IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Revue d'Intelligence Artificielle
Vol. 33, No. 2, April, 2019, pp. 145-149
Journal homepage: http://iieta.org/journals/ria

A Secure Query Protocol for Multi-layer Wireless Sensor Networks Based on Internet of Things

Tao Lin[1,2], Peng Wu[1,3*], Fengmei Gao[2], Linhong Wang[2]

[1] School of Automation, Chongqing University, Chongqing 400044, China
[2] Chongqing College of Electronic Engineering, Chongqing 401331, China
[3] Chongqing Chuanyi Automation Co., Ltd., Chongqing 401121, China

Corresponding Author Email: pwu@cqu.edu.cn

https://doi.org/10.18280/ria.330210

Received: 17 February 2019
Accepted: 10 March 2019

Keywords:
wireless sensor network (WSN), multi-layer, secure query protocol, Internet of Things (IoT)

ABSTRACT

This paper puts forward a secure query protocol for wide-range multi-layer WSNs based on the Internet of Things (IoT), which protects data transmission from three aspects: network security, attack mode and privacy protection. Drawing the merits from the Range Doppler (R-D) algorithm and Z-O encoding, our protocol only needs to compare the perceptual data nodes with the radius of the query interval once, which greatly improves the query efficiency. The embedded Z-O codes can transform the comparison into solving the intersection of the two sets. In addition, the hash-based message authentication code (HMAC) was introduced to achieve unidirectional protection of the original data. The simulation results show that our protocol outperforms the traditional protocols in query time, coding length and bucket mechanism, and enjoys high power efficiency, low storage cost and good query accuracy.

1. INTRODUCTION

Wireless sensor network (WSN) is an emerging intelligent private network with specific functions. It is the foundation of the Internet of Things (IoT) applications. The IoT is the network of physical objects, which connects and exchanges data over the Internet.

However, the WSN technology is not mature enough to withstand repeated attacks, which causes privacy leaks in data transmission and query. Most query protocols for the WSN involve a limited number of nodes and protect privacy with traditional maintenance strategy for network security, failing to fully consider data security.

Traditional secure query protocols are constrained by environmental factors, node resources and network topology. The attempts to improve these protocols mainly focus on the features of attack modes, without addressing new security threats like attacks on clustered and multi-layer networks [1-4].

With the proliferation of the WSN, it is a research hotspot to work out a low-cost, energy-efficient and high-precision secure query protocol for the network. Such a protocol should boast high fault tolerance, forward query requests via reasonable nodes, and identify the information forged by the attacker. The protocol design has been analyzed by many scholars [5-9].

Drawing on existing protocols, this paper puts forward a secure query protocol for wide-range multi-layer WSNs in the IoT. This protocol protects data transmission from three aspects: network security, attack mode and privacy protection. The research results provide theoretical reference for similar projects.

2. PROTOCOL DESIGN

2.1 Traditional methods for privacy protection and integrity verification

Traditionally, the privacy of WSN data transmission is typically protected by the bucket mechanism and the SafeQ protocol, while the data integrity is usually verified by the encoding number (E-N) mechanism and spatiotemporal crosscheck.

The bucket mechanism divides the target interval into multiple buckets. If a node shares a bucket with the user who queries a data record, the node will tag the data record based on the bucket of the record, forming a separate label. The user searches for the label and the corresponding encrypted data in the region, and decrypts the data to obtain the query results. Under the bucket mechanism, the sensors lack enough security, and the security level depends heavily on the length of the bucket interval [10].

The SafeQ protocol mainly relies on the prefix encoding algorithm. There are several things to do before the query: sorting the sensitive data in the region, splitting the region into intervals, encoding the intervals by certain rules, and conducting range search for minimum and maximum intervals.

Extended from the bucket mechanism, the E-N mechanism both protects data privacy and verifies the integrity of query results. However, the length of the bucket interval greatly affects the computing result of the mechanism, and increases power consumption and traffic loss [11-13].

