SDA: SECURE DATA AGGREGATION TECHNIQUE FOR WSN IN THE PRESENCE OF COLLUSION ATTACKS

Riyad A M, Assistant Professor of Computer Science, EMEA College of Arts & Science,

Kondotty, Malappuram(dist.), Kerala

Abstract:

Since the peoples have limited energy and computing capacity, To aggregate data from many sensor nodes using basic techniques such as averaging. WSNs are often unattended, making them very susceptible to node compromise attempts. As a result, determining the reliability of data and the reputation of sensor nodes is critical for WSN. Iterative Filtering methods were discovered to be very useful for this purpose. These algorithms aggregate data and offer trustworthiness evaluation to nodes in the form of weight factors. These algorithms aggregate data from many sources at the same time and give a trust estimate of these sources, often in the form of matching weight factors applied to data supplied by each source. In this article, examined various safe data aggregation methods and presented a novel complex collision attack with implications for wireless sensor networks.

Keywords: WSN, SDA, IF, Sensor Nodes, Collision Attack

1 INTRODUCTION

Wireless sensor networks are rapidly being used in a wide range of applications; nevertheless, computing power and energy resources are two major difficulties for wireless sensor networks. Due to their limitations, sensor networks rely on a basic technique known as averaging for data aggregation [1]. Data aggregation based on a basic averaging method is more vulnerable to flaws and malicious attempts. By controlling hacked nodes, an attacker may capture and compromise sensor nodes and execute a range of attacks. Cryptographic techniques cannot prevent this since attackers often get full access to information stored on compromised nodes. To counteract this danger, sensor node trust levels must be established and node trustworthiness ratings must be adjusted [2]. Trust and reputation systems play an essential role in assisting the functioning of a broad variety of distributed systems, from wireless sensor networks to social networks, by giving an estimate of the trustworthiness of such distributed system members. An estimate of trustworthiness at any one time reflects an aggregate of the conduct of the participants up to that point and must be resilient in the face of different kinds of flaws and malevolent activity.

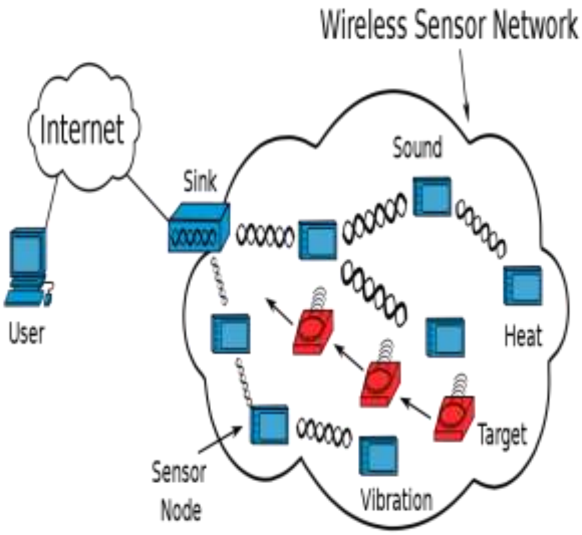


Figure 1: WSN Architecture

Attackers may influence the trust and reputation ratings of members in a distributed system in a variety of ways, and such manipulation generally has a negative impact on the system's performance. Iterative Filtering (IF) methods are an effective and trustworthy alternative for wireless sensor networks because they address both data aggregation and data trustworthiness estimation issues with a single iterative process [10]. Future aggregator nodes will be capable of executing increasingly complex data aggregation techniques when the computing capacity of extremely low power processors substantially increases, making wireless sensor networks less susceptible.

2 BACKGROUND WORK

Simple data aggregation is vulnerable to node compromise attacks and generates incorrect data. Furthermore, the current approach lacks accuracy and speed while aggregating data, lowering performance in the face of non-stochastic errors such as malfunctions and malicious attacks [3]. It is readily affected by leveraging fake data injection via a number of hacked nodes by simple data aggregation. Robust data aggregation is a major issue in WSNs, and a number of studies have been published that investigate harmful data injection while taking into consideration the different adversary models. The primary focus of the study is on iterative filtering algorithms, WSN certainty and reputation systems, and safe data aggregation with compromised node detection in WSNs. A number of papers have been published that introduce IF methods for addressing data aggregation problems [4], [5], and [6]. Li and colleagues the article presented by H.L.Shi [7] presents six alternative methods, all of which are iterative and identical, with the only variation being the choice of norm and aggregation function. In [8,] Ayday et al. presented a somewhat modified iterative method. Their major distinctions from previous algorithms are: 1) the ratings include a time-discount factor, so their significance fades with time; and 2) the system keeps a blacklist of individuals who are particularly poor raters. Liao et al. developed an iterative method that, in addition to the social network of users, makes use of the rating matrix. This paper presented various known iterative filtering algorithms that, although considerably more resistant to collusion attacks than simple averaging techniques, are nevertheless vulnerable to a new sophisticated collusion attack [9] this paper propose. Simulation on synthetically produced data sets is used to validate IF's performance. The simulation findings show that the robust aggregation method is both successful in terms of resilience against new sophisticated attacks and capable in terms of computing cost. Sensor errors are calculated based on biassed and unbiased measurements in a given location. IF outperforms the other approach in terms of accuracy and collusion resistance.

