Modified SafeQ Protocol based Technique for Maintaining Privacy & Integrity in Wireless Sensor Network

Hardik R Patel
Department of Computer Science & Engineering
Nar Narayan Shastri Institute of Technology, Jetalpur, Gujarat, India

Abstract

The architecture of two-tiered sensor networks, where storage nodes serve as an intermediate tier between sensors and a sink for storing data and processing queries, has been widely adopted because of the benefits of power and storage saving for sensors as well as the efficiency of query processing. However, the importance of storage nodes also makes them attractive to attackers. We use SafeQ, a protocol that prevents attackers from gaining information from both sensor collected data and sink issued queries. SafeQ also allows a sink to detect compromised storage nodes when they misbehave. To preserve privacy, SafeQ uses a novel technique to encode both data and queries such that a storage node can correctly process encoded queries over encoded data without knowing their values. To preserve integrity, we propose two schemes—one using Hidden Markov Model to generate integrity verification information so that a sink can use this information to verify whether the result of a query contains exactly the data items that satisfy the query. To improve performance, we use an optimization technique using Bloom filters to reduce the communication cost between sensors and storage nodes.

Keywords: SafeQ, Wireless sensor network systems (WSNs), S&L scheme

I. INTRODUCTION

Wireless sensor network systems (WSNs) have been broadly sent For different applications, For example, environment detecting, building security observing, seismic tremor Forecast, and so Forth we consider a two-layered sensor system construction modeling in which Storage Nodes assemble information From adjacent sensor and giving answer inquiries From the user of the network system. The Storage Node work as a moderate level between the sink and the sensors For putting away information or handling inquiries.

Fig. 1: Architecture of wsn[1]

Storage Node convey three primary advantages to sensor systems.

1) Sensors spare power by sending all gathered information to their nearest StorageNode instead of sending them to the sink through long courses.
2) Sensors can be memory-restricted on the grounds that information is chiefly put away on StorageNode.
3) Inquiry handling turns out to be more effective in light of the fact that the sink just speaks with StorageNode For inquiries.

The inclusion of Storage Nodes in sensor networks system was initial defined in[2] and has been widely defined[3]-[4]. Several products of Storage Nodes are StarGate [5] and RISE [6], are available in market. In any case, The consideration Of Storage Node additionally takes huge privacy and integrity challenges. As Storage Node Store information got from sensors and giving as an imperative part for noting inquiries, they are more powerless against be Compromised, particularly in a threatening
situation. A Compromised StorageNode Forces huge dangers to a sensor system. To start with, the assailant may get touchy information that has been, or will be, put away in the Storage Node. Second, the Compromised Storage Node may return Forged information For a question. Third, this Storage Node may exclude all information things that fulfill the inquiry.

In this way, we need to outline a convention that keeps assailants From picking up information From both sensor gathered information and sink issued inquiries, which regularly can be demonstrated as range questions, and permits the sink to identify Compromised Storage Node when they act up.

For Privacy, Compromising a Storage Node should not permit the attacker to acquire the delicate information that has been, and will be, put away in the node, and additionally the queries that the StorageNode has gotten, and will get. Note that we treat the queries From the sink as secret on the grounds that such inquiries may release basic information about inquiry guarantors' intrigues, which need to be secured particularly in military applications.

For integrity, the sink needs to recognize whether a query result From a StorageNode incorporates Forged information things or does exclude all the information that fulfill the inquiry. There are two key challenges to solving the privacy and integrity. Initial, a StorageNode needs to accurately process encoded inquiries over encoded information without knowing their real values. Second, a sink needs to confirm that the result of an inquiry contains all the information things that fulfill the question and does not contain any Forged information.