The spatiotemporal crosscheck fuses the spatial and temporal cross validation algorithms. The core idea of this method is to set up verification relationships between nodes.

In general, the existing query protocols may consume too much power and face instable connections for real-time communication, if the WSN has a large environmental range.

2.2 Secure query protocol for wide-range multi-layer WSN

To overcome the defects of the existing query protocols, this paper puts forward a secure query protocol for wide-range multi-layer WSN, which mainly considers network security, attack mode and privacy protection.

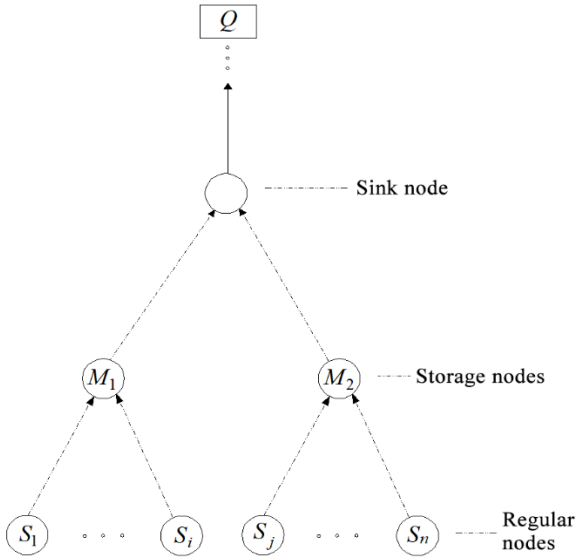

Figure 1. Structure of the multi-layer WSN

Figure 1 presents the structure of the multi-layer WSN, which consists of a sink node layer, a storage node layer and a regular node layer. As shown in the figure, there are numerous regular nodes scattering across the network environment. Each regular node has a limited computing power. Once inputted to the WSN, a query request will be forwarded by the sink node to the storage node layer. Then, each storage node will look for the result that matches the criteria in its own query unit.

In the WSN, the data transmission is mainly threatened by privacy attack and integrity attack. Some of these attacks come from outside the system, and some are collusions between the attacker and byzantine nodes [14-16]. The attacker tends to capture data records or control network nodes. Therefore, our protocol aims to disable the attacker to infer privacy data from the current perception data or deduce the upper and lower limits from the query interval, and detect the forged and tampered query results at any time.

Our protocol draws the merits from the Range Doppler (R-D) algorithm and Z-O encoding. Below is a brief introduction to the two methods.

The R-D algorithm basically compares the search result with the upper and lower limits, which are inferred from the radius of the query area, and transforms the comparison into the contrast between the search value and the median of the query interval.

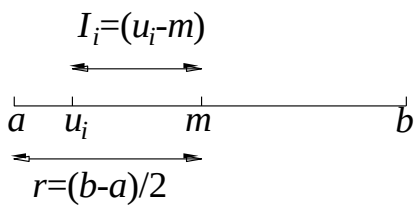

Figure 2. The principle of the R-D algorithm

The principle of the R-D algorithm is illustrated in Figure 2, where $[a, b]$ is the query range, u_i is the value to be queried, $m=(a+b)/2$ is the median of the query interval, r is half of the interval length, and $I_i=(u_i\text{-m})$ is the query index. The goal of the R-D algorithm is to judge if u_i fall into $[a, b]$. If $I_i<$r, then u_i belongs to this interval; otherwise, u_i does not belong to the interval. The R-D algorithm can obtain the desired result through a few comparisons between the query value and the limits, and is thus known for low computing complexity.

The Z-O encoding is an important way to query extreme values. By the R-D algorithm, the value comparison can be transformed into the problem of intersection between code sets. The Z-codes and O-codes can be respectively defined as:

$$Z_x = \left\{ x_n x_{n-1} \cdots x_{i+1} \middle| x_i = 0,1 \le i \le n \right\} \tag{1}$$

$$O_x = \left\{ x_n x_{n-1} \cdots x_i \middle| x_i = 1,1 \le i \le n \right\} \tag{2}$$

where, $x_i \in [0,1]$; n is the total number of Z-codes and O-codes. Both x and n are binary numbers. The value of x is greater than y, if the intersection between Z-codes and O-codes is non-empty.