3 SDA: SYSTEM MODEL

This method offers an unconditionally secure approach for multicast network coding that is resistant to duplicate data. This approach enables intermediate nodes and destinations to validate the data origin and integrity of incoming messages without decoding, detecting and discarding malicious messages that fail the validation. It is critical to remember that destinations must receive a sufficient number of uncorrupted messages in order to decode and retrieve the full file provided by the source. This approach, on the other hand, gives destinations the option to filter out corrupted messages and have them filtered out by intermediary nodes as well.

The network-coded content distribution method enables intermediary nodes to identify malicious packets injected into the network and to notify adjacent nodes when a malicious packet is discovered. It generates hash values of the encoded blocks of data using a homomorphic hash function, which are then transmitted to the intermediate nodes and destinations prior to the encoded data. These hash values are sent via a pre-established secure channel. This section focuses mostly on data origin authentication and data integrity. It provides security against false data and attacks in particular because selected intermediary nodes are able to check the authentication tags of the packets received while being unable to decode the contents, and therefore identify and reject the fake data while receiving the data.

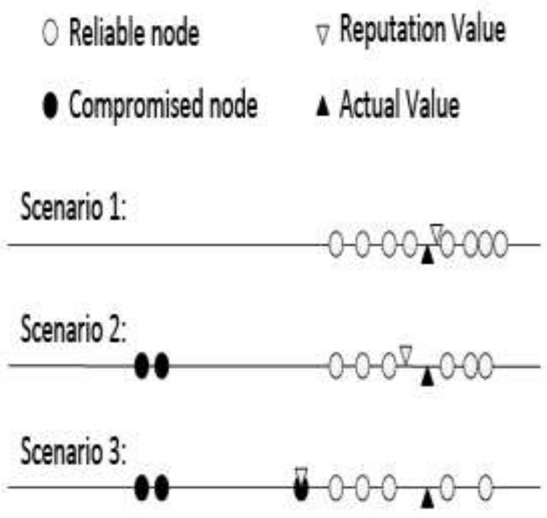


Figure 2: Attack scenario

3.1 Resource

A resource is made up of a large number of tiny nodes with limited computing capability, memory space, power resources, and short-range communication. Randomly, one node acts as the Resource-head, gathering data from other nodes and sending it to the Recipient through verifying nodes. Small nodes are classified as Resources. One node is chosen at random to be the Resource-head in each Resource. To balance energy usage, all nodes within a Resource rotate through the role of Resource-head. That is, there is no physical distinction between a Resource-head and a regular node since the Resource-head serves the same sensing function as the typical node.

