International Journal of Research
Available at https://edupediapublications.org/journals
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 03Issue 10
June2016

Computing Data Aggregation with Secure in Wireless Sensor Networks

[1] V.S.RAHUL KUMAR , [2] A.L.SREENIVASULU

[1] *M.Tech, Dept of CSE, Intellectual Institute of Technology, Affiliated to JNTUA, AP, India*

[2] *Associate Professor, Dept of CSE, Intellectual Institute of Technology, Affiliated to JNTUA, AP, India*

Abstract— Wireless sensor networks (WSNs) are constrained in terms of memory, computation, communication, and energy. To reduce communication overhead and energy expenditure in (WSNs), data aggregation is used. Data aggregation is a very important technique, but it gives extra opportunity to the adversary to attack the network, inject false messages into the network and trick the base station to accept false aggregation results.. Lately, the research neighborhood has proposed a amazing aggregation framework referred to as synopsis diffusion which combines multi path routing schemes with duplicate-insensitive algorithms to safely compute aggregates (e.g., predicate count, Sum) in spite of message losses as a consequence of node and transmission screw ups. Nevertheless, this aggregation framework does not advert-gown the concern of false sub-aggregate values contributed by com-promised nodes leading to tremendous blunders within the combination computed at the base station, which is the foundation node within the aggregation hierarchy. That is an most important obstacle on the grounds that sensor networks are tremendously prone to node compromises due to the unattended nature of sensor nodes and the shortage of tamper-resistant hardware.

In this paper, we make the synopsis diffusion strategy cozy in opposition to assaults where compromised nodes make a contribution false sub-combination values. In specific, we gift a novel lightweight verification algorithm wherein the bottom station can assess if the computed combination (predicate rely or Sum) includes any false contribution. Thorough theoretical analysis and wide simulation be trained show that our algorithm outperforms different current approaches. Regardless of the network size, the per-node verbal exchange overhead in our algorithm is $O(1)$.

Index Terms—Base station, data aggregation, hierarchical aggregation, in-network aggregation, sensor network security, synopsis diffusion.

I. INTRODUCTION

Wireless sensor networks (WSNs) gained popularity because of the fact that they can be used to solve the real world challenges with low cost. WSNs are used in a variety of applications such as habitat monitoring and target tracking etc but these networks are constrained in terms of resources such as memory, communication, computation, and energy. WSNs consist of a large number of low power and low energy sensing devices called nodes. In addition to these nodes, there may be one or more powerful devices called base stations. Base station controls the network and processes the data collected by the sensor nodes. The sensor nodes sense and collect data from the environment and send it to the base station which performs further query on the collected data. The sensor nodes may be deployed in vicinity to each other as the number of nodes in the network may be very large. Due to this vicinity sensors may collect and transmit redundant data. Since the transmission of data costs much high than the computation, it is usually advantageous to organize the sensor nodes in clusters. The data are processed locally within the network and the aggregated data are sent to the base station in cluster environment.

Many protocols have been proposed for secure data aggregation in WSNs to reduce the communication overhead and the energy expenditure. Generally the network is divided in to a tree topology that is rooted at the base station. The sensor nodes sense data from the environment, the aggregators aggregate the data from the sensor nodes and send the data to the base station. Base station performs further query on the data. Data aggregation reduces communication overhead significantly but it makes the security more difficult. Any compromise node can forge data or can inject false data in to the network and thus one compromised node can alter the final aggregation. In general the data aggregation reduces the communication overhead but it opens new doors to the adversary and the aggregated data can easily be attacked by the adversary

In this paper, we design an algorithm to compute aggregates, such as Count and Sum, and to enable the base station to verify if the computed aggregate is valid. We call this algorithm the *verification algorithm* , though strictly speaking, it is an aggre-gate computation and verification algorithm. The key observa-tion which we exploit to minimize the communication overhead of this algorithm is that to verify the correctness of the final synopsis (the aggregate of the whole network) the base station does not need to receive authentication messages from all of the nodes. We validate the performance of our algorithm via both theoretical analysis and simulation. Irrespective of the network size, the per-node communication overhead in our verification algorithm is $O(1)$, while that of the least expensive existing al-gorithm (which is [6]) is $O(\log A)$, where A is the value of the aggregate, Count or Sum.

