RESEARCH ARTICLE OPEN ACCESS

Secure Data Aggregation Using Multi objective Metaheuristic Approach for WSN

Jaybhaye Chaitali *,P.M.Pawar**

** Department of Information Technology, Smt.Kashibai Navle College of Engineering,Pune.*

*** Department of Information Technology, Smt.Kashibai Navle College of Engineering,Pune.*

--************************----------------------------------

Abstract:

Wireless Sensor Network is consisting of group of sensor nodes which are able to sense data and to convey sensed information to the base station. Sensor nodes have finite energy and finite memory. For improving the life of sensor node it is essential to minimize the communication process. Data Aggregation is the useful method for extending the life and minimizing communication. Data aggregation combines the data from various nodes and summarizes the data. Sensor nodes convey aggregated data using wireless medium Sensor nodes are placed in the hostile and inaccessible areas the security is necessary factor. There is a possibility of attacks for this confidentiality integrity of data is necessary. So security is required for data aggregation for reliable data. This paper discusses the Multiobjective optimization and Metaheuristic approach to provide secure data aggregation.

Keywords----**Data Aggregation, Secure data aggregation, Multi object, Meta heuristic.**

--************************----------------------------------

I. INTRODUCTION

Wireless Sensor Networks are very useful methodology .Sensor nodes in the wireless networks obtain the data from the areas which are not able to accessible and hostile. These sensor nodes are able to covey data which is sensed; using wireless medium. Health care, Military, Forecast monitoring, Habitat monitoring etc. these are the fields in which wireless sensor network is used. A sensor network senses the data and conveys to base station as it has small memory and finite energy. To convey this data to base station requires more energy which decreases the life of the sensor .Lifetime of sensor is extended by minimizing communication for this Data Aggregation is useful. Data Aggregation is process of binding of data from different sensor node and summarizing this data while conveying this data to the base station. Network has thousands of sensor nodes in same area. These nodes supervise the location and collect the data. Sensor nodes in same the location collect the mutual data. Collected data contains repetitive data. Data aggregation assists in removing redundancy and preserving energy of sensor nodes.

Sensor nodes in wireless sensor networks are put in the remote and critical areas for supervising the location and collecting the data. Sensor nodes convey this sensitive data to base station. There is always a possibility of attack and vulnerability of the data. Protection to the data is very necessary because attacker may put some fake data or access data while sending data .Data is transmitted to the base station but raw data is not visible to the base station. For reliability of data which is arriving at the base Security to data aggregation is necessary

This paper discusses the secure data aggregation, Multi objective Metaheuristic approach and analyze the other methods for secured data aggregation. Hierarchical based data aggregation is used for aggregating data. In which tree hierarchy is used in aggregation process. In wireless sensor networks sensors are arranged in tree hierarchy. In which Base is the root node. Leaf node collects the information and conveys to the non-leaf node. Non leaf node does the work of aggregation, which performs aggregation function on the arrived data from the leaf node. This data which is aggregated is provides to the base station. In this way the tree hierarchy works. This gives the secured and energy efficient aggregation of data.

This in this local search methods are used and divide and conquer is used for the formation of the clusters. This works on homogeneous stationary node. In this first clusters are formed. The node having more energy and maximum storage capacity is selected as a cluster head; it also works as a data aggregation node. Cluster head conveys this aggregated data to base station. In the deployment phase public keys are assigned by the base station to the sensor nodes and cluster head assigns the private keys to sensor nodes in the cluster. Cluster head does periodic authentication and validation of sensor nodes. Intruder is detected if any node fails in authentication or if there is delay in validation process.

Greedy search method and local search methods are used to minimize the delay and for establishing the security.

II. SECURE DATA AGGREGATION.

A. Data Aggregation

Sensor nodes are resides in the areas which are not accessible and hostile. Sensor nodes convey the sensed sensitive data to the base station. Sensor nodes have restricted resources. These resources cannot be changed, if the battery of the node gets over and, node dies. Sensor node also has limited amount of memory due to this node cannot store the sensed data. Node has to continuously transfer the sensed data to sink node but it increases communication overhead. Data transfer rate is minimized by performing data aggregation. Data aggregation combines data from various nodes. Data aggregation is the process of taking data from other nodes and put it into summary in order to reduce size of the collected data. This way data aggregation extends life of sensor node and life of sensor network. It also removes redundant data and prevents transferring of similar data, sensed by two nearby sensor nodes in the same area. Data aggregation also helps in reducing network traffic and use bandwidth efficiently. Various methods are used to perform data aggregation e.g., using tree hierarchy method, cluster formation.

