RESEARCH ARTICLE OPEN ACCESS

Watermarking and Multiobjective Based Secure Data Aggregation for Wireless Sensor Network

[1]Jaybhaye Chaitali, [2]Dr. P. M. Pawar
[1, 2] Department of Information Technology, Smt.Kashibai Navale College of Engineering,Pune

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

Abstract:

Wireless Sensor network is formed of sensor nodes, placed in defy and hostile environment. Many important fields are dependent on WSN for data gathering from human unreachable areas. Sensor nodes have restricted amount of resources like memory, battery. Data aggregation helps to reduce energy usage by minimizing data transmission and removing data redundancy. Research proposed Multiobjective based optimization for secure data aggregation. Multiobjective optimization is used for cluster level aggregation. The research proposed watermarking based mechanism for securing data Watermarking provides data confidentiality and data integrity with less computational cost. Results of the SDAMMO are compared with the MH-EESDA system. Results show that the SDAMMO system is more efficient than the MH-EESDA.

Keywords— **Wireless Sensor Network, Data Aggregation, Secure Data Aggregation, Multiobjective Optimization, Watermarking.**

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

I. INTRODUCTION

Wireless Sensor network is always been choice of multiple departments to collect data from scarce and challenging region. WSN is used by health care monitoring, weather forecasting, military, habitat monitoring etc. WSN is a group of sensor nodes, transmits data using wireless medium. Sensor node requires energy to transmit data to base station. A sensor node is a small device with resources like memory, energy in restricted amount, ability to sense and process data. Life of the sensor node is dependent on the resources. As the battery of the node dies, sensor node dies. Sensor node transmits data to base station as it has insufficient memory. Continuous transmission of sensed data requires lot of energy. To reduce the rate of data transmission and energy usage data aggregation is the basic method. Sensor nodes are captures information from hostile fields and also used in military surveillance, sensed information is prone to attack. To provide security to data aggregation is important concern. Making a aggregated data secure is secure data aggregation.

Data aggregation is the process of combining data by removing redundant data. Data aggregation reduces transfer of multiple and same data. As the data transfer is minimized energy is saved and life of node is increased. Data transfer takes more energy of sensor node than other tasks.

To reduce the data transmission rate and save the energy data aggregation is used. Assembling of the data coming from the various nodes is the data aggregation. Data aggregation removes redundancy and helps to increase the lifetime of the sensor node. To provide a security digital watermarking is used. Processing of inclusion of some information into the digital files like audio, video etc. to get the ownership of the data is the digital watermarking. Every time new watermark is created for the sensed data so there is no need to storage and computation power required for the watermark is less. Providing a security to the aggregated data is secured data aggregation.

More than two objects are optimized at the same time is called as Multiobjective optimization. Secure data aggregation helps to increase the lifetime of the network by reducing data transfer rate by providing security. Two objectives are achieved at the same time. Watermarking reduces the use of energy and memory. Watermark is generated for each sensed data packet and inserted in the data packet to be sent [3]. Watermark is generated using data, Media access control address (MAC address) of node, secure hash algorithm-1 (SHA-1) cryptographic algorithm. MAC address of the node is combined using XOR operation and SHA-1 algorithm is applied on the result to get 160 bit digest as a watermark. Cluster head performs verification of watermark and generates a watermark using each node's MAC address. To perform cluster level aggregation local search

method is used. Wireless sensor network localization (WSN localization) simulator is used for the simulation purpose.

This paper is organized as follows Section II contains related work to secure data aggregation, watermarking in wireless sensor network. Section III contains System model, provides design and flow of the proposed system. Section IV contains assumptions made while designing the system. Section V gives the mathematical modeling of the system. Section VI provides the Algorithms developed for system. Section VII gives Simulation results of proposed system.