It is to be noted that while our algorithm is designed having WSNs in mind, it is straightforward to extend our solution for secure aggregation query processing in a large-scale distributed database system over the Internet [6].

The rest of this paper is organized as follows. Section II reviews the body of related work, and Section III briefly presents the synopsis diffusion approach. Section IV describes the problem statement and the assumptions, and Section V discusses our verification protocol. Section VI presents the simulation results, and Section VII concludes this paper.

II. Related Work

Several researchers have studied problems related to data ag-gregation in WSNs.

A. Data Aggregation Without any Provision for Security

The tiny aggregation service (TAG) to compute aggregates, such as Count and Sum, using tree-based aggregation algorithms were proposed. Moreover, tree-based aggregation algorithms to compute an order-statistic have been proposed . To address the communication loss problem in tree -based algorithms the authors designed an aggregation framework called synopsis diffusion to compute Count and Sum, which uses a ring topology. Authors independently proposed very similar algorithms. These works use duplicate-insensitive algorithms for computing aggregates based on the algorithm for counting distinct elements in a multiset.

B. Secure Aggregation Techniques

Several secure aggregation algorithms have been proposed assuming that the base station is the only aggregator node in the network. It is not straightforward to extend these works for verifying in-network aggregation unless we direct each node to send an authentication message to the base station, which is a very expensive solution. Only recently, the research community has been paying attention to the security issues of hierarchical aggregation.

A tree- based verification algorithm was designed in by which the base station can detect if the final aggregate, Count or Sum, is falsified. We are unable to extend this idea for verifying a synopsis because the synopsis computation is duplicate - insensitive. A verification algorithm for computing Count and Sum within the synopsis diffusion approach was designed in [6]. Our algorithm has some similarity with [6] except the fact that our algorithm attempts to further reduce the communication

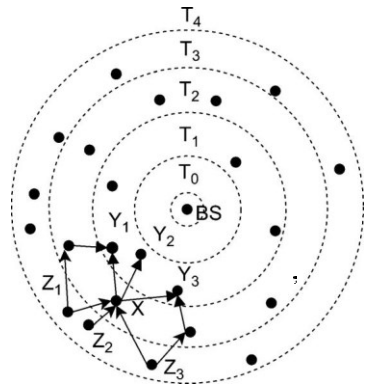

Fig. 1. Synopsis diffusion over a ring topology—A node may have multiple parents, e.g., X has three parents, Y_1 , Y_2 , Y_3 .

overhead in a novel approach. In addition, we provide extensive theoretical analysis to find the best tradeoff between the security and communication overhead. Recently, a few novel protocols have been proposed for "secure outsourced aggregation"; however, these algorithms are not designed for WSNs.

III. EXISTING SYSTEM

In a large sensor network, in-network data aggregation significantly reduces the amount of communication and energy consumption. Recently, the research community has proposed a robust aggregation framework called synopsis diffusion which combines multipath routing schemes with duplicate-insensitive algorithms to accurately compute aggregates (e.g., predicate Count, Sum) in spite of message losses resulting from node and transmission failures. However, this aggregation framework does not address the problem of false sub-aggregate values contributed by compromised nodes resulting in large errors in the aggregate computed at the base station, which is the root node in the aggregation hierarchy. This is an important problem since sensor networks are highly vulnerable to node compromises due to the unattended nature of sensor nodes and the lack of tamper-resistant hardware.