Cluster formation is efficient way for data transfer and saving energy. In cluster formation network is divided into clusters. Every cluster has a cluster head. A sensor node in the cluster transfers the data to the cluster head. Cluster head perform data aggregation and sends data to the base station. This is useful for large sensor networks. Cluster head requires more energy for aggregation, due to this cluster head dies early. New cluster head has to elect again.

***1) Centralized method*:** In this every node sends data through central node. This uses multihop protocol. Leader node receives packets from other nodes and aggregates that data for querying.

***2) In-Network method*:** In network way considers multihop network. Intermediate network performs processing of data to increase the lifetime of network. It uses two reductions way with size reduction and without size reduction. In with size reduction, length of packet is reduced for increasing lifetime of network. Without size reduction, only merges the data into a single packet without reducing packet length.

***3) Tree based hierarchy*:** In tree based root node is considered as a sink node, leaf nodes are source node and non-leaf nodes are aggregator node. Leaf nodes get data and transmit it to aggregator node. Aggregator node aggregates the data and transfers it to the sink node.

There are chances of attacks and putting some false data while conveying the data. So applying security to the aggregated data is secure data aggregation.

B. Secure Data Aggregation

Many secure algorithms are used for the secure data aggregation .Single aggregator model and multiple aggregator models are applied to secure Data aggregation .Only base station does the aggregation on the arrived data in single aggregated model. Base station collects the data which is validated and performs data aggregation. This aggregated data is conveyed to the trustworthy remote sever which does the verification of aggregated data. The multiple aggregator model more than one nodes does aggregation and conveys data to base station. Encryption gives confidentiality and message authentication gives message integrity. Data exchanged between two nodes is encrypted using public, encryption algorithm and message authentication code. Data confidentiality ensures that message is confidential it is not exposed while sending; Data confidentiality is achieved using public and private keys i.e. suing encryption process. Data Integrity ensures that data in not altered. Cyclic codes are and message authentication codes are used for integrity of data. Authentication mechanism used to detect the malicious attacks and spoofed packets. Intruder sends the data using fake identity so authentication is required. For authentication symmetric key cryptography is used. In symmetric key cryptography sender and receiver exchange their private keys. Using these private keys message authentication code is computed for doing the communication. Confidentiality, Integrity, data freshness, availability, authentication, non repudiation, data accuracy are the requirements for security. Data confidentiality does not provide access to unauthorized object. It is divided into hop by hop and end to end. Hop by hop cannot use in real work. Aggregation is directly used on encrypted data with the use of homomorphism function. Data freshness guarantees that data is fresh. Data availability tells about availability of network and accessibility of data. Authentication consists of entity and data authentication. Entity authentication detects if the data is received from authenticated server. Data authentication detects if the data is original or not. on repudiation states that packet is send and received by the correct person. Once the packet is sent, a person cannot deny.

Security is threatened by the two types of adversaries, active and passive adversaries.

***1) Active Adversary*:** Active adversary injects packets, destroys data, compromises data or node from nodes.

***2) Passive Adversary*:** Passive adversary works during data transmission process between two nodes. Passive adversary does eavesdropping on transmission of data and gets important information. Depend upon the network access adversaries' gets information. If adversary has total access then it is a strong adversary. This type of adversary has total access to the wireless sensor network. Passive adversary listen every communication in the network and active adversary accesses all the components of WSN. Partial access adversaries are less strong. Passive adversaries listen to all the communication between nodes subset. Active adversaries access subset of nodes

III. MULTIOBJECTIVE OPTIMIZATION AND METAHEURISTIC

A. Multiobjective Approach

Multiobjective function is considered when two or more than two objectives are functioning. Multiobjective optimization is the type of optimization. In Single objective optimization single objective is considered for maximizing and minimizing purpose. Multiobjective optimization considers more than two objects. Multiobjective optimization minimizes and maximizes the functions of objectives with vector which considers the number of constraints. Multiobjective optimization gives number of optimal solution. Pareto optimality, Weighted Sum Method, Utility method etc. are the methods of Multiobjective optimization. In secure data aggregation, two objectives are optimized. In this data transfer is minimized by performing data aggregation, helps to increase battery life. Second is security is provided to transferring data.