II. RELATED WORK

Privacy- and integrity-preserving range queries in WSNs have drawn people’s attention recently[4],[7],[8]. Sheng and Li proposed a scheme to work with the privacy and integrity of range queries in sensor networks[4]. This scheme uses the bucket-partitioning idea proposed by Hacigumus et al. in [13] For informationbase privacy. The fundamental thought is to divide the domain of information values into bucket, the range of which is figured in light of the conveyance of information values and the location of sensors. In each time-slot, a sensor gathers information things From nature, places them into Buckets, encryts them together in each bucket, and afterward sends each encrypted bucket along with its bucket ID to a close-by StorageNode. For each Bucket that has no information things, the sensor sends an encoding number, which can be utilized by the sink to confirm that the Bucket is empty, to a close-by Storage Node. At the point when the sink needs to per Form a Range query, it finds the smallest arrangement of bucket IDs that contains the range in the query, and then sends the set as the query to Storage Node. After accepting the bucket IDs, the Storage Node gives back the comparing encrypted information in every one of those buckets. The sink can then decode the encoded Buckets and check the trustworthiness utilizing encoding numbers The S&L scheme only considered one-dimensional information in [4], and it can be extended to handle multidimensional information by dividing the domain of each dimension into multiple buckets.

The S&L scheme has two main drawbacks inherited From the bucket-partitioning technique. First, as figured out in [14] the bucket-partitioning technique admit Compromised StorageNodes to obtain a reasonable estimation on the actual value of Both information items and queries. In SafeQ such estimations are very difficult. Second, For multidimensional information, the power utilization of both sensors and StorageNodes, as well as the space utilization of StorageNodes, increments exponentially with the number of dimensions because of the exponential increment of the quantity of buckets. In SafeQ, power and space utilization increments straightly with the quantity of measurements time the quantity of information items. Shi et al. proposed a streamlined Form of S&L’s integrity preserving plan planning to aiming the correspondence cost in the middle of sensors and StorageNode [7],[8]. The essential thought of their enhancement is that each sensor utilizes a bit guide to speak to which bucket have information and broadcast its bit guide to the close-by sensors. Each sensor joins the bit maps got From others to its own information things and encodes them together. The sink checks inquiry result trustworthiness For a sensor by inspecting the bit maps From its close-by sensors. In our examinations, we didn't pick the arrangements in [7],[8]For side-by-side comparision For two reasons. To start with, the methods utilized as a part of [7],[8] are like S&L plan aside From the enhancement For integrity verification. The way they stretch out S&L plan to handle multi-dimensional information is to separation the space of every measurement into numerous buckets. They acquire the same shortcoming of permitting Compromised StorageNode to gauge the estimations of information things and inquiries with S&L plan. Second, their advancement procedure permits a Compromised sensor to effFortlessly Compromised the trustworthiness confirmation usefulness of the system by sending misrepresented bit maps to sensors and StorageNode. Conversely, in S&L and our plans, a Compromised sensor can't endanger the queries and confirmation of information gathered by different sensors. A Compromised StorageNode powers huge dangers to a sensor system. Along these lines, to stay away From such sorts of attacks, a protocol is planned that keeps attackers From picking up information From both sensor gathered information and sink issued queries. We quantified the efficiency of SafeQ and S&L scheme on 1, 2, and 3 dimensional information. For better comparison, we directed our tests on the same information set that S&L utilized as a part of their experiment [4] . The information set was browsed an expansive genuine information set From Intel Lab [10] and it consist of the temperature, humidity and voltage information gathered by 44 nodes during March 1-10-2004. Note that S&L just directed analyses on the temperature information, while we experimented with both SafeQ and S&L schemes on 1-dimensional information (of temperature), 2-dimensional information (of temperature and humidity) and 3-dimensional information (of temperature, humidity, and voltage). As in [4], we just as partitioned 44 hubs into 4 gatherings and sent a StorageNode For every gathering. In actualizing SafeQ, we utilized HMAC-MD5 [11] with 128-bit keys as the hash capacity For hashing prefix numbers. We utilized the DES encryption calculation as a part of executing both SafeQ and S&L plan. In

All rights reserved by www.ijste.org 152
executing our Bloom filter technique, we picked the quantity of hash capacities to be 4 (i.e., k = 4), which ensures that the false positive rate actuated by the Bloom filter is under 1%. In actualizing S&L plan, we utilized the parameter values (i.e., VARp = 0.4 and ENp = 1) For processing optimal bucket partitions as in [4], and we utilized HMAC-MD5 with 128-bit keys as the hash function For figuring encoding number. For multi-dimensional information, we utilized their optimal bucket partition algorithm to segment multi-dimensional information along each dimension. In our trials, we tried different things with diverse sizes of time openings going From 10 minutes to 80 minutes. For every time space, we created 1,000 random range queries.