DISADVANTAGES:

In spite of the diverse applications, sensor networks pose a number of unique technical challenges due to the following factors:

Ad hoc deployment: Most sensor nodes are deployed in regions which have no infrastructure at all. A typical way of deployment in a forest would be tossing the sensor nodes from an aero plane. In such a situation, it is up to the nodes to identify its connectivity and distribution.:

Unattended operation: In most cases, once deployed, sensor networks have no human intervention. Hence the nodes themselves are responsible for reconfiguration in case of any changes.

Untethered: The sensor nodes are not connected to any energy source. There is only a finite source of energy, which must be optimally used for processing and communication.

An interesting fact is that communication dominates processing in energy consumption. Thus, in order to make optimal use of energy, communication should be minimized as much as possible.

Dynamic changes: It is required that a sensor network system be adaptable to changing connectivity (for e.g., due to addition of more nodes, failure of nodes etc.) as well as changing environmental stimuli.

Thus, unlike traditional networks, where the focus is on maximizing channel throughput or minimizing node deployment, the major consideration in a sensor network is to extend the system lifetime as well as the system robustness.

Survey Focus: A number of papers propose solutions to one or more of the above problems. Our survey focuses on the suggested solutions in the following areas:

Energy Efficiency: Energy efficiency is a dominant consideration no matter what the problem is. This is because sensor nodes only have a small and finite source of energy. Many solutions, both hardware and software related, have been proposed to optimize energy usage.

Localization: In most of the cases, sensor nodes are deployed in an ad hoc manner. It is up to the nodes to identify themselves in some spatial co-ordinate system. This problem is referred to as localization.

Routing: Communication costs play a great role in deciding the routing technique to be used.

Traditional routing schemes are no longer useful since energy considerations demand that only essential minimal routing be done

Besides the above topics, we will also look at some proposed sensor network systems. We also have a quick look at some of the simulators available today for simulating sensor networks

IV. PROPOSED SYSTEM

we make the synopsis diffusion approach secure against attacks in which compromised nodes contribute false sub aggregate values. In particular, we present a novel lightweight verification algorithm by which the base station can determine if the computed aggregate (predicate Count or Sum) includes any false contribution. Thorough theoretical analysis and extensive simulation study show that our algorithm outperforms other existing approaches. Irrespective of the network size, the per-node communication overhead in our algorithm. Here we are going to reduce communication hierarchy by selecting the effective parent and effective child selection manner from paper concept. By this manner the communication rate will be reduce in good manner for the same operation

VII. CONCLUSION

In this work we have studied the two most important parts of data communication in sensor networks- query processing, data aggregation and realized how communication in sensor networks is different from other wireless networks. Wireless sensor networks are energy constrained network. Since most of the energy consumed for transmitting and receiving data, the process of data aggregation becomes an important issue and optimization is needed. Efficient data aggregations not only provide energy conservation but also remove redundancy data and hence provide useful data only. The security issues of in-network aggregation algorithms to compute aggregates such as predicate Count and Sum. A compromised node can corrupt the aggregate estimate of the base station, keeping our focus on the ring-based hierarchical aggregation algorithms. To address this problem, we presented a lightweight verification algorithm which would enable the base station (BS) to verify whether the computed aggregate was valid. For future work, we plan to design an efficient attack-resilient computation algorithm. This algorithm would guarantee the successful computation of the aggregate even in the presence of an attack..

REFERENCES

[1] James Reserve Microclimate and Video Remote Sensing 2006 [Online]. Available: http://research.cens.ucla.edu

[2] S. Madden, M. J. Franklin, J. M. Hellerstein, and W. Hong, "TAG: A tiny aggregation service for ad hoc sensor networks," in *Proc. 5th USENIX Symp. Operating Systems Design and Implementation (OSDI)*, 2002.

[3] J. Zhao, R. Govindan, and D. Estrin, "Computing aggregates for monitoring sensor networks," in *Proc. 2nd Int. Workshop Sensor Network Protocols Applications*, 2003.