In recent years, the attackers can deduce the original data from the digital codes and then steal the key information. To prevent this attack, the hash-based message authentication code (HMAC) can be introduced to achieve unidirectional protection of the original data:

$$\begin{cases} HMAC_k\left(N\left(e_i\right)\right) = HMAC_k\left(N\left(e_j\right)\right) & e_i = e_j \\ HMAC_k\left(N\left(e_i\right)\right) \ne HMAC_k\left(N\left(e_j\right)\right) & e_i \ne e_j \end{cases} \tag{3}$$

where, e_i is an element in the code. Thus, the value of x is greater than y, if the intersection between HMAC(N(O_x)) and HMAC(N(Z_y)) is non-empty.

Our protocol is implemented in three steps: system initialization, local query and result verification. During system initialization, the user inputs the query request to the system; then, the query parameters (K, k_i, m) to the sink node S_i, where K is the secret key to compute the HMAC value, and k_i is the secret key to construct the authentication code. The system initialization formula can be established as:

$$Q \to S_i : \left(K, k_i, m\right) \tag{4}$$

The local query mainly computes the m value in $I_{i,t}=(v_{i,t}\text{-m})$, using the R-D algorithm, and determines the code set $HMAC_k$(N(Z($I_{i,t}$)) in $v_{i,t}$ based on the K value. In this step, the storage node M collects the set of query codes:

$$S_i \to M : HMAC_k\left(N\left(Z\left(I_{i,t}\right)\right)\right) \tag{5}$$

both S_i and M satisfy the following constraints:

$$\begin{cases} HMAC_k\left(N\left(O\left(r\right)\right)\right) \to S_i \\ S_i \to M : \left(v_{i,t}\right) \quad v_{i,t} \notin \left[a,b\right] \end{cases} \tag{6}$$

where, O is the set of numerical codes.

During result verification, all the qualified perceptual data in the set and encrypted data are queried, and fed back to the

querier. The verification formulas are as follows:

$$M \to Q: HMAC_k\left(N\left(Z\left(I_{i,t}\right)\right)\middle\|\left(v_{i,t}\right)_k\right) \quad v_{i,t} \in [a,b] \quad (7)$$

$$M \to Q: C_M = HMAC_k\left(\|c_j\right) \quad v_{j,t} \notin [a,b] \quad (8)$$

As shown in Figure 3, the verified results are fused by S_i under the query conditions. The fused result is transmitted to the storage unit to construct the corresponding authentication code, which retains the integrity of the query result.

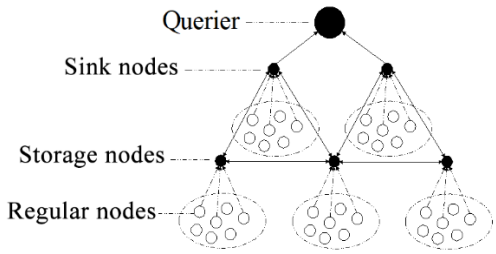

Figure 3. Fusion of the verified results

3. SIMULATION VERIFICATION

To verify its rationality and reliability, our protocol was firstly compared with three traditional query protocols (namely, bucket mechanism, SafeQ and top-k query) through simulations. The sensitive data of the four protocols were encrypted by Data Encryption Standard (DES) in C++, and proved to have the same HMAC power consumption.

Firstly, the four protocols were applied to submit the data for 100 random queries, and their power consumptions were measured within the same time interval. To reduce uncertainty, the measured data of the 100 queries were averaged as the final results.