II. REALATED WORK

In [4] a privacy management policy is proposed and an original aggregation algorithm deals with end-to-end data encryption is provided. Transmission queue is evaluated by each node periodically to aggregate the data to aggregate and this way controls the level of transmission queue. The aggregation process uses linear operations to merge the encryption based spatial correlated data and permits the sink node for estimation of the confidence level of the aggregated data. This provides anonymity management, data integrity check, data aggregation to reduce the network load, end-to-end secure data aggregation. In [5] proposed a cluster based packet aggregation which is bandwidth efficient. It decreases energy usage and increases bandwidth. Randomly distributed heterogeneous nodes are arranged into the number of clusters. Intra cluster and Inter cluster aggregation is also performed. Cluster head performs aggregation which is packet based. Random data is combined with variable packet generation rate. It uses compressible aggregation function for cluster head aggregation. It has less packet delivery ratio and throughput and better energy efficiency. In [6] secure encrypted data aggregation is proposed. Data redundancy is reduced by removing redundant readings without using encryption. Data secrecy and privacy is provided. Conventional aggregation functions are used when received readings are in plaintext. Readings are decrypted using key management if received readings are encrypted. This provides security and privacy, and duplicate instances of original readings will be aggregated into a single packet. It is resilient to known-plaintext attacks, chosen-plaintext attacks, cipher text-only attacks and man-in-the-middle attacks. Communication overhead is reduced and can be implemented in on-the-shelf sensor platforms. In [7] proposed a weighted sum based Multiobjective optimization. Performance of the metaheuristic algorithm is increased by the utility method and meta-heuristic search methods. This article proposes multi-objective. Divide and conquer algorithm is used for formation of secure clusters. These clusters perform secure data aggregation. The proposed method works in in three steps. In the first step, the clusters are formed, in the second step, the secure nodes are selected and in the third step, energy efficient data aggregation is done over the secure route paths of the network. Minimum degree of intrusions, threshold-based degree of intrusions and Maximum degree of intrusions is evaluated for node energy and rate of data aggregation. In [8] Adaptive Servilla, a middleware is proposed. It has to ability to coordinate the resources used by Wireless Sensor Network. This is achieved using binding strategies which automatically adapt the when the topology changes. Energy aware service binding decisions are energy efficient. In [9] tree structure is used to for sending data to the base station. Leaf node uses slicing and assembling technique to protect data while transferring. This uses additive function for aggregation purpose. This provides privacy protection for raw data and low overhead with better aggregation accuracy. In [10] proposed a digital watermarking based copyright protection mechanism. This helps to provide a security for data wireless sensor networks. In this the embedding capacity of data is expanded using method of manipulating both LSB and MSB bits .Here data related information is inserted into the data as a digital watermarking. To increase the speed of parsing look up table is used by base station. For providing copyright to user it generates two dimensional codes. In [11] fully distributed watermarking method is proposed. This method provides data integrity and data authenticity. The proposed method works on low watermark payload and computational complexity. It is energy efficient method or simulation purpose CupCarbon simulator is used. In [12] proposed a secure data aggregation using watermarking. This method helps to remove data redundancy and reduces data transmission and helps to increase life of sensor node. To generate a watermark MAC address of sensor node is combining with the sensed data packet. Then SHA-1 cryptographic algorithm is used to generate 160 bit digest as a watermark. In [13] proposed a data integrity protection strategy based on digital watermarking technologies. A one way hash function is used by source sensors to collect data from which watermark information is created. This generated watermark is embedded into redundant space of packets of certain bytes. At the base station side, an algorithm is designed for extraction of watermark information, watermarking algorithm is designed. Using this algorithm information is recalculated and compared with watermarking information to verify the integrity of the data during the transmission. This algorithm does not increase extra data storage space and remain data accuracy. According to the results, algorithm protects the integrity of the data and has more application values. In [14] proposed a reversible watermarking authentication. This authentication is based on the difference expansion of a generalized integer transform. A one way hash function is used by source sensors to a generate watermark information from the adjacent data. After generation, watermark is embedded into these data. Base station has one manager node. This manager node upon receiving data restores the original data and reliability is verified. This algorithm with low energy authenticates sensor data.

In [4] transmission buffer and geographical position is required to node .In [5] stationary nodes are used. Work is not applicable to mobile nodes. In [6] original readings and redundant readings are aggregated into two different packets.