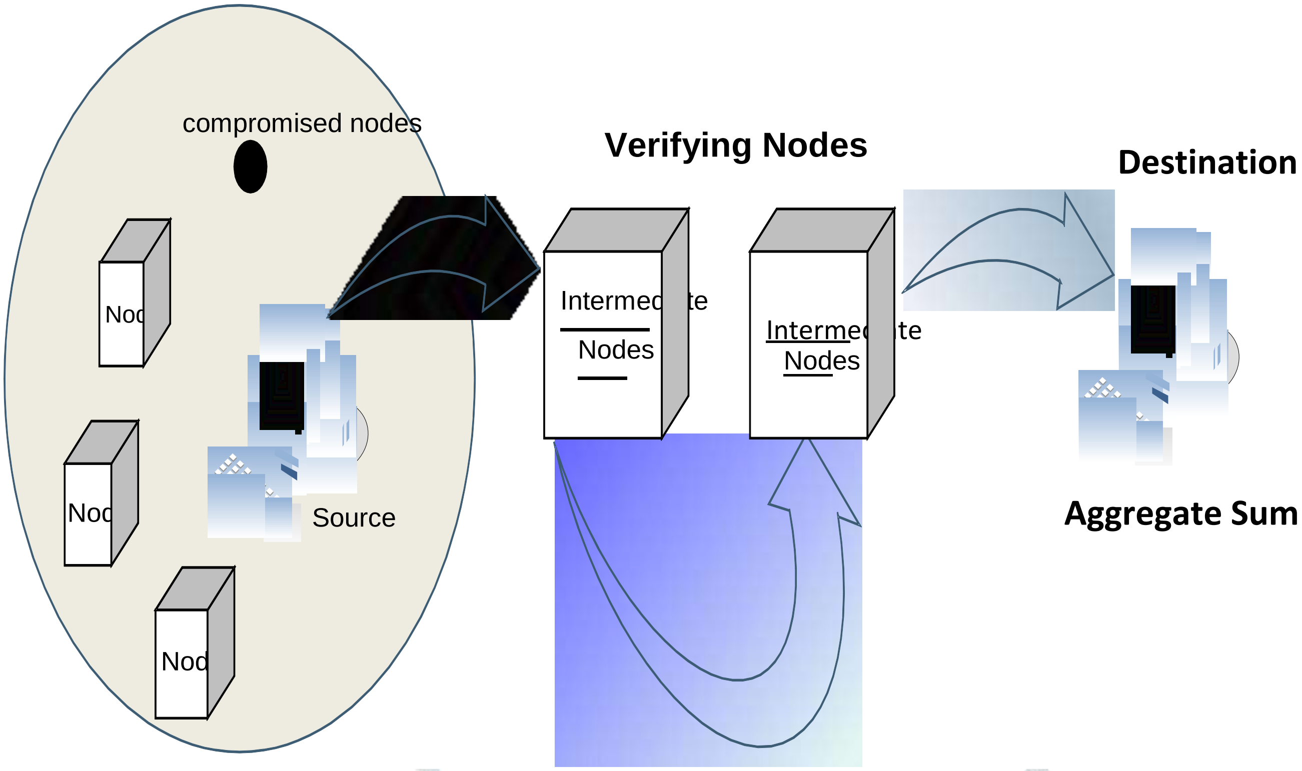


Figure 3: SDA Architecture

3.2 Authentication Code Generation

To meet these requirements, messages are appended at the source with either a digital signature, a message authentication code (MAC), or an authentication code (also called tag). To begin, MAC and authentication codes guarantee data integrity and data origin verification, while digital signatures ensure no repudiation. Second, MACs, authentication codes, and digital signatures should be classified according to the kind of security they provide: computational security (susceptible to an adversary with limitless computing capabilities) or unconditional security (i.e., robust against an attacker that has unlimited computational resources). The authentication is created here depending on the report provided by the resource node.

3.3 Verifying Nodes (or) Intermediate Nodes

It is especially important in the context of false information that may be sent to destination nodes, but intermediate nodes can also verify the validity of the packets. Such network nodes are referred to as verification nodes. Every node uses a fresh key to approve its reports and then exposes the key to verifying nodes. The verifying nodes may validate the reports using the dispersed and exposed keys. In this system, each node may monitor its neighbors by overhearing their broadcasts, preventing hacked nodes from altering their reports. At each hop, each verifying node performs report verification and key disclosure. Unless and until the reports are dropped or delivered to the base station. They may also function as verification nodes for other resource nodes.

3.4 Transmission Attacks

The hacked nodes may transmit fraudulent reports including fabricated or nonexistent events that "occur" in their resource. Furthermore, if they have enough secret knowledge, they may mimic certain uncompromised nodes of another resource or intermediary node and report the faked events "occurring" inside those nodes. These erroneous reports not only create false alarms at the Recipient, but they also deplete the limited energy of intermediate nodes. Completed data is lost as a result of node transmission failure.

3.5 Base Station

Fake report injection attacks are one kind, in which attackers inject into networks false data reports comprising nonexistent events or forged readings from compromised nodes. These attacks not only result in false alerts at the Recipient. The resource (cluster head) node receives data from other nodes and sends it to the recipient through intermediate nodes. As a result, the receiver may recover the whole original data by using the homomorphic hash function to validate the transmission at each node.

