International Journal of Scientific Engineering and Applied Science (IJSEAS) – Volume-2, Issue-1, January 2016
ISSN: 2395-3470
www.ijseas.com

REDUCE FALSIFIED SUB AGGREGATE ATTACK USING ATTACK RESILIENT TECNIQUE IN WSN

[1]G.Sangar [2]N.Vadivelan [3]K.Prasanth

1. Assistant professor, Department of CSE, Sakthimariamman Engineering College, Chennai.
2. Assistant professor, Department of CSE, Sakthimariamman Engineering College, Chennai.
3. Assistant professor, Department of CSE, Sakthimariamman Engineering College, Chennai

ABSTRACT

A wireless sensor network (WSN) is used to monitor physical or environmental conditions and transfer data through the network to destination.WSN are mostly used in many applications, such as volcano and fire monitoring, urban sensing etc. In this loss-resilient aggregation framework called synopsis diffusion concept is used to accurately compute aggregates such as predicate count, sum etc., But this framework does not address the problem of false sub aggregate values contributed by compromised nodes. Due to this negative impact, large errors may occur in the aggregate computed at the base station. In the proposed system, attack-resilient computation algorithm is used to compute the true aggregate by filtering out the contributions of compromised nodes. In addition to this, RPS algorithm is used to receive entire data at the receiver end without missing of any packets. The existing system is used to compute aggregates whereas proposed system securely computes aggregates, despite the falsified attack. The First phase is to derive initial estimate of the aggregate based on minimal authentication information received by base station and, the second phase derives more authenticated information from subset of nodes determined by the estimate of the first phase. Finally it reduces the amount of communication overhead and energy consumption.

Keywords: synopsis diffusion, attack-resilient, RPS algorithm Data aggregation, hierarchical aggregation, in-network aggregation, sensor network security.

INTRODUCTION

Wireless Sensor Network are increasingly used in several real-world applications, such as wild habitat monitoring, volcano and fire monitoring, urban sensing, and military surveillance. In most cases, the sensor nodes form a multi-hop network while the base station (BS) acts as the central point of control [1]. Typically, a sensor node has limitation in terms of computation capability and energy reserves. The BS wants to collect the sensed information from the network. One common way is to allow each sensor node to forward its reading to the BS, possibly via other intermediate nodes. Finally, the BS processes the received data.

In a large WSN, In-network data aggregation is used for combining partial results at intermediate nodes during message routing. So that it reduces the amount of communication overhead and energy consumption. One approach is to construct a spanning tree rooted at the BS, and then perform in-network aggregation along the tree [2]. It generalizes these aggregates to predicate Count and Sum. In addition, Average can be computed from Count and Sum. Further Sum algorithm is easily extended to compute Standard Deviation and Statistical Moment of any order.

However, communication losses resulting from node and transmission failures, can adversely affect tree-based aggregation approaches. To address this problem, multi-path routing techniques are used for forwarding the sub aggregates. For duplicate insensitive aggregates such as Min and Max, this approach provides a fault-tolerant solution. Unfortunately, for duplicate sensitive aggregates, such as Count and Sum, multi-path routing leads to double counting of sensor readings. Temporal over-sampling can simplify the analysis of sensitive aggregates at the base station [3].

Another major issue in WSN is progress in battery technology that has been much slower than increases in processing and communication rates, which emphasizes the importance of energy-efficient operation [4]. To solve this problem resilient algorithm is used to check where there is plenty of redundancy in the data for consistency.

Hence, this algorithm is also designed for securely computing the aggregates even in the presence of the attacks [5]. In addition to this algorithm, Representative Point Selection algorithm is used to

receive entire data at the receiver end without missing of any packets.

RELATED WORKS

The possibility of node compromise introduces more challenges because most of the existing in-network aggregation algorithms have no provisions for security. A compromised node might attempt to thwart the aggregation process by launching several attacks, such as eavesdropping, jamming, message dropping, message fabrication, and so on.