[4] J. Considine, F. Li, G. Kollios, and J. Byers, "Approximate aggrega-tion techniques for sensor databases," in *Proc. IEEE Int. Conf. Data Engineering (ICDE)*, 2004.

[5] S. Nath, P. B. Gibbons, S. Seshan, and Z. Anderson, "Synopsis diffusion for robust aggregation in sensor networks," in *Proc. 2nd Int. Conf. Embedded Networked Sensor Systems (SenSys)*, 2004.

[6] M. Garofalakis, J. M. Hellerstein, and P. Maniatis, "Proof sketches: Verifiable in-network aggregation," in *Proc. 23rd Int. Conf. Data Engineering (ICDE)*, 2007.

[7] M. B. Greenwald and S. Khanna, "Power-conservative computation of order-statistics over sensor networks," *Proc. 23th SIGMOD Principles of Database Systems (PODS)*, 2004.

[8] P. Flajolet and G. N. Martin, "Probabilistic counting algorithms for data base applications," *J. Computer Syst. Sci.*, vol. 31, no. 2, pp. 182–209, 1985.

[9] D. Wagner, "Resilient aggregation in sensor networks," in *Proc. ACM Workshop Security of Sensor and Adhoc Networks (SASN)*, 2004.

[10] L. Buttyan, P. Schaffer, and I. Vajda, "Resilient aggregation with attack detection in sensor networks," in *Proc. 2nd IEEE Workshop Sensor Networks and Systems for Pervasive Computing*, 2006.
[11] B. Przydatek, D. Song, and A. Perrig, "SIA: Secure information aggregation in sensor networks," in *Proc. 1st Int. Conf. Embedded Networked Sensor Systems (SenSys)*, 2003.
[12] H. Chan, A. Perrig, and D. Song, "Secure hierarchical in-network aggregation in sensor networks," in *Proc. ACM Conf. Computer and Communications Security (CCS)*, 2006.
[13] K. B. Frikken and J. A. Dougherty, "An efficient integrity-preserving scheme for hierarchical sensor aggregation," in *Proc. 1st ACM Conf. Wireless Network Security (WiSec)*, 2008.
[14] Y. Yang, X. Wang, S. Zhu, and G. Cao, "SDAP: A secure hop-by-hop data aggregation protocol for sensor networks," in *Proc. Seventh ACM Int. Symp. Mobile Ad Hoc Networking and Computing (MobiHoc)*, 2006.
[15] S. Nath, H. Yu, and H. Chan, "Secure outsourced aggregation via one-way chains," in *Proc. 35th SIGMOD Int. Conf. Management of Data*, 2009.
[16] L. Hu and D. Evans, "Secure aggregation for wireless networks," in *Proc. Workshop Security and Assurance in Ad hoc Networks*, 2003.
[17] H. Yu, "Secure and highly-available aggregation queries in large-scale sensor networks via set sampling," in *Proc. Int. Conf. Information Processing in Sensor Networks*, 2009.
[18] S. Roy, S. Setia, and S. Jajodia, "Attack-resilient hierarchical data aggregation in sensor networks," in *Proc. ACM Workshop Security of Sensor and Adhoc Networks (SASN)*, 2006.
[19] A. Perrig, R. Szewczyk, V. Wen, D. Culler, and J. D. Tygar, "SPINS: Security protocols for sensor networks," in *Proc. Int. Conf. Mobile Computing and Networks (MobiCom)*, 2001.
[20] W. He, X. Liu, H. Nguyen, K. Nahrstedt, and T. F. Abdelzaher, "Pda: Privacy-preserving data aggregation in wireless sensor networks," in *Proc. IEEE Int. Conf. Computer Communications (INFOCOM)*, 2007.
[21] S. Roy, M. Conti, S. Setia, and S. Jajodia, Secure data aggregation in wireless sensor networks 2011 [Online]. Available: http://mason.gmu.edu/~sroy1/AggVer.pdf, http://www.few.vu.nl/~mconti/papers/AggVer.pdf