A Survey Analysis of Safe Data Collection in Wireless Sensor Network in the Existence of Collision Attacks

Pooja Motgi
M. Tech Student
Department of Computer Science and Engineering
AIET, VTU University,Kalaburagi

Prof. Ramesh Jadhav
Associate Professor
Department of Computer Science and Engineering
AIET, VTU University,Kalaburagi

Abstract

Due to a need for robustness of monitoring, WSN are usually redundant. Data from multiple sensors is aggregated at an aggregator node which then forwards the base station only the aggregate values. At present, due to limitations of computing power and energy store of sensor nodes, data is aggregated by exceptionally simple algorithms such as averaging. The present algorithm is to allow the base station to securely compute build Count or Sum even in the occurrence of such an attack. Our attack-resilient calculation algorithm computes the right aggregate by filtering out the assistances of compromised nodes in the aggregation order.
Keywords: Aggregation, WSN

I. INTRODUCTION

Because of a requirement for fitness of checking and minimal effort of the hubs, remote sensor systems (WSNs) are normally excess. Information from numerous sensors is collected at an aggregator hub which then advances to the base station just the total qualities. At present, because of confinements of the figuring force and vitality asset of sensor hubs, information is collected by to great degree straightforward calculations, for example, averaging. In any case, such conglomeration is known not exceptionally defenseless against deficiencies, and all the more vitally, malignant assaults. This can't be cured by cryptographic techniques, in light of the fact that the assailants by and large increase complete access to data put away in the traded off hubs. Consequently information collection at the aggregator hub must be joined by an appraisal of dependability of information from individual sensor hubs. In this way, better, more advanced calculations are required for information conglomeration later on WSN. Such a calculation ought to have two elements.

1) In the vicinity of stochastic mistakes such calculation ought to create gauges which are near the ideal ones in data theoretic sense. Accordingly, for instance, if the clamor present in every sensor is a Gaussian freely circulated commotion with zero mean, then the appraisal created by such a calculation ought to have a difference near the Cramer-Rao lower bound (CRLB), i.e. it ought to be near the fluctuation of the Maximum Likelihood Estimator (MLE). Be that as it may, such estimation ought to be accomplished without supplying to the calculation the fluctuations of the sensors, inaccessible practically speaking.
2) The calculation ought to likewise be powerful in the vicinity of non-stochastic mistakes, such as flaws And vindictive assaults, and, other than amassing information; such calculation ought to additionally give an appraisal of the unwavering quality and reliability of the information got from every sensor hub.

II. LITERATURE SURVEY

A. Ignatovic et.al [1] in Remote sensor systems (WSNs) empowers the gathering of physical estimations over an extensive geographic range. It is frequently the case that we are occupied with registering and following the spatial-normal of the sensor estimations over a district of the WSN. Shockingly, the standard normal operation is not strong in light of the fact that it is exceedingly defenceless to sensor issues and heterogeneous estimation clamour. In this paper, we propose a computational effective technique to figure a weighted normal (which we will call hearty normal) of sensor estimations, which suitably takes sensor flaws and sensor clamor into thought. We accept that the sensors in the WSN use arbitrary projections to pack the information also, send the compacted data to the data mix centre. The key favourable position of our proposed technique is that the information combination focus just needs to perform decompression once to register the vigorous normal, in this way enormously lessening the computational prerequisites. We apply our proposed technique to the information gathered from two WSN arrangements to show its productivity and precision.

W. Zhou et.al [2] as the trust issue in remote sensor systems is rising as one imperative variable in security plans, it is important to investigate how to oppose assaults with a trust plan. In this paper we classify different sorts of assaults and countermeasures identified with trust plans in WSNs. Moreover, we give the improvement of trust components, give a short rundown of established trust approaches and stress the difficulties of trust plan in WSNs. A broad summarizing so as to write study is introduced best in

class trust instruments in two classifications: secure directing and secure information. In view of the examination of assaults and the current research, an open field and future heading with trust systems in WSNs is given.

H. Liao et.al [3] in the Internet period the data over-burden and the test to recognize quality substance has raised the issue of how to rank both assets and clients in online groups. In this paper we add to a general positioning technique that can all the while assess clients' notoriety and items' quality in an iterative system, and that adventures the trust connections and social colleagues of clients as an extra wellspring of data. We test our strategy on two genuine online groups, Physics gathering and the Last.fm music inventory, and decide how distinctive variations of the calculation impact the resultant positioning. We demonstrate the advantages of considering trust connections, and characterize the type of the calculation better adept to normal circumstances.