PROBLEM STATEMENT

The verification algorithm failed to compute the aggregate in the presence of an attack. It does not address the problem of false sub aggregate values contributed by compromised nodes. This negative impact causes large errors when aggregate computed at the base station.

The possibility of node compromise introduces more challenges because most of the in-network aggregation algorithms have no provisions for security. Duplicate sensitive aggregates, such as Count and Sum, multi-path routing leads to double counting of sensor readings.

In the proposed method, an algorithm is designed to securely compute aggregates, such as Count and Sum despite the falsified sub aggregate attack. In addition to attack-resilient algorithm, RPS algorithm is used to receive entire data at the receiver end without missing of any packets.

Representative Point Selection protects the authenticity of the temporal variation patterns in the aggregation results. It detects false temporal variation patterns. It also exploits the spatial correlations among the sensor readings in close proximity, a series of security mechanisms are also proposed to protect the sampling procedure. It reduces the communication cost through selective verifications of aggregation results.

The First phase is to derive initial estimate of the aggregate based on minimal authentication information received by base station and, the second phase derives more authenticated information from subset of nodes determined by the estimate of the first phase. Additionally, partial results have to be combined at intermediate nodes during message routing.

The key observation which exploits is that minimize the communication overhead to verify the correctness of the final synopsis the BS does not need to receive authentication messages from all of the nodes.

SCOPE OF STUDY

The scope of this proposed system, a compromised node C can falsify its own sensor reading with the goal of influencing the aggregate value. There are three cases.

Case (i): If the local value of an honest node can be any value, then a compromised node can pretend to sense any value. In this case, there is no way to detect the falsified local value attack.

Case (ii): If the local value of an honest node is bounded, and a compromised node falsifies the local value within the bound, there is no solution for detecting previous attack.

Case (iii): The local value of an honest node is bounded, and a compromised node falsifies the local value outside the bound. The proposed algorithm does detect and guard against an attack scenario.

A compromised node can falsify the sub-aggregate which is supposed to compute based on the messages received from its child nodes. It is challenging to guard against this attack, and addressing this challenge is the main focus of the proposed system.

SYSTEM IMPLEMENTATION

User Authentication -The User Authentication module mainly consists of sub-modules such as user registration and attach file. In this module user authentication, registration process, login process and attaching file are performed. With the help of user authentication, the user could be a member of server and his details will be stored in server's database. By this authenticated user can logged in and transfer the file securely.

User Registration - In this sub-module, Authorized id number is used as primary key. If a client wants to transfer data, the user must be a member of the server by the process of registration. The data provided by the user will be stored in the database. After registration, the user will be provided with a username and password for the logging in.

Attach File - In this sub-module, the user after successful login provided with two options. They are user details and transferring file. In user details user can view the login details along with time. And in transferring file user tries to browse the file from the user system and attach it for transferring.

File Transferring- In this module, after successful login the user tries to browse the file from the user system and attach it and then transfer the file from source to destination securely. Before transferring the file, user has to provide the destination IP address and the start the server to connect both sender and receiver nodes.

Start Server - In this sub-module, providing IP address of destination node is very important to transfer the file to the exact destination. After proving IP address user has to start the server which connects the sender and receiver. This connection last until entire data is transferred completely.

Transfer File - In this sub-module, completing the validation and verification of user it allows user to browse and attach the file. Then connect both source and destination by starting the server, user can transfer the file through wireless sensor network.

Predicting Attackers- In this module, hacker tries to hack the router; the router will be sending that information immediately to the client. If there is also more than one compromised node, the shortest path selection algorithm finds alternative paths and transfer the data securely to the destination.

Hacking Router - In this sub-module, the hacker tries to hake the routers, which are act as an intermediate between the Server and Client. When hacker tries to hack the router, the router will send that information immediately to the client when router working as DOS Node. The hacked node may tries to add false data in to the original data while transferring the data. Otherwise router forwards the packet to the next node.