Redundant readings are not removed. In[7] data aggregation done is not delay tolerant. In [8] energy cost equations used are only predictions. In [9] slicing and dicing technique is used. To get original data pieces of data have to gather. Original data is known to node itself only. In [10] watermark used is not highly robust. [11] Proposed work does not work on real wireless sensor network. In [12] to cluster formation requires more energy. Optimal equation is not used to form clusters. The is applicable to homogeneous and stationary nodes.

In [13] data lost is not detected once data is lost. If data field size equals to size of resolution of data then this method does not work. In [14] fragile watermarking is used. Watermarking should be strong such that attacker should not put any data.

III. Assumptions and System Model

A. Assumptions

The proposed method for secure data aggregation uses homogeneous stationary nodes. Source node itself generates watermark for data confidentiality and integrity. Network is divided into clusters. Cluster head is elected on the basis of maximum storage and energy. Cluster head does aggregation of data coming from member nodes of the cluster and sends data towards base station. Cluster head also generated watermark using each sensor nodes MAC address.

B. System Model

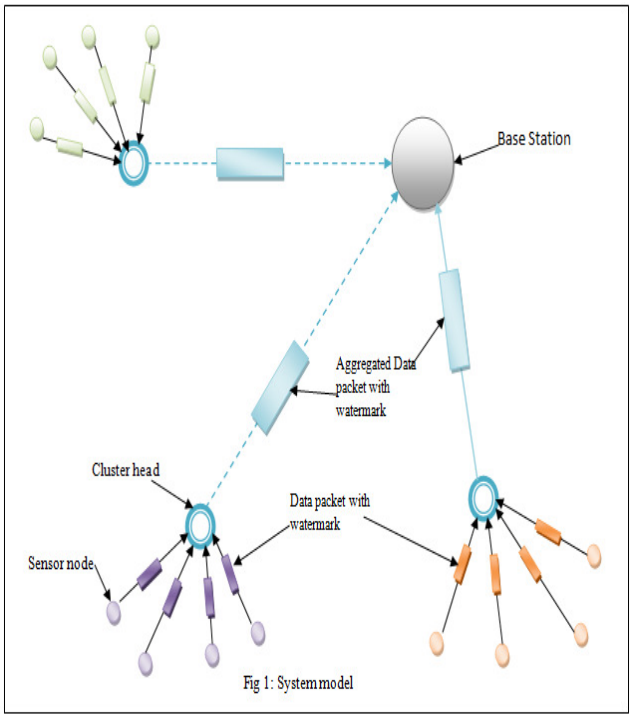

Fig 1: System model

Figure 1 shows the system model. Network is divided into the cluster head using divide and conquers algorithm. System works in three steps. In first step clusters are formed, Watermark Generation data transmission with watermark is second step, Verification and generation of watermark by cluster head is the third step.

Step 1: Network is divided into the clusters using divide and conquer algorithm. Sensor node with high energy and maximum memory is selected as a cluster head. Cluster's size is maintained equally.

Step 2: Source node captures the data from the placed area and send to the cluster head. Source node generates the watermark for captured data packet. To generate watermark, data is combined with MAC address by XOR function. SHA-1 algorithm is applied on the result to generate 160 bit digest. 160 bit digest is put into the data packet with data as a watermark and transmitted to the cluster head.

Step 3: Verification and Generation of watermark.

Cluster head receives data from the sensor nodes and disaggregates the received data. After receiving data, cluster head generates watermark from received data and compares this watermark with received watermark with data. If the watermarks are identical then data is not tampered. Cluster head generates the watermark for aggregated data. To generate a watermark MAC address of each node is combined with aggregated data using XOR operation. SHA-1 is applied upon the generated result to get the 160 bit digest as a watermark.