M. C. Vuran et.al [4] in Remote Sensor Networks (WSN) are chiefly described by thick sending of sensor hubs which altogether transmit data about detected occasions to the sink. Because of the spatial connection between's sensor hubs subject to watched occasions, it may not be essential for each sensor hub to transmit its information. This paper indicates how the spatial relationship can be abused on the Medium Access Control (MAC) layer. To the best of our insight, this is the principal exertion which abuses spatial relationship in WSN on the MAC layer. A hypothetical system is produced for transmission regulation of sensor hubs under a mutilation imperative. It is demonstrated that a sensor hub can go about as an agent hub for a few other sensor hubs watching the corresponded information. In light of the hypothetical system, a conveyed, spatial Correlation-based Collaborative Medium Access Control (CC-MAC) convention is then composed which has two parts: Event MAC (E-MAC) and Network MAC (N-MAC). E-MAC sift through the relationship in sensor records while N-MAC organizes the transmission of course through bundles. Recreation results demonstrate that CC-MAC accomplishes superior in wording vitality, bundle drop rate, and dormancy.

Fan ZHANG et.al [5] in online rating systems are attractive increasingly prevalent due to the service for decision making. In these systems, the negative effects of one-sided ratings have been commonly studied in current literatures. The corresponding results are mainly separated into detection-based and tolerance-based methods. Though, detection-based methods price lots of additional system resources for occupied out and are sensitive to holds. Though tolerance-based methods might solve such complications to some amount, they either cannot bear attacks vigorously or are restricted to distinct ratings. Moreover, most current methods supervise the favourites of users. In other words, even the equal scores rated by two users in the similar dataset have meanings of unlike interest extent permitting to their individual preferences. Also, it is unreasonable to give ratings rated in unlike times with the similar weight, as clearly the current ratings reflect users' view better than ratings valued a extensive time ago. In this paper, we suggest that this system which uses rating pre-processing to contract with the former two difficulties, and then highlight recent ratings by addition varying time weights. We determine the effectiveness of our method by the experimentations on the real-life datasets of the Movie Lens.

III. IMPLEMENTATION

The work mainly consists of 4 modules. They are:

1) Setting up Network Model
2) Robust Data Aggregation
3) Enhanced Iterative Filtering
4) Accuracy with A Collusion Attack

A. *Modules Description*

1) Setting up Network Model

Our first module is setting up the system model. We consider a substantial scale, homogeneous sensor system comprising of asset compelled sensor hubs. The sensor hubs are partitioned into disjoint groups, and every bunch has a group head which goes about as an aggregator. Information are intermittently gathered and collected by the aggregator.

2) Robust Data Aggregation

We give a vigorous starting estimation of the reliability of sensor hubs to be utilized as a part of the main cycle of the IF calculation. The greater part of the customary factual estimation strategies for change include utilization of the specimen mean. Hence, proposing a vigorous change estimation strategy on account of skewed specimen mean is a fundamental piece of our system. We accept that the stochastic parts of sensor blunders are free irregular variables with a Gaussian conveyance; in any case, our analyses demonstrate that our strategy works entirely well for different sorts of mistakes with no adjustment. In addition, if mistake appropriation of sensors is either known or evaluated, our calculations can be adjusted to different dispersions to accomplish an ideal execution. Taking into account such an estimation of the inclination and difference of every sensor, the predisposition appraisal is subtracted from sensors readings and in the following period of the proposed system, we give an underlying assessment of the notoriety vector ascertained utilizing the MLE.

3) Enhanced Iterative Filtering

As indicated by the proposed assault situation, the assailant abuses the weakness of the IF calculations which starts from a wrong suspicion about the underlying reliability of sensors. Besides, the underlying weights for all sensor hubs can be figured in light of the separation of sensors readings to such an underlying notoriety. Our exploratory results represent that this thought not just solidifies the IF calculations against the proposed assault situation, yet utilizing this underlying notoriety enhances the productivity of the IF calculations by diminishing the quantity of emphases expected to approach a stationary point inside of the recommended resilience.