Random Path Selection - In this sub-module, if one of the routers is hacked and while transferring the data it omits the hacked router and chooses a shortest path from routing table passes data through those routers and reaches the destination.

Data Storage - In this module, receiver has two options while transferring file, one is checking receiver details and the other one is selecting path to store the data in the data base. In order to receive the entire data correctly and filter out original data from false data RPS algorithm is used at the receiver end.

Receiver Details - In this sub-module, client can able to view the particular details of the receiver which is going to store the file in the database. IP address of the destination node helps the data to be transferred quickly.

Select Receiver Path - In this sub-module, at the receiver end receiver has to select the path where the data is being stored in the database. RPS algorithm detects false temporal pattern then filters original data and stores it into the specified location given by the receiver.

3.6 PROPOSED SYSTEM

In the proposed system, an algorithm to securely compute aggregates such as Count and Sum despite the falsified sub aggregate attack. In addition to attack-resilient algorithm, RPS algorithm is used to receive entire data at the receiver end without missing of any packets. The existing system is used to compute aggregates whereas proposed system securely computes aggregates, despite the falsified attack.

The First phase is to derive initial estimate of the aggregate based on minimal authentication information received by base station and, the second phase derives more authenticated information from subset of nodes determined by the estimate of the first phase.

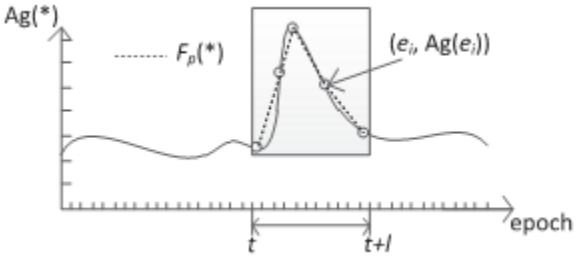
Representative Point Selection protects the authenticity of the temporal variation patterns in the aggregation results. It detects false temporal variation patterns. These representative application scenarios open up different schemes for integrating the WSNs into the Internet. Its exploits the spatial correlations among the sensor readings in close proximity, a series of security mechanisms are also proposed to protect the sampling procedure. It reduces the communication cost through selective verifications of aggregation results.

This algorithm takes space and time, considering the time of the computation. The BS selects some points from the series of aggregation results in the time window to be verified, and checks their correctness to detect any fabrication of temporal variation patterns.

The key observation which exploits is that minimize the communication overhead to verify the correctness of the final synopsis the BS does not need

to receive authentication messages from all of the nodes. Fig illustrates the Data Flow Diagram includes database where the received data is stored.

RPS Algorithm:



This algorithm takes space and time, considering the time of the computation.

The BS selects some points from the series of aggregation results in the time window to be verified, and checks their correctness to detect any fabrication of temporal variation patterns. Considering that the adversary can manipulate only a small number of aggregation results such as extreme points to tamper with the temporal variation pattern, it may be ineffective to check a set of randomly selected points to detect forged patterns because the selected points may not cover these manipulated points, which causes that the attack is not detected. Thus, to guarantee effective attack detection, the selected points should be able to capture the temporal variation pattern in the timewindow like extreme points. We refer to these points as representative points and the epoch of a representative point as representative epoch. After the selection of representative points, the BS broadcasts a verification request, which includes the representative epochs, the sampling ratio ϱ, and a nonce number $nonce_v$, to the WSN. Once receiving the verification request, each node decides whether to act as a sampled node. Before the sampled nodes send to the BS their sensor readings of every representative epoch, their neighboring nodes verify the correctness of sample data and authenticate the sample messages. the sensor reading samples, the BS checks the correctness of the aggregation results of each representative epoch by hypothesis testing. The general form of the hypothesis tests is