We make two key commitments in this paper. In the first place, we propose SafeQ, a novel and efficient protocol For taking care of range inquiries in two-layered sensor networks in a privacy and integrity saving design. SafeQ utilizes the procedures of three magic function and merkle hash tree and neighborhood chaining. Regarding security, SafeQ fundamentally reinpowers the security of two-layered sensor network systems. Dissimilar to Former workmanship, SafeQ prevents Compromised StorageNode From acquiring a sensible estimation on the genuine estimations of sensor gathered information things and sink issued inquiries. Regarding efficiency, our outcomes demonstrate that SafeQ altogether beats Former workmanship For multidimensional information as far as both power utilization and storage space.

III. EXISTING SYSTEM

the structural engineering of two-layered sensor systems, where serve as a transitional level in the middle of sensors and a sink For putting away information and preparing inquiries, has been generally embraced in view of the advantages of power and storage putting something aside For sensors and also the efficiency of query processing. On the other hand, the significance of StorageNode additionally makes them attractive to attackers. In this paper, the creators propose SafeQ, a convention that keeps attackers From picking up information From both sensor gathered information and sink issued inquiries. SafeQ additionally permits a sink to recognize Compromised StorageNode when they get out of hand. To work with privacy, SafeQ utilizes a novel method to encode both information and questions such that a StorageNode can effectively process encoded inquiries over encoded information without knowing their values. To work with integrity, the creators propose another information structure called a merkle hash tree and another is neighborhood chain that permits a sink to confirm whether the result of a query contains precisely the information things that fulfill the inquiry. Moreover, the creators proposed a solution For adjust SafeQ For event driven sensor systems.

A. Limitations of Existing System

The incorporation of StorageNode in wireless sensor network systems brings critical security challenges. As StorageNode store information got From sensors and serve as a critical part For noting inquiries, they are more defenseless against be Compromised, particularly in an antagonistic domain.

1) A Compromised StorageNode Forces noteworthy dangers to a sensor system. To begin with, the attacker may acquire delicate information that has been, or will be, put away in the StorageNode.
2) The Compromised StorageNode may return Forged information For a query.
3) This Storage Node may exclude all information things that fulfill the queries.

IV. PROPOSED WORK

A. Overview of Proposed Work

We define SafeQ, a protocol that keeps attackers From picking up information From both sensor gathered information and sink issued queries. SafeQ likewise permits a sink to identify Compromised StorageNodes when they act up. To work with privacy, SafeQ utilizes a novel procedure to encrypt the information and queries such that StorageNode can accurately process encrypted queries over encrypted information without knowing the values.

B. SafeQ Protocol Has Following Modules:

- StorageNode
- Sink
- Range Queries
- Integrity
- Privacy
C. StorageNode

StorageNodes are effective wireless devices that are outfitted with significantly more capacity limit and processing power than sensors. The StorageNode gathers all information from the Sensor Nodes. The StorageNode cannot see the genuine estimation of Sensor Node information. In the event that the StorageNode attempting to view the Sensor Node information, sink distinguishes act up of StorageNode.

D. Sink

The sink is the purpose of contact for clients of the sensor system. Each time the sink gets an query from a client, it first makes an interpretation of the query into different queries and afterward scatters the queries to the comparing StorageNodes, which handle the queries in light of their information and return the query results to the sink. The sink brings together the query results from different StorageNodes into the last answer and sends it back to the client. Sink can distinguish Compromised StorageNodes when they get into mischief.

E. Range Queries

The queries from the sink are range queries. A reach queries "discovering all the information items gathered at time-space in the extent" is denoted as. Note that the queries in most of the sensor network system applications can be easily modeled as range queries.

F. Complexity Analysis:

Accept that a sensor gathers -dimensional data things in a period opening, every trait of a data thing is a -bit number, and the HMAC aftereffect of each numericalized prefix is a number. The calculation cost, correspondence expense, and storage room of SafeQ are portrayed that the correspondence expense means the quantity of bytes sent for every accommodation or question, and the storage room indicates the quantity of bytes put away in a StorageNode for every accommodation. Besides, take note of that whether Sensor Nodes report to StorageNodes intermittently or upon a few occasions has no effect on these expenses of one time sending of data things.

G. Privacy Analysis:

In a SafeQ ensured two-layered sensor system, bargaining a StorageNode does not permit the assailant to get the real estimations of sensor gathered data and sink issued questions. The accuracy of this case is in light of the way that the hash capacities and encryption calculations utilized as a part of SafeQ are secure.