Accuracy with a Collusion Attack

In order to speak to the life of the proposed data all out system in the region of complex strikes, we misleadingly create a couple data sets by implanting the proposed course of action attacks. Thusly, we expect that the adversary uses (c < n) exchanged off sensor centres to dispatch the perplexing strike circumstance proposed. The aggressor uses the at first bartered centres to make peculiarity readings remembering the finished objective to skew the clear ordinary of all sensor readings. The foe then mutilates the last sensor readings by mixing the qualities close such skewed typical. This course of action attack circumstance makes the IF estimation to unite to a wrong stationary point. Remembering the final objective to look at the exactness of the IF counts with this interest strike circumstance, we misleadingly deliver a couple data sets with different qualities for sensors vacillations and moreover diverse number of exchanged off centres. The eventual outcomes of this test favour that our advanced attack circumstance is brought on by the discovered powerlessness in the IF estimations which mightily diminishes the duties of accommodating sensor centres when one of the sensor centres reports a quality close to the essential typical.

IV. METHODOLOGY

The following steps are used to perform encryption and decryption on a file.

- For a specified algorithm generate a key for a given array of bytes.
- Generate a cipher text for a specified algorithm.
- For encryption and decryption initialize the cipher text.

The algorithm includes 4 steps: key generation, key distribution, encryption and decryption. The algorithm includes a *public key* and a *private key*. In the System every user is aware of the public key and it is used to encrypt the messages. Here the main goal is that the messages which are encrypted by using the public key are only decrypted by using the user's private key but in a specified amount of time. The criteria on using this algorithm is that we have to select 3 large positive numbers as e, d, and n such that it should be modular exponentiation for m.

– Step 1: Key distribution

In key distribution mechanism the users has to send their public key as (n, e) in order to send their encrypted messages. But private key is not distributed. If Alice is sending message to Bob, then, Alice should generate a cipher text using his public key, where as Bob can decrypt the message using her private key, hence Alice's public key is known to all.

– Step 2: Encryption

Consider a message M that is to be transmitted between 2 users Alice and Bob. Firstly we need to calculate the integer value of M as m and find gcd (*m*, *n*) =1. Then generate the cipher text c for the given message m by using the corresponding users public key.

– Step 3: Decryption

The user Alice can regenerate m from c by applying their respective private key as exponent of d. For *m*, she can regenerate the original message *M* by reversing the padding scheme.

– Step 4: Key Generation

The keys in this algorithm are calculated in the following manner:
Select at random any 2 prime numbers such as p and q.
Compute n = pq.
Compute (n) = (p − 1)(q − 1) = n − (p + q − 1),
Select an integer e where 1 < e < (n) and gcd(e, (n)) = 1; i.e., e is coprime.
Calculate d as $d \equiv e^{-1} \pmod{(n)}$;
$d \cdot e \equiv 1 \pmod{(n)}$
e is released as the public key exponent.
d is kept as the private key exponent.

Public key includes mod n and the public (encryption) exponent e. The private key includes mod n and the private (decryption) exponent d, and should be kept secret. p, q, and (n) should also be kept secret as they are used to find the value of d.

V. CONCLUSION

In this paper, we presented a novel conspiracy assault situation against various existing IF calculations. Besides, we proposed a change for the IF calculations by giving an underlying estimation of the reliability of sensor hubs which makes the calculations conspiracy powerful, as well as more exact and speedier merging. In future work, we will explore whether our methodology can ensure against traded off aggregators. We likewise plan to actualize our methodology in a deploy sensor system.

REFERENCES

[1] S. Ozdemir and Y. Xiao, "Secure data aggregation in wireless sensor networks: A comprehensive overview," Computer. Networking., vol. 53, no. 12, pp. 2022–2037, Aug. 2009.
[2] L. Wasserman, All of Statistics : A Concise Course in Statistical Inference. New York, NY, USA: Springer,.
[3] Jøsang and J. Golbeck, "Challenges for robust trust and reputation systems," in Proc. 5th Int. Workshop Security Trust Manage., Saint Malo, France, 2009, pp. 253–262.

[4] K. Hoffman, D. Zage, and C. Nita-Rotaru, "A survey of attack and defense techniques for reputation systems," ACM Computer. Surveys, vol. 42, no. 1, pp. 1:1–1:31, Dec. 2009.
[5] R. Roman, C. Fernandez-Gago, J. Lopez, and H. H. Chen, "Trust and reputation systems for wireless sensor networks," in Security and Privacy in Mobile and Wireless Networking, S. Gritzalis, T. Karygiannis, and C. Skianis, eds.,Leicester, U.K.: Troubador Publishing Ltd, 2009 pp. 105–128.