Multiobjective optimization does not have feasible solution that minimizes all functions of objectives at once like Pareto optimizes. Pareto Optimality is on the basis of concept of dominance. Pareto optimal solutions are enhanced by decreasing the one of the objective. Weighted sum method takes the sum of all the objectives and transforms the problem of Multiobjective optimization of vectors into the scalar problem.

To solve multi-objective problems some techniques are invented.

1) Optimization Methods: Optimization algorithms divided as finitely terminating, Convergent iterative algorithms and using heuristics or metaheuristic.

- Finitely terminating: These algorithms consist of simple algorithms and combinatorial algorithms.

- Convergent Iterative algorithms: These algorithms evaluate hessians and gradients.

2) Heuristic or Metaheuristic: algorithms: Heuristic algorithms are approximate algorithms, provides approximate solution. It gives solution to problems don't have an optimal solution. Metaheuristic also gives near optimal solution. Many algorithms metaheuristic do multiobjective optimization. Metaheuristic gives solution to difficult problems in reasonable time but it does not give optimal solution. Metaheuristic algorithms are not suitable for all the problems.

A. Metaheuristic approach

Metaheuristic arranges the communication between the local procedures and higher level methodologies which gives better solution of optimization problem. It is the heuristic of heuristic. Metaheuristic is the set of algorithms. These algorithms used to solve the NP-hard problem. NP-hard problems do not have solution an optimal solution and not solved in polynomial time. Metaheuristic guide heuristics which gives approximate solution. Metaheuristic even provides a solution when the data is fallible and underdone. Metaheuristic works by assuming some assumptions in

Metaheuristic works by assuming some assumptions in problem and it samples very large solution set wholly, but it does not assure that solution is globally optimal. Metaheuristic consists of intensification and diversification. It explores large data space to generate optimal solution. Local search methods in Metaheuristic are tabu search, simulated annealing, iterated local search etc. Evolutionary computation, ant colony optimization, genetic algorithms and particle swarm methods are global search methods. Metaheuristic is categorized into global search method, Local search, Population based, memetic and hybridization, Nature inspired heuristic and parallel heuristic etc.

Population based heuristic preserves multiple solutions. Hybridization algorithm gives a solution by combining metaheuristic with other optimization strategies. Parallel programming is used for multiple parallel metaheuristic searches.

Tabu search is on hypothesis that for problem solving requires the blend of adaptive memory and responsive exploration for qualifying intelligence. Simulation annealing, takes the initial solution. It goes through solution set and iteratively finds new solution near to the recent one. Ant colony optimization is based on the behavior of ants. It is like a swarm intelligence in which it does not need any central control on swarms, they co act using self organization. Ant colony optimization is invented from behavior of ants and their way of communication through chemical pheromone. Every ant lays pheromone for communication and other ants follow path marked with chemical. This is used while food searching.

Bee’s algorithm invented from the behavior of bees. Bees use Waggle dance for collecting food as pheromone. Bee does waggle dance to guide other bees to take to location of food. Cooku search is one of the nature inspired algorithms. It is based on the reproduction strategy of the Cooku bird. Cooku lays a single egg at one time and put this into nest of another bird. Differential evolution is extension of genetic algorithms. It is based on the vector based evolutionary algorithms. It does operation on each component.

TABLE I
Comparative Analysis of Secure Data Aggregation.

