedded in ciphertext can be used to verify whether it has been replaced with some old versions, whereby replay attacks can be easily detected.

B. QUERY PROCESSING

In query processing, $\mathcal{M}$ executes the query commands received from the base station and makes responses to them. Upon receiving a query $Q_t = (\text{MAX}, t, C, \Gamma_t)$, $\mathcal{M}$ transmits a message as follows to the base station for each queried sensor s_i where $i \in \Gamma_t$.

$$\mathcal{M} \rightarrow \text{base station} :< i, (t \parallel d_i)_{k_i} >$$

Because each queried sensor will contribute a piece of ciphertext for Q_t, $\mathcal{M}$ will return $|\Gamma_t|$ pieces to the base station. When the base station receives all messages from $\mathcal{M}$, it can easily obtain the query result $R_t = max\{d_i | i \in \Gamma_t\}$ after decrypting the ciphertext in each message.

C. QUERY RESULT VERIFICATION

Query result verification concerns how the base station verifies the integrity of the query result. Upon receiving messages from $\mathcal{M}$, the base station decrypts the ciphertext of each message. Assuming that $< i, c_i >$ is a message received where c_i is the ciphertext, we denote $\tau(c_i)$ and $\nu(c_i)$ as the embedded time slot number and collected data item in c_i, respectively. R_t should be considered to satisfy the integrity requirement only if the following two conditions are both met:

Condition 1: For each message $< i, c_i >$, there are $i \in \Gamma_t$ and $\tau(c_i) = t$ holds.

Condition 2: The base station has received $|\Gamma_t|$ distinct messages from $\mathcal{M}$, each of which corresponds to a unique queried sensor in Γ_t.

The first condition checks whether each message is indeed contributed by a queried sensor in t, while the second one verifies whether the received messages are complete. Because every sensor shares its key only with the base station, not only can $\mathcal{M}$ obtain no plaintext of any collected data but also any misbehavior by $\mathcal{M}$, such as forging or omitting qualified messages, will be discovered when verifying the above two conditions.

D. PERFORMANCE ANALYSIS

Here, we first derive the in-cell communication cost Φ_{ic} incurred by NSMQ. We assume that each time slot number is of l_t bits, each node ID is of l_{id} bits, the average hops between a sensor node and $\mathcal{M}$ is L, and each collected data item is of w bits. We adopt a symmetrical encryption algorithm such as DES, AES, et al. Assume that the length of each block is l_c bits. Because each sensor submits a message to $\mathcal{M}$ in each time slot, Thus, we have

$$\mathbf{\Phi}_{ic} = n \cdot \left(l_{id} + l_t + \left\lceil \frac{l_t + w}{l_c} \right\rceil \cdot l_c \right) \cdot L. \tag{3}$$

We then derive the query communication cost Φ_{qc} incurred by transmitting messages from $\mathcal{M}$ to the base station. Assume that there are δ queried sensor nodes in the cell. Because each queried sensor contributes only one message for a query, we then have

$$\mathbf{\Phi}_{qc} = \boldsymbol{\delta} \cdot \left(l_{id} + \left\lceil \frac{l_t + w}{l_c} \right\rceil \cdot l_c \right). \tag{4}$$

V. COMPLICATED SECURE MAX/MIN QUERY SCHEME

NSMQ works well when δ is small, which means that only a few sensor nodes are queried. However, if δ is very large, the returned messages from $\mathcal{M}$ will dramatically increase, which will cause Φ_{qc} to increase significantly. Because the messages are transmitted on the costly and low-rate on-demand wireless link, it is necessary to develop some alternatives to reduce Φ_{qc}, while the abilities of privacy protection and integrity verification remain unchanged.

In this section, we propose a complicated secure MAX/MIN query scheme, denoted as CSMQ, to reduce Φ_{qc}. The basic idea of CSMQ is to let $\mathcal{M}$ determine the very ciphertext satisfying the query request without knowing the real value of any of the collected data items, which will be sent to the base station by $\mathcal{M}$ alongside some proof information to check the integrity of the query result. For instance, assume that the cell C contains 4 sensor nodes whose ID set is $\{1, 2, 3, 4\}$, and that each of them generates two data items as shown in Fig. 2. When the query $Q_t = (\text{MAX}, t, C, \{1, 2, 3, 4\})$ is requested, $\mathcal{M}$ will return 4 messages to the base station in NSMQ, while only one message $<3, (t \parallel 9 \parallel \Psi)_{k_i}>$ will be returned in CSMQ, where Ψ can prove that 9 is the correct query result.

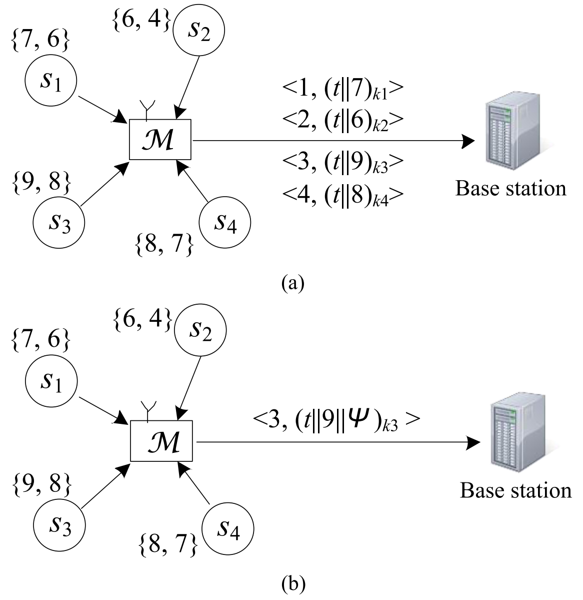


FIGURE 2. Examples of NSMQ and CSMQ. (a) NSMQ. (b) CSMQ.

To achieve CSMQ, there are two challenges as follows

- How to let $\mathcal{M}$ compare the collected data items without knowing their actual values.
- How to generate the proof information with the magic ability to verify the integrity of the query result.