IV. Mathematical Model

Notations	Descriptions
N	Number of nodes.
IE	Initial Energy of each node.
Tee	Total Initial Energy
TCe	Total Energy Required for Communication
RE	Remaining Energy
AE	Average Energy

Consider the N as the number of nodes having initial energy IE. Total initial energy TI_E is

$$TI_E = IE \times N$$

Energy required for the communication is TC_E. Therefore Remaining energy RE will be.

$$RE = TI_E - TC_E$$

Average Energy AE required for the communication is

$$AE = RM \div N$$

V. ALGORITHMS

Clusters are formed using divide and conquer algorithm. Cluster head is selected on the basis of high energy and maximum memory.

Source node generates the watermark for every data packet it captures. To generate a watermark it requires MAC address of the node, data packet and cryptographic algorithms SHA-1. To generate a watermark data is combined with MAC address of the node.SHA-1 algorithms is applied on the generated result to get 160 bit digest as a watermark. This watermark is added into the data packet of particular size to be sent.

Algorithm 1. Generation of watermark by the source node

Input: Sensed data, MAC address of the node
Output: 160 bit digest as a watermark
Begin
D=Sensed Data
$M_{sn}=MAC_{sn}$
WM=Generated Watermark
X= (M) XOR (D)
WM=SHA-1(X)
Data Packet=WM+D
end

Watermark is generated by the cluster for cluster level aggregation using second algorithm. To generate watermark cluster head requires the MAC addresses of all the clusters, Aggregation of collected data from every sensor node.MAC addresses all the nodes and cluster head is combined using XOR operation. SHA-1 algorithm is applied on the generated result to get 160 bit digest as a watermark. Generated watermark is added into the data packet of fixed size along with the aggregated data. Cluster head sends this packet to base station.

Algorithm 2. Generation of watermark by the cluster head.

Input: Aggregated Data, MAC addresses of all the nodes and cluster head
Output: 160 bit digest as a watermark
Begin
AD=Aggregated Data
$M=MAC_{ch}$
n= No. of sensor nodes in the cluster
WM=Generated Watermark

For i=1 to n-1 **do**
M=(M) XOR (MAC_i)
End

X= (X) XOR (AD)
WM=SHA-1(X)
Data Packet=WM+AD
end

Sensor node sends data to cluster head. Cluster head verifies watermark using watermark verification algorithm. To generate watermark it requires data ,watermark of sensor node, MAC address of the node. Cluster head generates the watermark and compares with the watermark of sensor node. If the extracted watermark is equals to generated watermark then data is not tampered, else data is tampered.

Algorithm 3. Watermark Verification Algorithm

Input: Data, MAC address of node, Watermark generated by node
Output: Data is tampered or not.
Begin
EW= Extracted watermark.
GW= Generated Watermark

if (EW=GW) **then**
Accept Data
else
Reject Data

end if
end

VI. SIMULATION PARAMETERS AND RESULT

A. *Siulation Parameters*

Energy consumption , Throughput, Delay parameters are considered for the simulation. To perform he simulation WSN localization simulator tool is used. Initially 1000J energy is provided to the every node. AODV is used as a routing protocol. Channel type considered is wireless channel. Every clusters consists of 12-13 sensor nodes.

B. *Simulation Result*

Simulation results are taken for the energy, throughput and delay.

1) Energy Results.