Paper Title	Mechanism	Advantages	Disadvantages	Simulation tool	Measurement
[1]	apply hash chain multi path process to obtain security of WSN and Developed network expanding model to reduce communication cost by multi path routing	1. Sensor nodes have low energy cost	1. Not for realistic application	Staple vs. INSENS	Data sink Ratio Vs False node ratio
[2]	Data aggregation design removes repetitive senor readings except doing encryption and keeps data secrecy and privacy during sending data	1. Security and privacy is given. 2. Avoids many attacks.	1. original readings will be aggregated into a Single packet not removed.	NS2	Estimation time to brute force with different key lengths
[3]	Data aggregation is obtained using without releasing private sensor readings and without announcing private senor head.	1.Protection to Data Privacy 2.More Accurate	1. suites to applications that have relative loose requirements of privacy-preservation	NS2	Percentage of left energy Vs Time
[4]	1. It includes a privacy management policy with an original aggregation 2. periodically every node evaluates some data for aggregation and use linear operations for encryptions	1. Data integrity is checked 2. End to end secure data aggregation is given	1.Geographical location should be known to node 2. Requires buffer.	Computer simulation	1.Buffer Vs time 2.Overall transmitted message Vs parameter set
[5]	Two adaption mechanisms are used 1. Energy aware selection strategy is identified and evaluated. 2. Service user and provider automatically adjust changes in topology.	1. Framework to address limitations of resources.	1. To avoid overhead passive adaptation is used. 2.Energy cost expressions are predictions	1.TinyOS 2.TelosB 3.Imote 2	1. Energy footprint Vs Count 2.Energy cost compared with Duty cycle and sensing period.
[6]	1. Every report has signed with private keys for not altering the message. 2. Parents and sibling send the aggregation result to child for verification.	Identifies malicious node. Base station does not receive tampered data due to use of pair wise keys are exchanged between parent and child.	Possibility of sending a false data by child to parent node. Chances of grand parent is malicious as it send data to children for verification	NS2	Network size overhead, average overhead per node, Network size energy consumption
[7]	1. Cluster are formed of heterogeneous node. 2. CH decides region for intra and inter cluster. 3. CH aggregates packets of fixed size 4. Cluster aggregation and sink communication is done by CH	1.Energy is saved by reducing no. of transmissions of from node to sink 2. Packet aggregation gives good bandwidth use.	1. This algorithm is not for mobile nodes.	NS2	Comparison of throughput , Average energy consumption, Packet delivery ratio, Residual energy, Lifetime with packet generation rate
[8]	To form cluster formation, Divide and conquer method 1. Formation of cluster. 2. Selection Secure node. 3. Data aggregation.	1. Reduction in communication overhead. 2. Secure channel is provided	1 Data aggregation is not delay tolerant.	NS-2	Energy level comparisons of MH-EESDA with existing data aggregation protocols.

IV.IDENTIFIED CHALLENGES

[1] Security is provided using hash chain multipath mechanism but cannot implement on realistic applications.
[2] Provides security and privacy during transferring and eliminates redundant reading. It is resilient to various attacks during data transfer. These are aggregated into single packet without removal.[3] more suitable for application require less security. [4] Nodes necessary to know geographical locations and transmission buffer. [5] Passive adaptation is provided. [6] Parent sibling relation is used and verification is done by grandparent, more energy consumed. [7]Aggregation is performed using cluster formation. Nodes considered are stationary. [8]Data aggregation performed is not delay tolerant.

V. CONCLUSION

This paper shows the Multiobjective and Metaheuristic approach for secure data aggregation. First it gives the idea of secured data aggregation. Then it gives the notion of Multiobjective and Metaheuristic approach. Data aggregation bind collected data from various node and reduces retransmission of similar data multiple times. Security is important for data, as data is sensed from rare and hostile area. This data is sensitive and prone to attack. Security can be provided using multiple encryption techniques. Multiobjective optimization provides the solution when more than two objects are functioning .It is useful to optimize more than two objects simultaneously. It gives the solution using higher level of information. In wireless sensor networks there are multiple objects to maximize and minimize. Due to limited resources like energy, memory etc., and these resources should be used properly. Multiobjective optimization helps save energy and memory by reducing data transmission. Metaheuristic approach gives solution for the fallible and underdone data. It gives optimal solution from large set of workable solutions. It is able to sample the large sample set, minimizes the latency and establish the security therefore it is useful in wireless sensor network.