In the remainder of this section, we apply comparable factors to solve the first challenge, which is based on the 0-1 encoding verification. The proof information named minor-node-set is introduced to solve the second challenge.

A. DEFINITIONS

We use the 0-1 encoding verification that was first proposed by Lin et al. to solve the classic millionaires' problem [28]. It can be utilized to compare data items without knowing their values.

Let $x = b_1 b_2 \ldots b_{w-1} b_w \in \{0, 1\}^w$ be a binary string with w bits. The 0-encoding and 1-encoding of x are denoted as $E_0(x) = \{b_1 b_2 \ldots b_{i-1} 1 | b_i = 0 \wedge 1 \leq i \leq w\}$ and $E_1(x) = \{b_1 b_2 \ldots b_i | b_i = 1 \wedge 1 \leq i \leq w\}$ respectively, where $|E_0(x)| + |E_1(x)| = w$. For two data items x and y, $x > y$ if and only if $E_1(x) \cap E_0(y) \neq \emptyset$; otherwise, $x \leq y$. Obviously, if the codes of x and y are of different types, they can be compared; otherwise, they cannot.

To improve the efficiency of computing the intersection, numeralization functions are usually applied to convert the 0-1 encoded binary strings into numbers. Thus, we adopt the same numeralization function N(*) as in [23], which ensures that for any two 0-encoding or 1-encoding binary strings a and b, $a = b$ if and only if N(p) =N(q). Additionally, we utilize HMAC to realize one-wayness and collision resistance in encoding the data. We denote the HMAC function as $H_g(*)$, where g is the secret key of HMAC, which is only shared in sensor nodes.

Definition 1 [Comparable Factors (c-Factors)]: For data x, after applying 0-1 encoding, numeralization and HMAC, the two generated code sets are denoted as the comparable factors of x, which are abbreviated as c-factors. We denote $CF_0(x)$ and $CF_1(x)$ as the 0-encoding and 1-encoding c-factors respectively, i.e., $CF_0(x) = H_g(\mathrm{N}(E_0(x)))$, $CF_1(x) = H_g(\mathrm{N}(E_1(x)))$.

According to the 0-1 encoding verification properties and Definition 1, we can easily established the following Lemma 1.

Lemma 1: For data x and y,

1) If $CF_1(x) \cap CF_0(y) \neq \emptyset$, then $x > y$; otherwise $x \leq y$.
2) $0 \leq |CF_0(x)| \leq w \wedge 0 \leq |CF_1(x)| \leq w \wedge |CF_0(x)| + |CF_1(x)| = w$

As shown in Lemma 1, any two data items can be compared using their c-factors instead of their real values.

Definition 2 (Minor-Node-Set:) Given a sensor s_i in cell C, the minor-node-set of s_i, denoted as Ω_i, is the set of the IDs of sensor nodes whose maximum is smaller than s_i in C, and we have

$$\Omega_i = \{j \in \Gamma | j \neq i \wedge d_j < d_i\} \tag{5}$$

where d_i and d_j are the maximums of s_i and s_j in t, respectively.

As shown in Definition 2, the maximum of s_i is larger than any other sensor node whose IDs are in Ω_i and $\Omega_i \subset \Gamma$. Such a minor-node-set is exactly the critical evidence for query result verification in CSMQ, which will be discussed in the next section.

B. DATA SUBMISSION

In data submission, each sensor first generates the minor-node-set, and then submits the encrypted data embedding the minor-node-set and the collected data item to $\mathcal{M}$. The detailed procedures are as in the following two phases.

Phase 1:

- For each sensor s_i in C, after obtaining the maximum d_i in t, s_i computes the c-factors of d_i, $CF^0(d_i)$ and $CF^1(d_i)$, where the key g is shared with all sensor nodes but not $\mathcal{M}$. Then, s_i broadcasts the following message in cell C, in which $min\{A, B\}$ represents the set of A or B that has fewer items and $*$ represents all other sensor nodes in the cell.

$$s_i \rightarrow * :< i, min\{CF^0(d_i), CF^1(d_i)\} >$$

- When s_i receives the message $< j, min\{CF^0(d_j), CF^1(d_j)\} >$ broadcasted by s_j, s_i compares its maximum d_i with d_j according to Lemma 1. If $d_i > d_j$, add $j$to Ω_i.
- Upon receiving the broadcasted messages from all the other sensor nodes in C, each sensor will generate its minor-node-set completely. Taking the scenario in Fig. 2 as an example, the *minor-sets* of s_1, s_2, s_3 and s_4 are $\Omega_1 = \{2\}$, $\Omega_2 = \emptyset$, $\Omega_3 = \{1, 2, 4\}$ and $\Omega_4 = \{1, 2\}$, respectively.

Phase 2:

- Each sensor s_i in C, after generating its minor-node-set Ω_i, submits the following message to $\mathcal{M}$ in which d_i is the maximum of s_i in t, $CF^0(d_i)$ and $CF^1(d_i)$ are the computed c-factors of d_i, and || is the concatenation operator.

$$s_i \rightarrow \mathcal{M} \ :< i, t, (t||d_i||\Omega_i)_{k_i}, CF^0(d_i), CF^1(d_i) >$$

As we discussed in Section IV, it is also infeasible for $\mathcal{M}$ to obtain the collected data item because it is encrypted. Meanwhile, because the HMAC function has one-wayness and collision resistance properties and $\mathcal{M}$ has no idea of the HMAC key, it is computationally infeasible for $\mathcal{M}$ to obtain the corresponding values from the c-factors. Therefore, the privacy of collected data items can be preserved from $\mathcal{M}$ even if they are compromised.