In the accommodation convention, a StorageNode just gets encoded data things and the protected hash estimations of prefixes changed over From the data things. Without knowing the keys utilized as a part of the encryption and secure hashing, it is computationally infeasible to register the real estimations of sensor gathered data and the relating prefixes.

In the question convention, a StorageNode just gets the protected hash estimations of prefixes changed over From a reach inquiry. Without knowing the key utilized as a part of the safe hashing, it is computationally infeasible to process the genuine estimations of sink issued questions.

Note that if a StorageNode and a sensor are both Compromised, the StorageNode may uncover the sensor gathered data and sink issued inquiries by utilizing animal power assaults. For this situation, the StorageNode knows the mutual mystery key for the capacity. Because of the one-wayness property of, the StorageNode can't uncover specifically.

Be that as it may, it can figure the consequences of the numericalized prefixes For every conceivable esteem in the data space in a beast power way, and afterward contrast the outcomes and they got data and questions. In light of the correlation, the StorageNode can uncover the sensor gathered data and sink issued questions. On the other hand, by and by, this computational expense could be restrictive for a vast data space.

H. Integrity Analysis

In Existing System To Work with Integrity Safeq Uses The Merkle Hash Tree And Neighbourhood Chains Technique. Here, In Our Propose System We Are Replase Merklehash Tree With Hidden Markov Model.

I. Hidden Markov Model

1) Set of states: \( \{S_1, S_2, \ldots, S_N\} \)
2) Process moves from one state to another generating a sequence of states $S_{i_1}, S_{i_2}, \ldots, S_{i_k}, \ldots$

3) Markov chain property: likelihood of each ensuing state depends just on what was the past state:
   \[ P(S_{i_k} | S_{i_1}, S_{i_2}, \ldots, S_{i_{k-1}}) = P(S_{i_k} | S_{i_{k-1}}) \]

4) States are not obvious, but rather every state haphazardly produces one of $M$ perceptions (or unmistakable states) $\{V_1, V_2, \ldots, V_m\}$

5) To characterize shrouded Markov model, the accompanying probabilities must be indicated: lattice of move probabilities $A=(a_{ij})$, $a_{ij}= P(s_i | s_j)$, grid of perception probabilities $B=(b_{i}(v_m))$, $b_{i}(v_m) = P(v_m | s_i)$ and a vector of introductory probabilities $p=(p_i)$, $p_i = P(s_i)$. Model is spoken to by $M=(A, B, p)$.

6) Evaluation issue. Given the HMM $M=(A, B, p)$ and the perception succession $O=o_1, o_2, \ldots$ Alright compute the likelihood that model $M$ has produced arrangement $O$. Decoding problem. Given the HMM $M=(A, B, \pi)$ and the observation sequence $O=o_1, o_2, \ldots$ oK calculate the most likely sequence of hidden states $S_i$ that produced this observation sequence $O$.

7) Decoding issue. Given the HMM $M=(A, B, p)$ and the perception succession $O=o_1, o_2, \ldots$ Alright compute the in all probability arrangement of concealed states $S_i$ that delivered this perception grouping $O$.

8) Learning issue. Given some preparation perception successions $O=o_1, o_2, \ldots$ Alright and general structure of HMM (quantities of concealed and unmistakable states), focus HMM parameters $M=(A, B, p)$ that best fit preparing data.

J. Work Flow of Proposed Work

1) SafeQ Working Steps Using HMM

Fig. 12: Working of SafeQ Using HMM

2) Steps

1) Sensor Collect information $(d_1, d_2, \ldots, d_n)$ and apply the secret key $((d_1)_{k_1}, \ldots, (d_n)_{k_n})$.

2) Sensor also apply the hashing function on the information $H(d_1, \ldots, d_n)$.

3) StorageNode perform Bucket Partation method on sensor’s collected information.

4) Now Sink issue the Query $\{t, \{a, b\}\}$.

5) Sink apply another magic function $\{t, G(\{a, b\})\}$ and sending query to the StorageNode.

6) Now hidden markov model apply to the sink issues queries.

7) HMM generate the final state result of query.