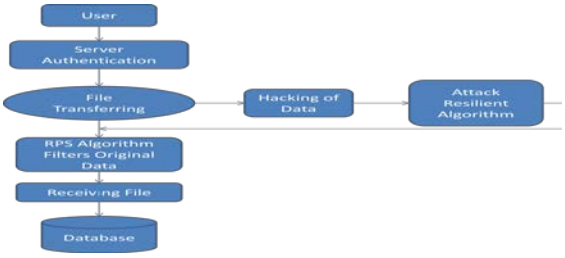
$$H_0 : A(t) = Ag(t) \text{ versus } H_a : A(t) \neq Ag(t).$$

If the aggregation results in all representative epochs are verified as correct, the temporal variation pattern in the time window is assumed to be authentic. we suppose that the time window to be verified is from epoch t to epoch $t+l$, denoted by $[t, t+l]$, and we always take the points at the boundary epochs t and $t+l$, as two representative points. Obviously, $l+1 < l_{max}$.

Representative Point Selection:

Representative Point Selection protect the authenticity of the temporal variation patterns in the aggregation results. it detect false temporal variation patterns. These representative application scenarios open up different schemes for integrating the WSNs into the Internet, which we present and compare in Section. Its exploits the spatial correlations among the sensor readings in close proximity, a series of security mechanisms are also proposed to protect the sampling procedure. It reduces the communication cost through selective verifications of aggregation results.

Input : A set of points S
Output : k clusters
For every cluster u (each input point), in u.mean and u.rep store the mean of the points in the cluster and a set of c representative points of the cluster (initially c = 1 since each cluster has one data point). Also u.closest stores the cluster closest to u.
All the input points are inserted into a k-d tree T
Treat each input point as separate cluster, compute u.closest for each u and then insert each cluster into the heap Q. (clusters are arranged in increasing order of distances between u and u.closest).
While size(Q) > k
Remove the top element of Q(say u) and merge it with its closest cluster u.closest(say v) and compute the new representative points for the merged cluster w.
Also remove u and v from T and Q.
Also for all the clusters x in Q, update x.closest and relocate x
insert w into Q
repeat



Data Flow Diagram

Advantages - A better secure communication and Confidentiality,
Resilience against node capture
Replication attacks using reduced resources
Minimum overhead and energy consumption
Aggregate are computed at the intermediate nodes

LITERATURE REVIEW

In order to achieve the assurances to enable the BS to obtain the true estimate of the aggregate even in the presence of an attack, there are various efficient methods that were designed for secure transmission of data were investigated.

The wireless sensor network (WSN) which is used to monitor physical or environmental conditions and transfer data through the network to destination [1]. In this loss-resilient aggregation framework called synopsis diffusion concept is used to accurately compute aggregates such as predicate count, sum etc., But this framework does not address the problem of false sub aggregate values contributed by compromised nodes. A compromised node C can falsify its own sensor reading with the goal of influencing the aggregate value. Due to this negative impact, large errors may occur in the aggregate computed at the base station. The attack-resilient computation algorithm consists of two phases. In the first phase, the BS derives a preliminary estimate of the aggregate based on minimal authentication information received from the nodes. In the second phase, the BS demands more authentication information from only a subset of nodes while this subset is determined by the estimate of the first phase. At the end of the second phase, the BS can filter out the false contributions of the compromised nodes from the aggregate.