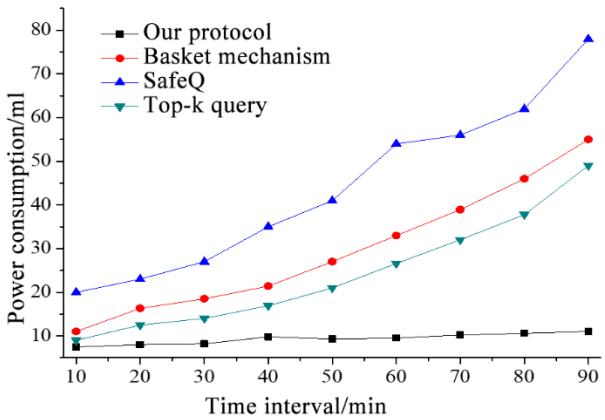

Figure 4. Mean power consumption of regular nodes

As shown in Figure 4, the total amount of data queried increased with the elapse of time, and the power consumptions of the four protocols increased to different degrees. The four protocols could be ranked as SafeQ, bucket mechanism, top-k query and our protocol in descending order of power consumption. On average, our protocol consumed only 1/3 of the power required by SafeQ. This is because SafeQ needs redundant data for its storage nodes to receive the encrypted files. The ensuing extra data chains expand the size of the feedback files, which is 2~3 times that of the original data. Bucket mechanism and top-k query also face similar problems. By contrast, our protocol only provides encrypted files meeting the query conditions, which only account for a small portion of the total data.

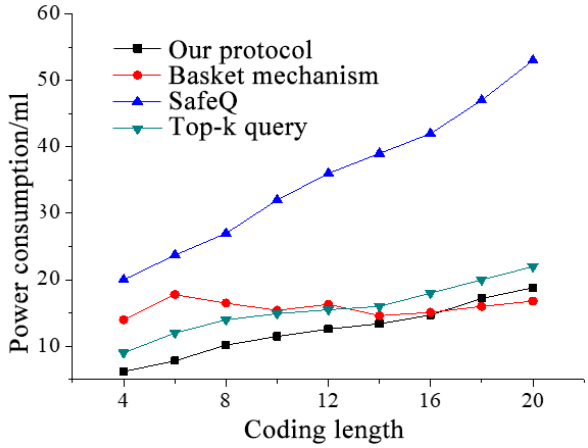

Figure 5. Mean power consumption at different coding lengths

Next, the power consumption of each protocol was measured at different coding lengths within the query time of 25min. The measured results (Figure 5) show that, with the increase in the coding length n, the power consumptions of all protocols were on the rise, for n is positively correlated with the number of HMAC codes. SafeQ still consumed the most power, about 3.6 times that of our protocol. The power-efficiency of our protocol comes from the following facts: The total number of Z-codes and O-codes is 2n, while our protocol reduces the total number of elements below that level by the HMAC authentical mechanism.

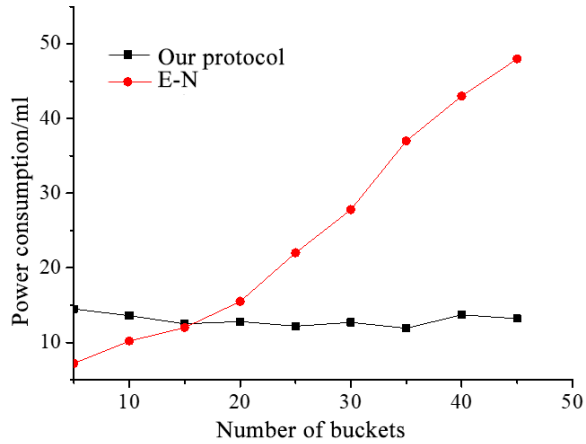

Figure 6. Mean power consumption at different number of buckets

To further verify the data integrity, our protocol and the E-N mechanism were simulated at different number of buckets. As shown in Figure 6, with the growth in the number of buckets, the power consumption of the E-N rocketed up, while that of our protocol did not change significantly. The superiority of our protocol is attributable to is immunity to the length of bucket interval. The query time of the E-N is seriously affected by the length of the bucket interval. If the

bucket interval is short, the number of bucket mechanisms increases, forcing the nodes to produce more authentication codes.