C. QUERY PROCESSING

When $\mathcal{M}$ receives a query $Q_t = (\mathrm{MAX}, t, C, \Gamma_t)$ from the base station, it processes Q_t on its stored data following the two steps below:

- Loads the stored data items $\{(t||d_i||\Omega_i)_{k_i}, CF^0(d_i), CF^1(d_i)|i \in \Gamma_t\}$ that are received from the queried sensor nodes in t.
- Compares the data items with the loaded c-factors according to Lemma 1, and finds the ciphertext whose corresponding data is the maximum. We assume the ciphertext is$(t||d_i||\mathbf{\Omega}_i)_{k_i}$. Then, $\mathcal{M}$ transmits the following response to the base station.

$$\mathcal{M} \rightarrow \text{base station} \ :< i, (t||d_i||\Omega_i)_{k_i} >$$

Upon receiving the above response $(t||d_i||\Omega_i)_{k_i}$, the base station decrypts it with the key k_i shared with s_i to obtain the query result $R_t = d_i$, which is the maximum of the data item collected by the queried sensor nodes in t. Obviously, there is only one response message returned from $\mathcal{M}$ to the base station.

Lemma 2: If $\mathcal{M}$ follows the query processing scheme and returns the correct response, which is assumed to be $< i, (t||d_i||\Omega_i)_{k_i} >$, then we have

$$|\Omega_i| \in \{|\Gamma_t| - 1, |\Gamma_t|, \ldots, n - 1\}. \tag{6}$$

Proof: According to the assumption, s_i is the queried sensor whose ID is in Γ_t and all the data collected by the other queried sensor nodes in Γ_t are smaller than d_i. Definition 2 shows that Ω_i is composed of the IDs of sensor nodes whose corresponding collected data is smaller than d_i in the cell. Thus, we have Γ_t−{i}$\subseteq \Omega_i \subseteq \Gamma$−{i}, which indicates that $|\Omega_i| \in \{|\Gamma_t| - 1, |\Gamma_t|, \ldots, n\text{-}1\}$. **(End)**

D. QUERY RESULT VERIFICATION

Upon receiving a message $< i, c_i >$ from $\mathcal{M}$ for the query Q_t, where c_i is the ciphertext, the base station decrypts c_i with the corresponding key k_i to obtain the embedded time slot number $\tau(c_i)$, the collected data item $\nu(c_i)$ and the minor-node-set $\zeta(c_i)$. Only if the following conditions stand, $\nu(c_i)$ is the correct query result and satisfies the integrity requirement.

Condition 1: $i \in \Gamma_t \wedge \tau(c_i) = t$

Condition 2: $\Gamma_t - \{i\} \subseteq \zeta(c_i)$

Lemma 3: The condition 1 and condition 2 can verify the integrity of the query results.

Proof: If a query result is falsified, which makes condition 1 not hold, it means that the data items returned by $\mathcal{M}$ are not contributed in t or are not collected by the queried sensor nodes in Γ_t. If condition 2 does not hold, it means that the maximum data items collected by all the queried sensor nodes of Γ_t could not be returned. Condition 1 is to check the authenticity of the received query response, while Condition 2 is to check the correctness of the query result. When a falsified message is returned to the base station, the time slot number, the collected data items and the minor-node-set are obtained through decryption. It is easy for the base station to check whether condition 1 and condition 2 both stand. As a result, we can see that condition 1 and condition 2 can verify the integrity of the query results. **(End)**

Taking the scenario of Fig. 2 as an example again, the correct response that should be returned from $\mathcal{M}$ for the query $Q_t = (\text{MAX}, t, C, \{1, 2, 3, 4\})$ is $<3,\ (t||9||\{1, 2, 4\})_{k_3}>$. Because the keys owned by the sensor nodes are unknown to it, $\mathcal{M}$ cannot modify the collected data or the minor-node-set in the ciphertext $(t||9||\{1, 2, 4\})_{k_3}$ without being detected in the first condition verification. The only option left for $\mathcal{M}$ is to replace $<3,\ (t||9||\{1, 2, 4\})_{k_3}>$ with the responses contributed by other sensor nodes. Assume that $<4,\ (t||8||\{1, 2\})_{k_4}>$ is chosen to make the replacement. Although it will pass the first condition verification, it cannot escape being detected during the second condition verification, where $\Gamma_t - \{4\} = \{1, 2, 3\} \subseteq \{1, 2\}$ happens.

E. PERFORMANCE ANALYSIS

Now we derive Φ_{ic} incurred by CSMQ, which can be divided into two parts, denoted as Φ_{icb} and Φ_{ics}. The former is incurred by data broadcasting in the cell, while the latter is introduced by data submission to $\mathcal{M}$. We assume that the simplest broadcast scheme is used, in which each node forwards a received broadcast packet once [33]. The notation of n, l_{id}, l_t, w, L and δ are the same as in Section IV. In addition, we assume that a sensor node s_i broadcasts a c-factor having ρ_i HMAC codes, each of which is of l_h bits. Then, we have

$$\boldsymbol{\Phi}_{icb} = n \cdot \sum_{i=1}^{n} (l_{id} + \rho_i \cdot l_h). \tag{7}$$

We assume that the HMAC code is randomly distributed and each collected data item has w binary bits. According to Lemma 1, for any collected data item x, $0 \leq |CF_0(x)| \leq w \wedge 0 \leq |CF_1(x)| \leq w \wedge |CF_0(x)| + |CF_1(x)| = w$ holds; then we can derive that the broadcasted c-factor $min\{CF^0(x), CF^1(x)\}$ has approximately $w/4$ HMAC codes on average, i.e., $\rho_i \approx w/4$. Therefore, Φ_{icb} can be approximated as follows:

$$\boldsymbol{\Phi}_{icb} \approx n^2 \cdot (l_{id} + w/4 \cdot l_h) \tag{8}$$

Comparing with NSMQ, c-factors and minor-node-set information for each sensor node are additionally submitted to $\mathcal{M}$ in CSMQ, and then we have

$$\begin{aligned}\boldsymbol{\Phi}_{ics} &= \left(n \cdot (l_{id} + l_t + w \cdot l_h) + \left\lceil \frac{\sum_{i=0}^{n-1} (l_t + w + i \cdot l_{id})}{l_c} \right\rceil \cdot l_c \right) \cdot L \\ &= \left(n \cdot (l_{id} + l_t + w \cdot l_h) + \left\lceil \frac{\frac{(n-1)(n-2)}{2} \cdot l_{id} + (l_t + w) \cdot n}{l_c} \right\rceil \cdot l_c \right) \cdot L\end{aligned} \tag{9}$$