Assume we are given a Hidden Markov Model (HMM) with state space $S$, starting probabilities $\pi_{i0}$ being in state $i$ and move probabilities $a_{ij}$ of transitioning from state $i$ to state $j$. Let's assume we watch Output $y_1, \ldots, y_T$. The doubtlessly state sequences $X_1, \ldots, X_T$ that delivers the perceptions is given by the repeat relations

\[ V_{1,k} = P(y_1 | k) \cdot \pi_k \]
\[ V_{t,k} = \max_{x \in S} \left( P(y_t | k) \cdot a_{x,k} \cdot V_{t-1,x} \right) \]

Here $V_{t,k}$ is the likelihood of the most plausible state arrangement in charge of the first $t$ perceptions that has $k$ as its last state. The Viterbi way can be recovered by sparing back pointers that recollect which state $X$ was utilized as a part
of the second mathematical statement. Let $\text{Ptr}(k, t)$ be the capacity that profits the estimation of $x$ used to process $V_t, k$ if $t > 1$, or $k$ if $t = 1$. At that point:

$$x_I = \arg\max_{x \in S}(V_t, x)$$

$$x_{t-1} = \text{Ptr}(x_t, t)$$

8) Final state sending to the sink.
9) Now Sink check the final state = initial state.
10) If yes then decrypt the initial state using the secret key.

K. Experimental Setup

The implementation results from our system show that SafeQ significantly performs the S&L scheme for multidimensional information in terms of energy and space utilization. For the integrity-maintaining schemes, the Hidden Markov Model is better than neighborhood chain and Merkle hash tree method in terms of energy and space consumption. The rationale for us to include the Hidden Markov Model perform both Forward and backward chaining are the typical approach to achieving integrity. We perform SafeQ Hidden Markov Model. Now we are seen the implementation there are 11 sensors are collect the information from environment and transmit those information to the StorageNode. StorageNode also collect the query from sink which is declared by base station (BC). Now Sensor Node 1, 2, 3, 4, 5, 6, 7, 9, 10, 11 sending collected information to the StorageNode1.

Now base station send the query 52(115), 215(207) to StorageNode. StorageNode find the query result and QR is available from sensor node 6 information. Then StorageNode send the information to sink.
L. Packet Delivery Rate

The experimental results from our system comparisons show that SafeQ PerForm For multidimensional information in terms of energy consumption, space consumption, and packet delivery ratio in existing system and our system. Experimental results are show that Packet delivery ratio in existing system is 1 902 883 124 115 116 107 128 119 11 seconds and in our system is 1 1002 1003 04 05 06 07 08 09 0 seconds. So the average packet delivery ratio in existing system is 227.82 seconds and in our system is 185.91 seconds.
M. Space Occupancy

Now we perform the space occupancy show in graph. The space occupancy in existing system is 1 372 533 914 1225 1416 1547 1568 1499 152 bytes and in our system is 1 212 413 714 1025 1246 1367 1368 1369 136 bytes. So the average space occupancy in existing system is 838.82 bytes and in our system is 713.73 bytes.

N. Energy Consumption
Now we consider the energy consumption. The energy consumption in existing system is 1122 343 334 285 347 308 369 32 joule and in our system is 1122 183 174 125 106 187 148 209 16 joule. So the average energy consumption in existing system is 218.82 joule and in our system is 115.55 joule. This results conform to theoretical analysis that the packet delivery ratio, space occupancy, energy consumption in our system are better grows than existing system.

V. CONCLUSION

We define three key commitments in the paper. In the First place, We create SafeQ a Novel and effective convention for taking care of reach queries in two-layered sensor organizes in a privacy and integrity preserving design. SafeQ utilizes the methods of prefix membership verification, and Hidden Markov model. As far as privacy, SafeQ essentially fortifies the security of two-layered sensor systems. Dissimilar to Former workmanship, SafeQ prevents a Compromised StorageNode From acquiring a sensible estimation on the real estimations of sensor gathered information things and sink issued queries. Regarding efficiency, our outcomes demonstrate that SafeQ altogether outflanks Former workmanship for multidimensional information as far as both power utilization and storagespace. Second, we use an improvement procedure utilizing Bloom channels to fundamentally diminish the communication cost in the middle of sensors and StorageNode.

REFERENCES

[10] Intel Lab Information. http://berkrley.intel-research.net/labinformation,