3.6 Aggregation

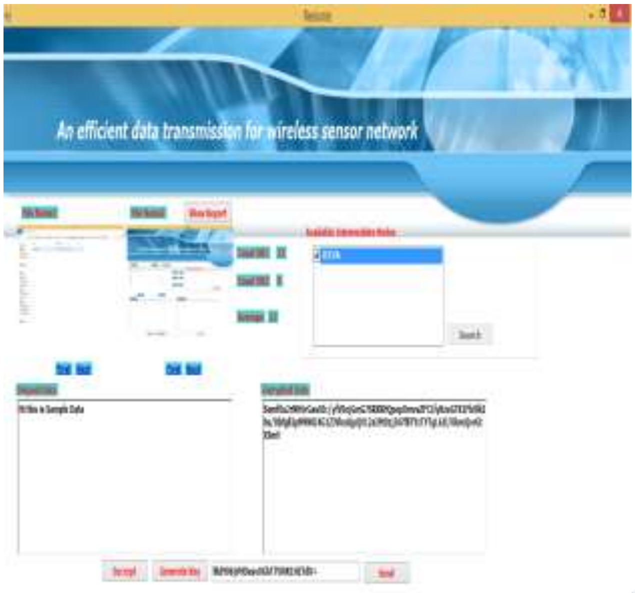
The scientific community considers Count and Sum to be significant aggregates. It is worth noting that these aggregates are easily generalizable to the predicate Count (e.g., number of sensors whose transferred data received in base station without any loss). In addition, Average may be calculated using Count and Sum. A Sum method may easily be modified to calculate Standard Deviation and Statistical Moments of any order.

3.7 Algorithm 1: Iterative filtering algorithm.

```
Input: X; n;m.
Output: The reputation vector r
        l← 0;
        $w^{(0)}$ ←1;
repeat
        Compute $r^{(l+1)}$;
        Compute d;
        Compute $w^{(l+1)}$;
        l ←l + 1;
until reputation has converged;
```

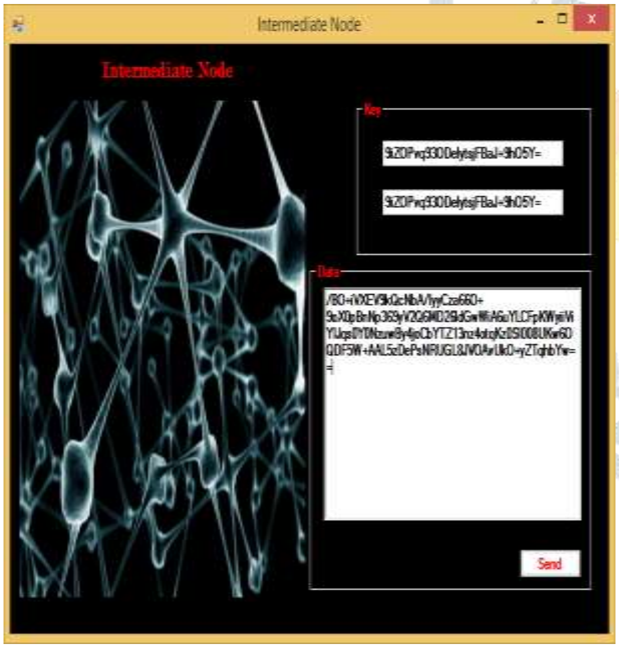
4 RESULTS AND DISCUSSION

In the case of the basic assault scenario, the results of this experiment clearly demonstrate that this initial trustworthiness has no detrimental impact on the performance of the IF algorithm with both discriminant functions. In the next section, this paper demonstrate how these starting values enhance the IF algorithm in the suggested collusion attack scenario. The findings of this experiment confirm that this complex attack scenario is triggered by a weakness identified in the IF algorithms, which drastically reduces the contributions of benign sensor nodes when one of the sensor nodes returns a number extremely near to the simple average.


Source
An efficient data transmission for wireless sensor network
Show Report
Available Intermediate Nodes
Search
Original Data
Encrypted Data
Encrypt
Generate Key
Send