Adding Φ_{icb} and Φ_{ics}, we have

$$\begin{aligned}\boldsymbol{\Phi}_{ic} &\approx n^2 \cdot (l_{id} + w/4 \cdot l_h) \\ &\quad + \left(n \cdot (l_{id} + l_t + w \cdot l_h) + \left\lceil \frac{\frac{(n-1)(n-2)}{2} \cdot l_{id} + (l_t + w) \cdot n}{l_c} \right\rceil \cdot l_c \right) \cdot L\end{aligned} \tag{10}$$

We then derive Φ_{qc} incurred by returning the message from $\mathcal{M}$ to the base station. Assume that the minor-node-set in the returned response has μ IDs. Then, we have

$$\boldsymbol{\Phi}_{qc} = l_{id} + \left\lceil \frac{l_t + w + \mu \cdot l_{id}}{l_c} \right\rceil \cdot l_c. \tag{11}$$

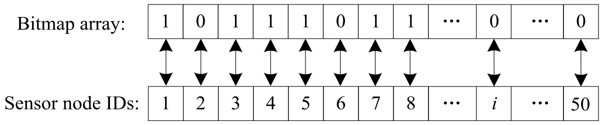
Here, $\mu \in \{\delta - 1, \delta, \ldots, n - 1\}$ holds, according to Lemma 2.

VI. OPTIMIZING CSMQ

Although CSMQ can reduce Φ_{qc} efficiently, its Φ_{ic} increases dramatically compared to NSMQ because each sensor node must broadcast a c-factor and submit the extra c-factors and encrypted minor-node-set. In this section, we concentrate on the optimizations for CSMQ to reduce its Φ_{ic}. We first use a bit-mapping method to compress the minor-node-sets, which are both transmitted from sensor nodes to $\mathcal{M}$ and from $\mathcal{M}$ to the base station; therefore Φ_{ic} and Φ_{qc} will both be reduced by such optimization. Then, we apply the random c-factor selection strategy of [24] to reduce Φ_{ic} further. We denote the optimized CSMQ as OSMQ.

A. COMPRESSING MINOR-NODE-SETS

A bit-mapping mechanism is introduced to reduce the length of the minor-node-set of each sensor node. Given a sensor node s_i, we map its minor-node-set Ω_i to an n-bit bitmap



Bitmap array:	1	0	1	1	1	0	1	1	···	0	···	0
Sensor node IDs:	1	2	3	4	5	6	7	8	···	i	···	50

FIGURE 3. **Bitmap array.**

array, denoted as $bm(\Omega_i)$, which satisfies $j \in \Omega_i$ if the j-th bit of the array is 1, otherwise $j \notin \Omega_i$.

For instance, assuming that there are 50 sensor nodes in cell C and the minor-node-set of s_i is $\Omega_i = \{1, 3, 4, 5, 7, 8\}$, then we have the corresponding bitmap array, as shown in Fig. 3. If each sensor ID is 16 bits in length, the set Ω_i will take up 96 bits while $bm(\Omega_i)$ only requires 50 bits. It is obvious that as the number of IDs in Ω_i increases, more space will be saved using the bit-mapping mechanism. As a result, after applying the compressing minor-node-set optimization, Φ_{ic} and Φ_{qc} will both be reduced because the minor-node-set must be transmitted in both the data submission and the query processing procedures.

The basic precondition of the bit-mapping mechanism is that each sensor node in the same cell uses the same bit-mapping mechanism, which is also shared with the base station. It is not complex to build up such a bit-mapping mechanism. In the initialization associated with deploying the network, the base station can map the sensor IDs into a corresponding bitmap array and synchronize the bit-mapping information with the sensor nodes.

B. RANDOM c-FACTOR SELECTION

We introduce the random c-factor selection that was proposed in our prior work [24] to decrease Φ_{ic}. In data submission, every sensor node randomly selects the 0-encoding or 1-encoding c-factor of its collected data in each time slot to submit to $\mathcal{M}$. In query processing, $\mathcal{M}$ also uses Lemma 1 to determine the query response for the base station. Because there is only a 0-encoding or 1-encoding c-factor for each collected data item stored in $\mathcal{M}$, the query response from $\mathcal{M}$ consists of the minimal set of candidate ciphertext R. There are one or more pieces of ciphertext in R, and the types of the corresponding c-factors of the data embedded in R are the same. Based on the experimental statistics in [24], the mean quantity of the pieces of ciphertext in R is close to 2 when the amount of test samples becomes large, which agrees with the theoretical analysis.

Compared to CSMQ, after applying random c-factor selection, Φ_{ic} can be decreased significantly, because only half of the c-factors need to be submitted from the sensor nodes to $\mathcal{M}$. It seems to be reasonable that Φ_{qc} could be increased because of the growth of the quantity of returned ciphertext in the query response. However, the optimization of compressing the minor-node-set can decrease Φ_{qc}. The larger the network scale is, which means more sensor nodes are deployed in the network, the more significant the decrease of Φ_{qc} that will be incurred. Furthermore, the candidate ciphertext returned to the base station from $\mathcal{M}$ is only approximately 2 on average. As a result, Φ_{qc}can be reduced after applying the above two optimizations, especially when the number of sensor nodes deployed in the network is large.

In addition, we also introduce a hash-based code compression method to reduce the communication cost, as proposed in our prior work [23].

C. APPLYING OPTIMIZATIONS

Based on the above optimizations, we transform CSMQ into OSMQ as follows.