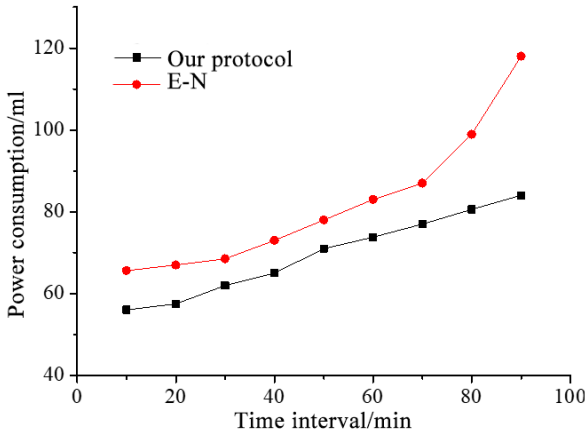

Figure 7. Mean power consumption at different time intervals

Fixing the number of buckets at 16, the author measured the power consumptions of our protocol and the E-N mechanism for integrity verification within different query times. It can be seen from Figure 7 that, with the elapse of query time, both protocols consumed more power to verify the data integrity. Overall, our protocol consumed 1.2 times fewer power than the E-N.

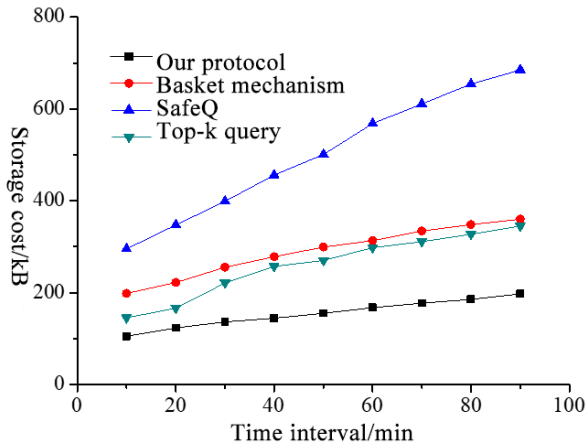

Figure 8. Mean storage cost of each protocol

Finally, the storage costs of our protocol, bucket mechanism, SafeQ and top-k query were compared (Figure 8). The comparison shows that, the longer the query time, the greater the storage cost for data submission in each protocol. Our protocol achieved the smallest mean storage cost, about 0.8 time that of bucket mechanism and 0.4 time that the SafeQ.

4. CONCLUSIONS

Drawing on existing secure query protocols for the WSN, this paper proposes a secure query protocol for wide-range multi-layer WSNs in the IoT. In this protocol, data transmission is protected from three aspects; network security, attack mode and privacy protection. The main conclusions are as follows:

(1) Our protocol only needs to compare the perceptual data nodes with the radius of the query interval once, which greatly improves the query efficiency. The embedded Z-O codes can transform the comparison into solving the intersection of the two sets.

(2) The HMAC authentication mechanism was introduced to improve the data security and privacy of our protocol.

(3) The simulation results show that our protocol outperforms the traditional protocols in query time, coding length and bucket mechanism, and enjoys high power efficiency, low storage cost and good query accuracy.

ACKNOWLEDGMENT

This work was supported by the Scientific and Technological Research Program of Chongqing Municipal Education Commission (Grant Numbers: KJ1602901, KJQN201803102) and Chongqing College of Electronic Engineering Scientific Research Project (Grant Number: XJZK201809).