Hi this is Sample Data

Figure 4: Sender sends the Encrypted Data with Key

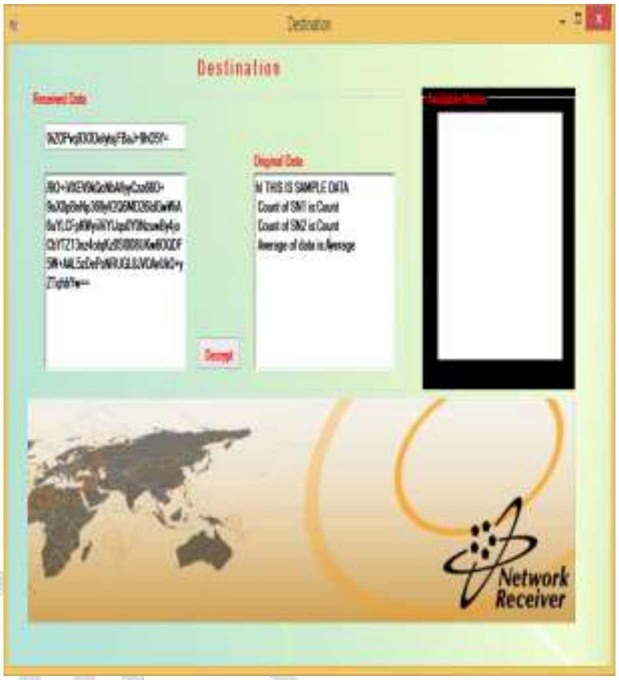


9tZOPwq930DelytsjFBaJ-9hO5Y=

9tZOPwq930DelytsjFBaJ-9hO5Y=

/BO+iVXEV9kQcMbA/lyyCza66O+
9eX0pBnNp369yV2Q6MD29dGwWiA6uYLCFpKWyaVi
YUqs0YtNzuwBy4joCbYTZ13nz4otqKz0SI008UKw6O
QDF5W+AAL5zDePsNRUGL8JVOAwUkO+yZTqhbYw=
=

Figure 5: Intermediate Node transferring the data and key to receiver



9tZOPwq930DelytsjFBaJ-9hO5Y=

/BO+iVXEV9kQcMbAlyyCza66O+
9eX0pBnNp369yV2Q6MD29dGwWiA
6uYLCFpKWyaVi YUqs0YtNzuwBy4jo
CbYTZ13nz4otqKz0SI008UKw6OQDF
5W+AAL5zDePsNRUGL8JVOAwUkO+y
ZTqhbYw==

hi THIS IS SAMPLE DATA
Count of SN1 is Count
Count of SN2 is Count
Average of data is Average

Figure 6: Receiver receives the encrypted data and key for decrypting the content

5 CONCLUSIONS

Data aggregation methods and data averaging strategies are investigated. The assumptions of adversary models are examined. New complex collusion attack scenarios are described, as well as their effect on wireless sensor networks. Future aggregator nodes will be capable of executing increasingly complex data aggregation techniques when the computing capacity of extremely low power processors substantially increases, making wireless sensor networks less susceptible. In the future, an improved approach against collusion attacks will be implemented, making collusion not only more resilient, but also more accurate and quicker convergence.

6 REFERENCES

[1] S. Ozdemir and Y. Xiao, “Secure data aggregation in wireless sensor networks: A comprehensive overview,” Comput. Netw., vol. 53, no. 12, pp. 2022–2037, Aug. 2009.

[2] A. Jøsang and J. Golbeck, "Challenges for robust trust and reputation systems," in Proceedings of the 5 th International Workshop on Security and Trust Management, Saint Malo, France, 2009.

[3] H.-S. Lim, Y.-S. Moon, and E. Bertino, "Provenance-based trustworthiness assessment in sensor networks," in Proceedings of the Seventh International Workshop on Data Management for Sensor Networks, ser. DMSN '10, 2010, pp. 2–7.

[4] K. Hoffman, D. Zage, and C. Nita-Rotaru, "A survey of attack and defense techniques for reputation systems," ACM Comput. Surv., vol. 42, no. 1, pp. 1:1–1:31, Dec. 2009.

[5] C. de Kerchove and P. Van Dooren, "Iterative filtering in reputation systems," SIAM J. Matrix Anal. Appl., vol. 31, no. 4, pp. 1812–1834, Mar. 2010.

[6] C. T. Chou, A. Ignatovic, and W. Hu, "Efficient computation of robust average of compressive sensing data in wireless sensor networks in the presence of sensor faults," Parallel and Distributed Systems, IEEE Transactions on, vol. 24, no. 8, pp. 1525–1534, Aug 2013.

[7] Y. Zhou, T. Lei, and T. Zhou, "A robust ranking algorithm to spamming," CoRR, vol. abs/1012.3793, 2010.

[8] E. Ayday, H. Lee, and F. Fekri, "An iterative algorithm for trust and reputation management," in Proceedings of the 2009 IEEE international conference on Symposium on Information Theory,Volume 3, ser. ISIT'09, 2009, pp. 2051–2055

[9] Y.-K. Yu, Y.-C. Zhang, P. Laureti, and L. Moret, "Decoding information from noisy, redundant, and intentionally distorted sources," Physica A Statistical Mechanics and its Applications, vol. 371, pp. 732–744, Nov. 2006.

[10] 11. P. Laureti, L. Moret, Y.-C. Zhang, and Y.-K. Yu, "Information filtering via Iterative Refinement," EPL (Europhysics Letters), vol. 75, pp. 1006–1012, Sep. 2006.