1) In the phase 2 data submission of Section V, the message transmitted from s_i to $\mathcal{M}$ is changed into the following:

$$s_i \rightarrow \mathcal{M} :< i, t, (t||d_i||bm(\Omega_i))_{k_i}, rnd\{CF_0(d_i), CF_1(d_i)\} >$$

where $rnd\{*\}$ indicates a random element selection from a set.

2) In the query processing of Section V, the response message transmitted from $\mathcal{M}$ to the base station is changed into the following:

$$\mathcal{M} \rightarrow \text{base station} : \{< i, (t||d_i||bm(\Omega_i))_{k_i} > | \\ (t||d_i||bm(\Omega_i))_{k_i} \in R\}$$

where R is the minimal set of candidate ciphertext that can be determined by $\mathcal{M}$ based on the algorithms in [13].

Upon receiving the above response $\{<i, (t||d_i||bm(\Omega_i))_{k_i}> |(t||d_i||bm(\Omega_i))_{k_i} \in R\}$, the base station decrypts the ciphertext with the keys shared with the corresponding sensor nodes to obtain the plaintext data items. Then, the query result R_t is the maximum of these data items, which can be easily obtained.

3) In the query result verification of Section V, the verification procedures are changed as follows. Assume that the base station receives $W = \{< i, c_i > | c_i \in R\}$ from $\mathcal{M}$, where c_i is the ciphertext. The base station decrypts c_i with the corresponding key k_i to obtain the embedded time slot number $\tau(c_i)$, the collected data item $v(c_i)$ and the bitmap array of the minor-node-set $\zeta(c_i)$. If the following two conditions hold, the obtained query result is correct and satisfies the integrity requirement. Here, & represents the bitwise operator AND.

Condition 1: $\forall < i, c_i > \in \boldsymbol{W} \rightarrow i \in \boldsymbol{\Gamma}_t \wedge \boldsymbol{\tau}(c_i) = t$

Condition 2: $\exists < i, c_i > \in \boldsymbol{W} \rightarrow \left(bm\left(\boldsymbol{\Gamma}_t - \{i\}\right) \& \zeta'\left(c_i\right)\right) = bm\left(\boldsymbol{\Gamma}_t - \{i\}\right)$

Condition 1 is to check whether all returned messages are indeed contributed by the queried sensor nodes in t, while condition 2 is to check whether there is a message $<i, c_i>$ that it is contributed by a queried sensor in Γ_t and its embedded data $v(c_i)$ is larger than the data collected by other sensor nodes in Γ_t.

With the optimizations, OSMQ can save the space of the minor-node-sets and c-factors, which will reduce the communication cost of the networks. In addition, because the security settlements are kept the same as with CSMQ, OSMQ can preserve the privacy of collected data and verify the integrity of the query results.

D. COMMUNICATION COST ANALYSIS

In this section, we will discuss the communication cost Φ_{ic} and Φ_{qc} of OSMQ. We assume that the length of the bitmap array is equal to the quantity of sensor nodes, i.e., n. According to Lemma 1, for any data x, $|CF_0(x)| + |CF_1(x)| = w$ holds, which means that the pair of c-factors of x have w items. It is reasonable that the random selected c-factor of a collected data item has $w/2$ items on average. Other parameters are the same as above. Then, we have

$$\mathbf{\Phi}_{ic} = n \cdot \left(l_{id} + l_t + w/2 \cdot l_h + \left\lceil \frac{n + l_t + w}{l_c} \right\rceil \cdot l_c \right) \cdot L + n^2 \cdot (l_{id} + w/4 \cdot l_h) \quad (12)$$

and

$$\mathbf{\Phi}_{qc} \approx 2 \cdot (l_{id} + \left\lceil \frac{n + l_t + w}{l_c} \right\rceil \cdot l_c). \quad (13)$$

VII. PERFORMANCE EVALUATIONS

As the first work on secure MAX/MIN queries supporting privacy preservation and query result verification, we evaluate the performance of our proposed NSMQ, CSMQ and OSMQ. We implement these schemes on the simulator of [34] with the Intel lab data set [35]. We evaluate and analyze the in-cell communication cost Φ_{ic} and the query communication cost Φ_{qc} of these three schemes. We also assume that the packet transmissions are both collision-free and error-free in our experiments.

The evaluations are conducted on a PC with a P4 2.6GHz CPU and 4G memory running the Windows 7 operating system. We carry out evaluations on a MAX query in a cell with n sensor nodes and a master node. The placement of the sensor nodes follows a uniform distribution over a $100 \times 100\text{m}^2$ area, and the radius of sensor communication is assumed to be 10m. The query region covers the whole cell, which means that all sensor nodes are queried. Default parameters used in the experiment are shown in Table 2. In each measurement, we generate 10 different networks with different IDs. In each generated network, sensor nodes are randomly distributed. The result is based on the average of the 10 different networks.

TABLE 2. Default parameters setting.

Para.	*n*	l_t	l_{id}	*w*	l_c
Val.	80	32 bits	24 bits	8 bits	256 bits

A. EVALUATIONS ON Φic

We take the network ID, the node number n, the length of the node ID l_{id} and the length of the collected data item w as the independent variables to measure Φ_{ic} for NSMQ, CSMQ and OSMQ. The results are shown in Fig. 4 - Fig. 7.