REFERENCES

[1] Nike Gui, Ruichuan Chen, Zhuhua Cai, Jianbin Hu, Zhong Chen, “A Secure routing and Aggregation protocol with low energy cost for Sensor nodes”, In *IEEE International symposium on information Engineering and electronic commerce*, Ternopil, Ukraine (pp. 79-84),2009.

[2] Shih-I Huang Æ Shiuhpyng Shieh Æ J. D. Tygar, “Secure encrypted Data aggregation for WSN”, *Journal of Wireless Networks*, 16(4), 915–927, 2009.

[3] Hongjuan Li,Kai Lin,Kequi Li,“Energy- efficient and high accuracy Secure data aggregationin Wireless Sensor Networks” ,*Journal of Computer Communications*, 34(4), 591–597,2009

[4] Sabrina Sicari, Luigi Alfredo Griecob, Gennaro Boggiab, Alberto Coen-Porisinia, “DyDAP: A dynamic data aggregation scheme for privacy aware wireless sensor Networks”, *Journal of Systems and Software*, 85(1), 152–166,2012

[5] Chien-liang Fok, Gruia-Catalin Roman, Chenyang Lu, “Adaptive service provisioning for enhanced energy efficiency and flexibility in wireless sensor networks”, *Journal of Science of Computer Programming*,78(2), 195–217,2013.

[6] Hongjuan, L., Keqiu, L. ., Wenyu, Q., & Ivan, S “Secure and energy-efficient data aggregation with malicious aggregator identification in wireless sensor network”, *Future Generation Computer Systems*, 37, 108–116, 2014.

[7] Dnyaneshwar Mantri, Neeli Rashmi Prasad, Ramjee Prasad, “ BECPA: Bandwidth Efficient Cluster Based Packet Aggregation”, *Wireless Personal Communications*, 76(3), 335–349, 2014

[8] M. Bala Krishna · M. N. Doja, “Multi-Objective Meta-Heuristic Approach for Energy- Efficient Secure Data Aggregationin Wireless Sensor Network”,.In *Springer Science+Business Media New York*,USA,2014

[9] Gonzalez, T.F, “Handbook of approximate algorithms and metaheuristics”, *Computer and information science series. 6000BrokenSoundParkwayNW*, Suite300, BocaRaton, FL, USA: Chapman & Hall/CRC, Taylor & Francis Group, 2007.

[10] Hani Alzaid, Ernest Foo, Juan Gonzalez Nieto, “Secure Data aggregation in Wireless Sensor Network: a survey”, *Australian Information Security conference*, vol 81, 2008.

[11] Muhammad Iqbal, Muhammad Naeem, Alagan Anpalagan, Ashfaq Ahmed, Muhammad Azam, “Wirelss sensor Network Optimization: Multi-Objective Paradigm”, Senosors, 17572-17620, 2015.

[12] Nandini s. Patil, Prof.P.R.Patil, “Data aggregation in Wireless Sensor Network”, *IEEE International Conference of Computational Intelligence &computing Research”*, 2010.

[13] Ankit Tripathi,Sanjeev Gupta,Bharti Chorasiya, “Survey on Data Aggregation Techniques for Wireless Sensor Network” , *International Journal of Advanced research in computer and communication engineering* ,vol 3,Issue 7,2014.

[14] Anindita Ray, Debashis De, “Data Aggregation Techniques in Wireless Sensor Network: A Survey”, *Journal of Engineering Innovation and Research,* vol 1, 2277-5668, 2012.

[15] Zesong Fei, Bin Li,Shaoshi Yang, Chengwen Xing,Hongbin Chen, Lajos Hanzo, “A Survey of Multiobjective Optimization in Wireless Sensor Networks : Metrics, Algorithms and Open Problems” ,*Accepted to appear on IEEE Communication Surveys& Tutorials,*2016.

[16] Michel, G., &Jean-Yves, P. (2010).Handbook of metaheuristics. In Hillier, F.*series in operations research and management science* (2nd ed., Vol. 146), Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA.

[17] Sankardas Roy, Mauro Conti, Sanjeev Setia, Sushil Jajodia, “Secure Data aggregation in Wireless Sensor Networks” ,*IEEE Transactions on Information Forensics and Security,* Vol 7,No 3,2012.