To reduce energy consumption, during collection of data from individual nodes is aggregated at a base station or host computer [2], many systems also perform in network aggregation of sensor data at intermediate nodes enroute to the base station. The compromised nodes can be used to inject false data that leads to incorrect aggregates being computed at the base station. An approach used by several data acquisition systems for sensor networks is to construct a spanning tree rooted at the data sink, and then perform in-network aggregation along the tree. Tree-based aggregation approaches are not robust to communication losses which result from node and transmission failures and are relatively common in sensor networks. Because each communication failure loses an entire sub tree of readings, a large fraction of sensor readings are potentially unaccounted for at the data sink, leading to a significant error in the aggregate computed. To address this problem, researchers have proposed novel algorithms that work in conjunction with multiple path routing for computing aggregates in lossy networks. A compromised node can be used to launch a variety of security attacks. These attacks include jamming at the physical or link layer as well as other resource consumption attacks at higher layers of the network software. Compromised nodes can also be used to disrupt routing protocols and topology maintenance protocols that are critical to the operation of the network. To design the attack-resilient COUNT protocol, two issues need to be addressed. First, since a node cannot locally determine the position of bit r in the final synopsis, the base station needs to specify a global criterion which determines if a node needs to include a MAC along with its synopsis. Second, this criterion should be designed so that the number of such nodes who include a MAC is minimized.

The imaging sensors embedded in the natural environment enable remote collection of large quantities of data, thus easing the design and deployment of sensing systems in a variety of application domains [3]. The data collected from such imagers is difficult to interpret due to a variety of “nuisance factors” in the data formation process, such as illumination, vantage point, partial occlusions, etc. There are three applications that exemplify these problems and the solutions are developed. First, it

explained how temporal over-sampling can simplify the analysis of a slow process such as the avian nesting cycle. Then, described how to overcome temporal under-sampling in order to detect birds at a feeder station. Finally, explains how to exploit temporal consistency to reliably detect pollinators as they visit flowers in the field. Embedded sensing system technologies are readily applicable to the visual monitoring of the natural environment. The ability to measure nesting patterns accurately and in a scalable manner is broadly relevant to ecosystem, including responses to climate change and land use. Currently, avian biologists manually inspect nesting locations and visually log the stage of the nest for future analysis. But some indicator variables include the number of eggs that are laid and eventually hatch and the occupancy of the nest over the different nesting stages. By visiting the sites to collect this information, biologists can incorporate domain knowledge to filter unwanted data and detect important events. However, not only is this time consuming but it limits the number of observations a biologist can collect. Moreover, observations are typically limited to the behavior of birds outside their nests because biologists "lacked the capability to peer inside the private lives of birds" for long durations for fear of disturbing the birds. Cameras are an ideal sensor for gathering these observations.

The application of wireless sensor network (WSN) technology to long-duration and large-scale environmental monitoring [4], sensor networks face a number of challenges. The field arguably emerged due to the commoditization of cheap, low-power, single-chip microcontrollers and radios. These components emerged due to the rapid growth of global industries such as cell phones, wireless remotes, and car locks. The Holy Grail is a system that can be deployed and operated by domain specialists not engineers, but this remains some distance into the future. Reliability and productivity are key concerns and influence the design choices for system hardware and software. Energy has been and remains a challenge for sensor network deployments. The energy state of a node places a constraint on the performance that a node can deliver. A node's energy state reflects its stored battery energy, actual and predicted harvested energy, and its energy load. Progress in battery technology has been much slower than increases in processing and communication rates, which emphasizes the importance of energy-efficient operation.

In security for data aggregation in sensor networks, scientific data collection, environmental monitoring, building health monitoring, burglar and fire alarm systems, and many other applications involving distributed interaction are used with the physical environment [5]. Current aggregation schemes were designed without security in mind and there are easy attacks against them. An aggregation transaction begins by broadcasting the query down the tree from the base station to the leaves. Then, the sensor nodes measure their environment and send their measurement back up the tree to the base station. Finally, the base station performs an aggregation computation to obtain the aggregate. Thus, sensor nodes act as data sources, and the base station acts as a sink. Data aggregation can be viewed as an important building block in sensor networks. One naive way to make an aggregation function more robust against spoofed sensor readings is to place upper and lower bounds on the acceptable range of a sensor reading. Resilient aggregation is best-suited to settings where there is plenty of redundancy in the data, so that cross-check of the sensor readings for consistency can be done. There are several approaches for making these aggregation schemes more resilient against certain attacks, and formed a mathematical framework for formally evaluating their security. Compromise of sensor nodes is indeed a real threat in real sensor networks. Because sensor nodes must be low-cost, cannot afford to mount them in physical packaging that provides a high level of tamper resistance.