Fig. 4 shows that Φ_{ic} of NSMQ, CSMQ and OSMQ are all uniformly distributed in the different networks. NSMQ and OSMQ are obviously lower than CSMQ. NSMQ has the lowest Φ_{ic}, and it takes only 3.68% of CSMQ and 13.19% of OSMQ. Compared with CSMQ, OSMQ saves approximately 71.55% cost on average. The reason is that each sensor node needs to submit only a piece of ciphertext in NSMQ, but in CSMQ, much extra data, including c-factors and minor-node-sets, will be submitted for secure comparison and verifica-

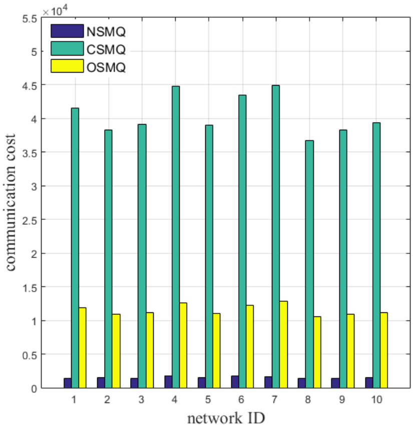


FIGURE 4. Φ_{ic} **versus Network ID.**

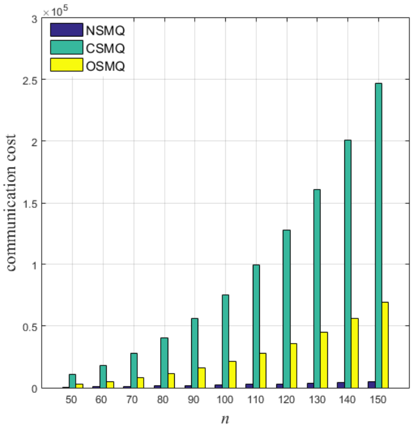


FIGURE 5. Φ_{ic} **versus** n**.**

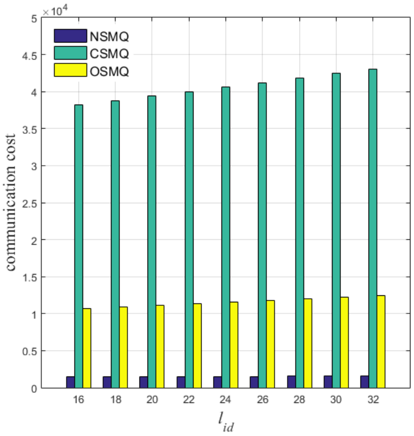


FIGURE 6. Φ_{ic} **versus** l_{id}**.**

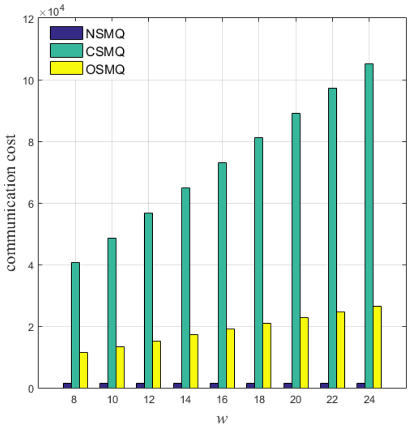


FIGURE 7. Φ_{ic} versus w.

tion proof generation, which makes Φ_{ic} of CSMQ obviously larger than the others. OSMQ optimizes CSMQ by decreasing the submitted c-factors and compressing the minor-node-sets, which makes it better than CSMQ in Φ_{ic}.

According to the results shown in Fig. 5 - Fig. 7, we can see that Φ_{ic} of NSMQ, CSMQ and OSMQ are all increased as n, l_{id} and w grow, but the growth of NSMQ and OSMQ is obviously slower than that of CSMQ. Φ_{ic} of NSMQ and OSMQ are apparently lower than CSMQ, and NSMQ performs the best. On average, NSMQ saves 96.62%, 96.24% and 97.71% compared to CSMQ in Fig. 5 - Fig. 7, respectively, while OSMQ saves 71.65%, 71.59% and 73.68% compared to CSMQ, respectively. The reason is similar to Fig. 4. In addition, from Fig. 5 - Fig. 7, we have that n is the most important factor for Φ_{ic}, while w and l_{id} are the lesser and least ones, respectively. The reason is that the network scale becomes large as n grows, and the broadcast in CSMQ and OSMQ is quadratic with n, which makes Φ_{ic} increase significantly, especially in CSMQ. Φ_{ic} in all three schemes is approximately linear with l_{id} and w, based on the equations in communication cost analysis.

B. EVALUATIONS ON Φqc

To evaluate Φ_{qc} of NSMQ, CSMQ and OSMQ, we take n, l_{id} and w as the independent variables. The results are shown in Fig. 8 - Fig. 10.

According to Fig. 8 - Fig. 10, we can see that the Φ_{qc} of NSMQ, CSMQ and OSMQ all increase as n, l_{id} and w grow. The increments are obvious for n and l_{id} in NSMQ and CSMQ, but hardly observable in the other situation. Φ_{qc} of OSMQ is visibly lower than CSMQ and NSMQ. On average, OSMQ saves 97.75%, 97.50% and 97.50% compared to NSMQ in Fig. 8 - Fig. 10, respectively, while it saves 75.73%, 75.31% and 72.97% compared to CSMQ, respectively. The reason is that $\mathcal{M}$ needs to return n pieces of ciphertext in NSMQ, while there is only one and on average two pieces in CSMQ and OSMQ, respectively. Although the returned ciphertext in OSMQ is probably more than that of CSMQ, Φ_{qc} of OSMQ is lower than CSMQ because the embedded minor-node-set of the returned ciphertext of OSMQ is compressed into a bitmap array, as opposed to in CSMQ.

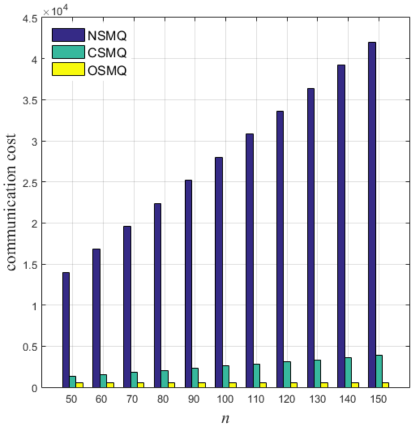


FIGURE 8. Φ_{qc} versus n.

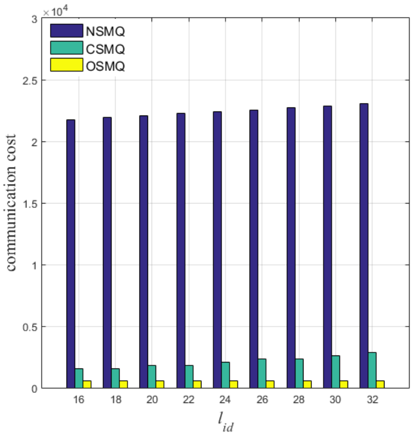


FIGURE 9. Φ_{qc} versus l_{id}.