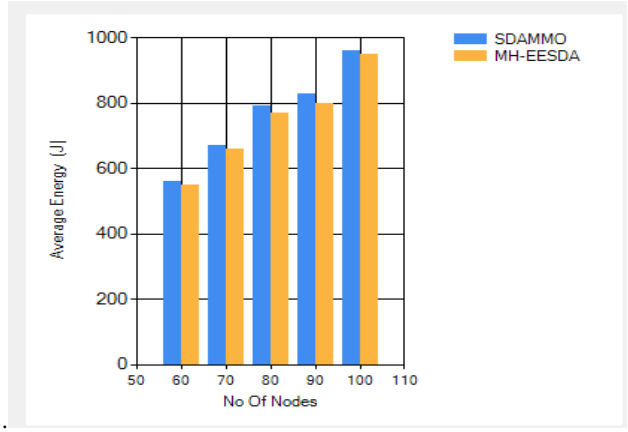

Fig 2: Two cluster Energy Result

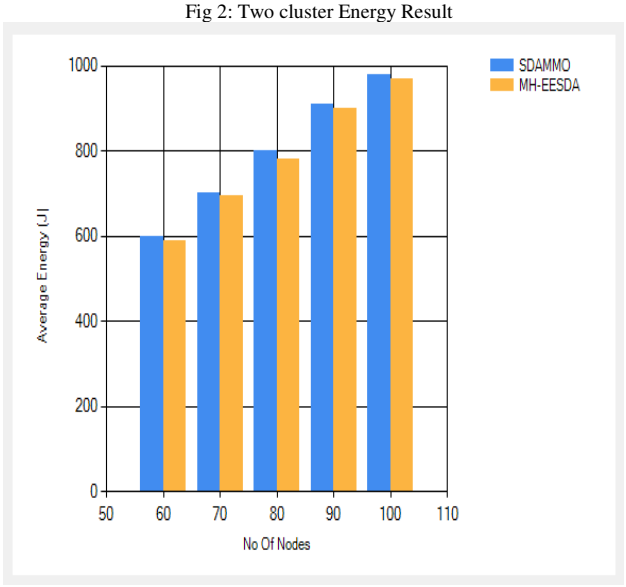

Fig 3: Four cluster Energy Result

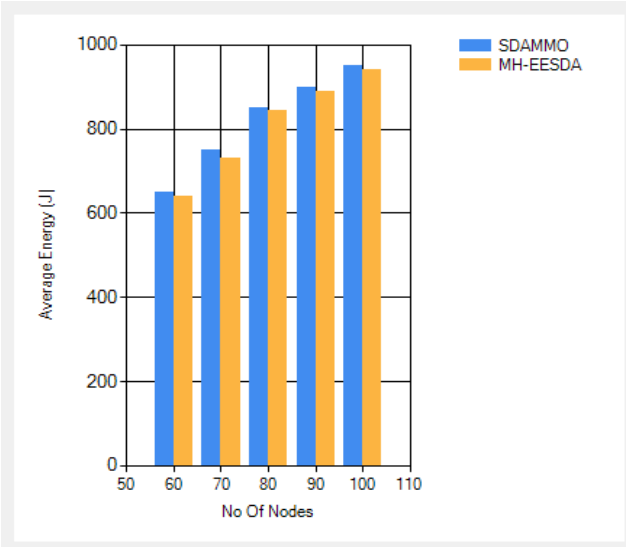

Fig 4: Six cluster Energy Result

Figure 2,3,4 shows the comparative results of average residual energy between the proposed system i.e. SDAMMO and Existing system. The results are taken for 60, 70, 80,90,100 by dividing into two , Four and six clusters. The blue color represents the proposed system and yellow color represents the existing system. The results of the SDAMMO are compared with the MH-EESDA system. The energy required by the proposed system for the communication and to generate the watermark is less than the existing system is less. Existing system provides security using encryption technique, requires more computation energy than generation of watermark. Results shows that residual energy of the MH-EESDA system is less than the SDAMMO. The proposed system is more efficient than the existing system.

2) Throughput

Figure 5,6,7 shows the comparative results of the SDAMMO i.e. proposed system and MH-EESDA i.e. existing system for throughput. Throughput for the 60 to 100 nodes is taken by diving the 60 to 100 nodes into the two, four and six clusters. Results show that the SDAMMO has more throughput than the existing system.