CONCLUSION

The specific focus of the project is towards the security issues of in-network aggregation algorithms to compute aggregates such as predicate Count and Sum. The falsified sub-aggregate attack launched by a few compromised nodes can inject arbitrary amount of error in the base station's estimate of the aggregate. This concept is committed an attack-resilient computation algorithm which would guarantee the successful computation of the aggregate even in the presence of the attack and RPS algorithm is used at the receiver end to filter out the original data and stores it into the database.

REFERENCES

[1] Sankardas Roy, Mauro Conti, Member, Sanjeev Setia, and Sushil Jajodia, "Secure Data Aggregation in Wireless Sensor Networks: Filtering out the Attacker's Impact" Proc. IEEE, vol. 9, no. 4, pp.681-694, Apr 2014.

[2] S. Roy, M. Conti, S. Setia, and S. Jajodia, "Secure data aggregation in wireless sensor networks," Proc. IEEE, vol. 7, no. 3, pp. 1040–1052, Jun. 2012.

[3] T. Ko, J. Hyman, E. Graham, M. Hansen, S. Soatto, and D.Estrin,"Embedded imagers: Detecting, localizing, and recognizing objects and events in natural habitats," Proc. IEEE, vol. 98, no. 11, pp.1934–1946, Nov. 2010.

[4] P. Corke, T. Wark, R. Jurdak, W. Hu, P. Valen, cia, and D.Moore, "Environmental wireless sensor networks," Proc. IEEE, vol. 98, no. 11, pp. 1903–1917, Nov. 2010.

[5] D. Wagner, "Resilient aggregation in sensor networks," in Proc. ACM Workshop Security Sensor Adhoc Netw. (SASN), October 25.2004, pp. 68–79.

[6] Lei Yu, Jianzhong Li, Siyao Cheng, Shuguang Xiong, and Haiying Shen, " Secure Continuous Aggregation in Wireless Sensor Networks" Proc. IEEE, vol. 25, no. 3, pp. 762–774, Nov. 2004.

[7] L. Hu and D. Evans, "Secure aggregation for wireless networks," in *Proc. Workshop Security Assurance Ad Hoc Netw.*, 2003, pp. 384–391.

[8] H. Chan, A. Perrig, and D. Song, "Secure hierarchical in-network aggregation in sensor networks," in *Proc. ACM Conf. Comput. Commun. Security (CCS)*, 2006, pp. 278–287.

[9] B. Chen and H. Yu, "Secure aggregation with malicious node revocation in sensor networks," in *Proc. 31st Int. Conf. Distrib. Comput. Syst. (ICDCS)*, 2011, pp. 581–592.

[10] L. Buttyan, P. Schaffer, and I. Vajda, "Resilient aggregation with attack detection in sensor networks," in *Proc. 2nd IEEE Workshop Sensor Netw.Syst. Pervasive Comput.*, Mar. 2006, pp. 331–336.

[11] B. Przydatek, D. Song, and A. Perrig, "SIA: Secure information aggregation in sensor networks," in *Proc. 1st Int. Conf. Embedded Netw. Sensor Syst. (SenSys)*, 2003, pp. 255–265.

[12] K. Frikken and J. A. Dougherty, "An efficient integrity-preserving scheme for hierarchical sensor aggregation," in *Proc. 1st ACM Conf Wireless Netw. Security (WiSec)*, 2008, pp. 68–76.

[13] S. Nath, H. Yu, and H. Chan, "Secure outsourced aggregation via one-way chains," in *Proc. 35th SIGMOD Int. Conf. Manag. Data*, 2009, pp. 31–44.