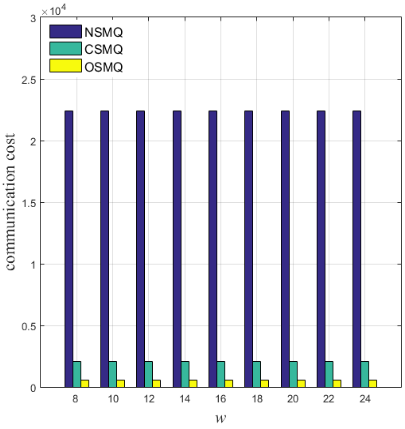


FIGURE 10. Φ_{qc} versus w.

As a result, among our proposed schemes, we can see that NSMQ is the most cost-saving scheme for Φ_{ic}, but it is very wasteful for Φ_{qc}. CSMQ significantly saves Φ_{qc} over NSMQ, but its Φ_{ic} dramatically increases, and OSMQ, an optimized version of CSMQ, is the balanced scheme for Φ_{ic} and Φ_{qc}.

VIII. CONCLUSION

Secure MAX/MIN query processing is an important issue in wireless sensor networks, and it can be utilized in fields where security is necessary. In this article, we propose secure MAX/MIN query processing in a TWSN that is the first work to solve the problems of data privacy protection and query result integrity verification at the same time. Three schemes, NSMQ, CSMQ and OSMQ, are designed to achieve secure MAX/MIN queries. They can prevent a compromised master node from peeking at the hosted data and check whether the query result satisfies the integrity requirement. Built upon symmetric encryption and hash-based message authentication coding primitives, our proposed schemes are effective for resource-constrained sensor networks.

REFERENCES

[1] O. Gnawali *et al.*, "The Tenet architecture for tiered sensor networks," in *Proc. 4th Int. Conf. Embedded Netw. Sensor Syst. (SenSys)*, Boulder, CO, USA, Oct./Nov. 2006, pp. 153–166.

[2] D. Yang, S. Misra, X. Fang, G. Xue, and J. Zhang, "Two-tiered constrained relay node placement in wireless sensor networks: Computational complexity and efficient approximations," *IEEE Trans. Mobile Comput.*, vol. 11, no. 8, pp. 1399–1411, Aug. 2012.

[3] L. Dong, J. Zhu, X. Zhang, H. Chen, C. Li, and H. Sun, "SEMR: Secure and efficient multi-dimensional range query processing in two-tiered wireless sensor networks," in *Proc. WAIM*, 2015, pp. 520–524.

[4] Y.-T. Tsou, C.-S. Lu, and S.-Y. Kuo, "SER: Secure and efficient retrieval for anonymous range query in wireless sensor networks," *Comput. Commun.*, vol. 108, pp. 1–16, Aug. 2017.

[5] H. Dai, Q. Ye, G. Yang, J. Xu, and R. He, "CSRQ: Communication-efficient secure range queries in two-tiered sensor networks," *Sensors*, vol. 16, no. 2, p. 259, 2016.

[6] X. Zhang, L. Dong, H. Peng, H. Chen, D. Li, and C. Li, "Achieving efficient and secure range query in two-tiered wireless sensor networks," in *Proc. IWQoS*, Hong Kong, May 2014, pp. 380–388.

[7] F. Chen and A. X. Liu, "Privacy- and integrity-preserving range queries in sensor networks," *IEEE/ACM Trans. Netw.*, vol. 20, no. 6, pp. 1774–1787, Dec. 2012.

[8] Y. T. Tsou, C. S. Lu, and S. Y. Kuo, "Privacy-and integrity-preserving range query in wireless sensor networks," in *Proc. GLOBECOM*, Anaheim, CA, USA, Dec. 2012, pp. 328–334.

[9] Y. Yi, Y. Yi, R. Li, F. Chen, A. X. Liu, and Y. Lin, "A digital watermarking approach to secure and precise range query processing in sensor networks," in *Proc. INFOCOM*, Turin, Italy, 2013, pp. 1950–1958.

[10] X. Zhang, L. Dong, H. Peng, H. Chen, S. Zhao, and C. Li, "Collusion-aware privacy-preserving range query in tiered wireless sensor networks," *Sensors*, vol. 14, no. 12, pp. 23905–23932, 2014.

[11] L. Dong, X. Chen, J. Zhu, H. Chen, K. Wang, and C. Li, "A secure collusion-aware and probability-aware range query processing in tiered sensor networks," in *Proc. SRDS*, Montreal, QC, Canada, Sep./Oct. 2015, pp. 110–119.

[12] H. Dai, Q. Ye, X. Yi, R. He, G. Yang, and J. Pan, "VP2RQ: Efficient verifiable privacy-preserving range query processing in two-tiered wireless sensor networks," *Int. J. Distrib. Sensor Netw.*, vol. 12, no. 10, pp. 1–15, 2016.

[13] R. Li *et al.*, "Privacy and integrity preserving top-k query processing for two-tiered sensor networks," *IEEE/ACM Trans. Netw.*, to be published.

[14] X. Liao and J. Li, "Privacy-preserving and secure top-*k* query in two-tier wireless sensor network," in *Proc. GLOBECOM*, Anaheim, CA, USA, Dec. 2012, pp. 335–341.

[15] C.-M. Yu, G.-K. Ni, I.-Y. Chen, E. Gelenbe, and S.-Y. Kuo, "Top-*k* query result completeness verification in sensor networks," in *Proc. ICC*, Budapest, Hungary, Jun. 2013, pp. 1026–1030.

[16] C.-M. Yu, G.-K. Ni, I.-Y. Chen, E. Gelenbe, and S.-Y. Kuo, "Top-*k* query result completeness verification in tiered sensor networks," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 1, pp. 1026–1030, Jan. 2013.