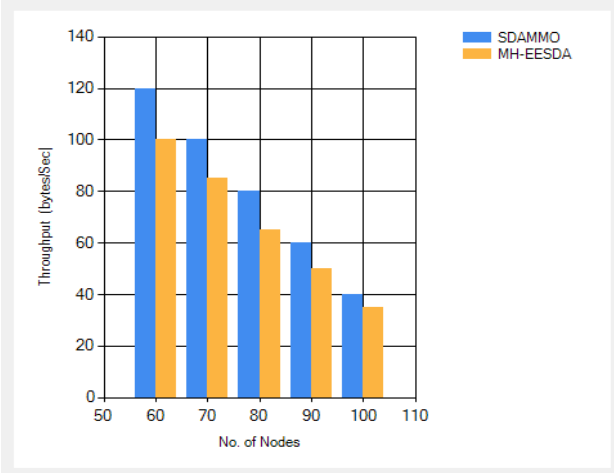

Fig 5: Two cluster Throughput

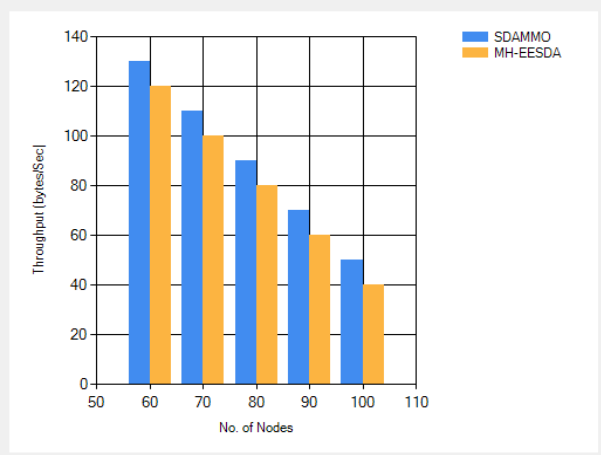

Fig 6: Four cluster Throughput

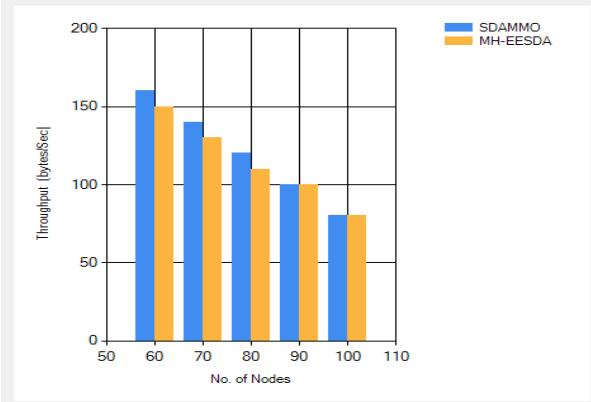

Fig 7: Six cluster Throughput

3) Delay.

Figure 8,9 and 10 shows the comparative results of the SDAMMO i.e. proposed system and MH-EESDA i.e. existing system for delay. Delay time required for the 60 to 100 nodes is taken by diving the 60 to 100 nodes into the two, four and six clusters. Results show that the SDAMMO has less delay than the existing system

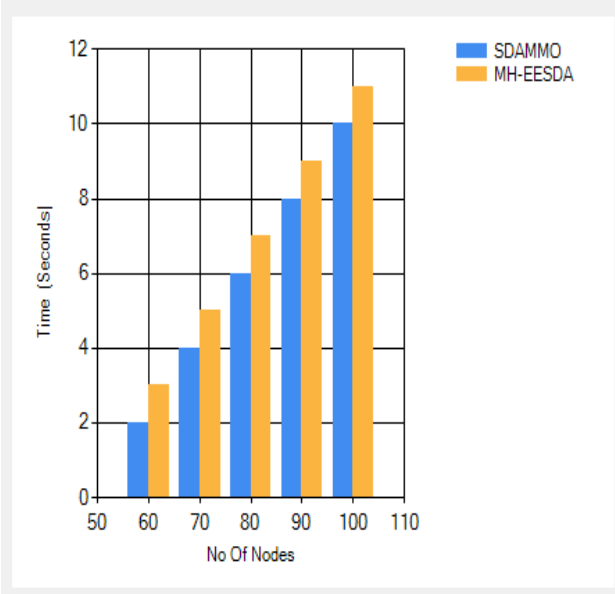

Fig 8: Two cluster Delay

.