REFERENCES

[1] Hua, D., Ye, Q., Geng, Y., Jia, X., He, R. (2016). CSRQ: Communication-efficient secure range queries in two-tiered sensor networks. Sensors, 16(2): 259-263. https://doi.org/10.3390/s16020259

[2] Bagaa, M., Challal, Y., Ouadjaout, A., Lasla, N., Badache, N. (2012). Efficient data aggregation with in-network integrity control for WSN. Journal of Parallel & Distributed Computing, 72(10): 1157-1170. http://dx.doi.org/10.1016/j.jpdc.2012.06.006

[3] Yang, D., Misra, S., Fang, X., Xue, G., Zhang, J. (2012). Two-tiered constrained relay node placement in wireless sensor networks: Computational complexity and efficient approximations. IEEE Transactions on Mobile Computing, 11(8): 1399-1411. https://doi.org/10.1109/TMC.2011.126

[4] Shan, D.S., Wan, Z.H., Qiao, L. (2011). Optimal sensors placement for long-span railway steel truss cable-stayed bridge. Journal of Civil Architectural & Environmental Engineering, 2: 795-798. https://doi.org/10.1109/ICMTMA.2011.482

[5] Stavrou, E., Pitsillides, A. (2010). A survey on secure multipath routing protocols in WSNs. Computer Networks, 54(13): 2215-2238. https://doi.org/10.1016/j.comnet.2010.02.015

[6] Othman, J.B., Mokdad, L. (2010). Enhancing data security in ad hoc networks based on multipath routing. Journal of Parallel & Distributed Computing, 70(3): 309-316. https://doi.org/10.1016/j.jpdc.2009.02.010

[7] Nasser, N., Chen, Y. (2007). SEEM: Secure and energy-efficient multipath routing protocol for wireless sensor networks. Computer Communications, 30(11-12): 2401-2412. https://doi.org/10.1016/j.comcom.2007.04.014

[8] Madria, S., Yin, J. (2009). SeRWA: A secure routing protocol against wormhole attacks in sensor networks. Ad Hoc Networks, 7(6), 1051-1063. https://doi.org/10.1016/j.adhoc.2008.09.005

[9] Boukerch, A., Xu, L., El-Khatib, K. (2007). Trust-based security for wireless ad hoc and sensor networks. Computer Communications, 30(11-12): 2413-2427.

https://doi.org/10.1016/j.comcom.2007.04.022
[10] Karlof, C., Wagner, D. (2003). Secure routing in wireless sensor networks: Attacks and countermeasures. Ad Hoc Networks, 1(2-3): 293-315. https://doi.org/10.1016/S1570-8705(03)00008-8
[11] Lopez, J., Roman, R., Agudo, I., Fernandez-Gago, C. (2010). Trust management systems for wireless sensor networks: Best practices. Computer Communications, 33(9): 1086-1093. https://doi.org/10.1016/j.comcom.2010.02.006
[12] Hsieh, M.Y., Huang, Y.M., Chao, H.C. (2007). Adaptive security design with malicious node detection in cluster-based sensor networks. Computer Communications, 30(11-12): 2385-2400. https://doi.org/10.1016/j.comcom.2007.04.008
[13] Safa, H., Artail, H., Tabet, D. (2010). A cluster-based trust-aware routing protocol for mobile ad hoc networks. Wireless Networks, 16(4): 969-984. https://doi.org/10.1007/s11276-009-0182-1
[14] Deng, J., Han, Y.S. (2007). Multipath key establishment for wireless sensor networks using just-enough redundancy transmission. IEEE Transactions on Dependable & Secure Computing, 5(3): 177-190. https://doi.org/10.1109/TDSC.2007.70233
[15] Khalil, I., Bagchi, S., Rotaru, C.N., Shroff, N.B. (2010). UnMask: Utilizing neighbor monitoring for attack mitigation in multihop wireless sensor networks. Ad Hoc Networks, 8(2): 148-164. http://dx.doi.org/10.1016/j.adhoc.2009.06.002
[16] Yang, Y., Zhong, C., Sun, Y., Yang, J. (2010). Network coding based reliable disjoint and braided multipath routing for sensor networks. Journal of Network & Computer Applications, 33(4): 422-432. https://doi.org/10.1016/j.jnca.2010.02.003