[17] R. Zhang, J. Shi, Y. Zhang, and X. Huang, "Secure top-*k* query processing in unattended tiered sensor networks," *IEEE Trans. Veh. Technol.*, vol. 63, no. 9, pp. 4681–4693, Nov. 2014.

[18] X. Ma, H. Song, J. Wang, J. Gao, and G. Min, "A novel verification scheme for fine-grained top-*k* queries in two-tiered sensor networks," *Wireless Pers. Commun.*, vol. 75, no. 3, pp. 1809–1826, 2014.

[19] H. Peng, X. Zhang, H. Chen, Y. Wu, Y. Wu, and J. Zeng, "Enable privacy preservation and result verification for top-*k* query in two-tiered sensor networks," in *Proc. IEEE TrustCom/BigDataSE/ISPA*, Helsinki, Finland, Aug. 2015, pp. 555–562.

[20] Y.-T. Tsou, Y.-L. Hu, Y. Huang, and S.-Y. Kuo, "PCTopk: Privacy- and correctness-preserving functional top-*k* query on un-trusted data storage in two-tiered sensor networks," in *Proc. SRDS*, Nara, Japan, Oct. 2014, pp. 191–200.

[21] H. Dai, G. Yang, H. Huang, and F. Xiao, "Efficient verifiable top-*k* queries in two-tiered wireless sensor networks," *KSII Trans. Internet Inf. Syst.*, vol. 9, no. 6, pp. 2111–2131, 2015.

[22] Y. Yao, N. Xiong, J. H. Park, L. Ma, and J. Liu, "Privacy-preserving max/min query in two-tiered wireless sensor networks," *Comput. Math. Appl.*, vol. 65, no. 9, pp. 1318–1325, May 2013.

[23] H. Dai, G. Yang, and X. Qin, "EMQP: An energy-efficient privacy-preserving MAX/MIN query processing in tiered wireless sensor networks," *Int. J. Distrib. Sensor Netw.*, vol. 2013, no. 1, pp. 1492–1519, 2013.

[24] H. Dai, T. Wei, Y. Huang, J. Xu, and G. Yang, "Random secure comparator selection based privacy-preserving MAX/MIN query processing in two-tiered sensor networks," *J. Sensors*, vol. 2016, no. 6, pp. 1–13, 2015.

[25] J. Cheng, H. Yang, S. H. Y. Wong, P. Zerfos, and S. Lu, "Design and implementation of cross-domain cooperative firewall," in *Proc. ICNP*, Beijing, China, Oct. 2007, pp. 284–293.

[26] A. X. Liu and F. Chen, "Collaborative enforcement of firewall policies in virtual private networks," in *Proc. PODC*, Toronto, ON, Canada, Aug. 2008, pp. 95–104.

[27] H. Krawczyk, R. Canetti, and M. Bellare, *HMAC: Keyed-Hashing for Message Authentication*, document RFC 2104, 1997.

[28] H.-Y. Lin and W.-G. Tzeng, "An efficient solution to the millionaires' problem based on homomorphic encryption," in *Proc. ACNS*, New York, NY, USA, Jun. 2005, pp. 456–466.

[29] L. Ertaul and V. Kedlaya, "Computing aggregation function minimum/maximum using homomorphic encryption schemes in wireless sensor networks (WSNs)," in *Proc. Int. Conf. Wireless Netw.*, Las Vegas, NV, USA, Jun. 2007, pp. 186–192.

[30] M. M. Groat, W. Hey, and S. Forrest, "KIPDA: k-indistinguishable privacy-preserving data aggregation in wireless sensor networks," in *Proc. 30th IEEE Int. Conf. Comput. Commun.*, Shanghai, China, Apr. 2011, pp. 2024–2032.

[31] Y. Yao, L. Ma, and J. Liu, "Privacy-preserving MAX/MIN aggregation in wireless sensor networks," *Adv. Inf. Sci. Service Sci.*, vol. 73, no. 3, p. 23, 2012.

[32] B. K. Samanthula, W. Jiang, and S. Madria, "A probabilistic encryption based MIN/MAX computation in wireless sensor networks," in *Proc. 14th Int. Conf. Mobile Data Manage.*, Milan, Italy, Jun. 2013, pp. 77–86.

[33] J. Wu and F. Dai, "A generic distributed broadcast scheme in ad hoc wireless networks," *IEEE Trans. Comput.*, vol. 53, no. 10, pp. 1343–1354, Oct. 2004.

[34] A. Coman, M. A. Nascimento, and J. Sander, "A framework for spatio-temporal query processing over wireless sensor networks," in *Proc. 1st Int. Workshop Data Manage. Sensor Netw. (DMSN)*, Toronto, ON, Canada, 2004, pp. 104–110.

[35] P. Bodik, W. Hong, C. Guestrin, S. Madden, M. Paskin, and R. Thibaux. *Intel Lab Data*, accessed on Feb. 28, 2004. [Online]. Available: http://db.csail.mit.edu/labdata/labdata.html

HUA DAI was born in 1982. He is currently an Associate Professor with the Nanjing University of Posts and Telecommunications. His research interests include data management and security and database security. He is a member of CCF.

MIN WANG was born in 1992. She received the M.S. degree from the Nanjing University of Posts and Telecommunications. Her research interests include data management and security in wireless sensor networks.

XUN YI was born in 1967. He is currently a Professor and the Ph.D. Supervisor with the Royal Melbourne Institute of Technology University. His research interests include information security and distributed data processing.

GENG YANG was born in 1961. He is currently a Professor and the Ph.D. Supervisor with the Nanjing University of Posts and Telecommunications. His research interests include cloud computing and security, data security, and privacy protection. He is the Senior Member of CCF.

JINGJING BAO was born in 1993. She received the M.S. degree from the Nanjing University of Posts and Telecommunications. Her research interests include data management and security in wireless sensor networks.