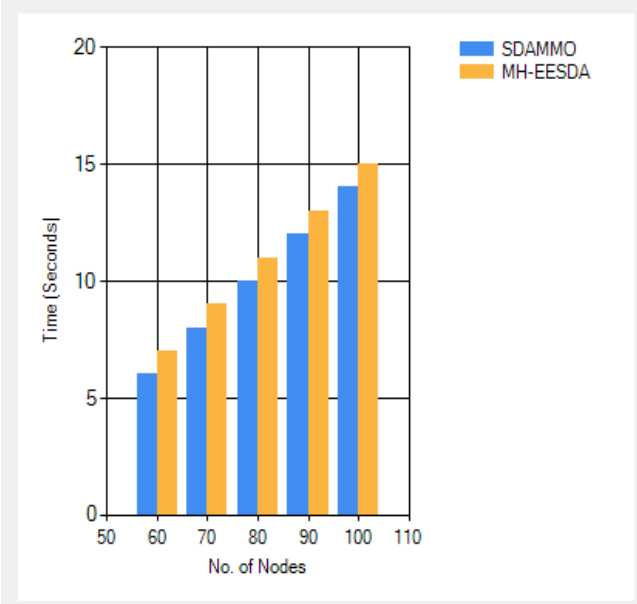

Fig 9: Four cluster Delay

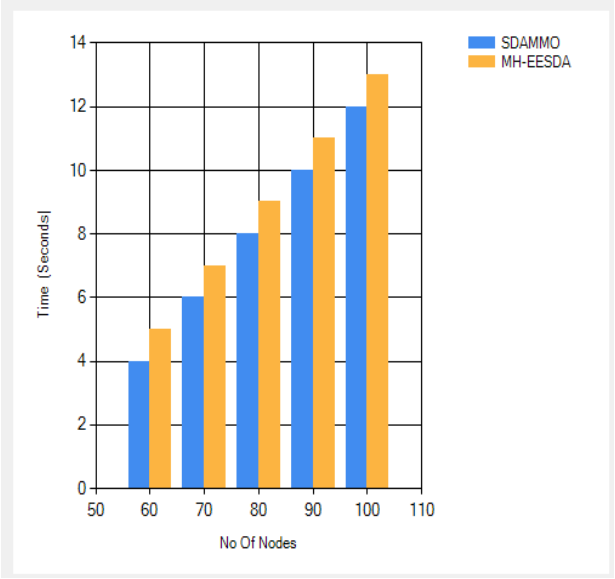

Fig 8: Six cluster Delay

VII. Conclusion

Proposed work uses watermarking technique to provide security to aggregated data in homogeneous sensor network. Watermarking used is of fragile type. This is provided using fixed byte of packet for data transmission. Proposed work also uses Multiobjective optimization and metaheuristic approach for secure data aggregation. Multiobjective optimization optimizes the multiple objectives at one time. Securities, lifetime of sensor, memory and energy efficiency are the objectives considered for optimization. Data aggregation helps to increase lifetime by reducing redundancy and data transmission. Watermark requires less computation energy and memory than cryptographic keys. It is expected that proposed work will requires less energy to generate clusters and watermark. This work can be further extended to work using mobile nodes and heterogeneous node. [Discuss something about the result].

References

[1] Suat Ozdemir, Yang Xiao. "Secure data aggregation in wireless sensor networks: A comprehensive overview". In *journal of Computer Networks 53 2022–2037,2009.*

[2] Michel, G., &Jean-Yves, P..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.2010.

[3] Gaurav Chawla, Ravi Siani, Rajkumar Yadav, "Classification of Watermarking Based upon Various Parameters", *International Journal of Computer Applications & Information Technology*, *Vol.I, Issue II* ,pp – 202-208, September 2012.

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

[6] 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.

[7] M. Bala Krishna · M. N. Doja, "Multi-Objective Meta-Heuristic Approach for Energy- Efficient Secure Data Aggregationin Wireless Sensor Network",. *Wireless Personal Communication Volume 81,Issue -1, pp 1-16, 2015.*

[8] 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.

[9] Hongjuan Li,Kai Lin,Kequi Li,"Energy- efficient and high accuracy Secure data aggregationin Wireless Sensor Networks" ,*Journal of Computer Communication*s, 34(4), 591–597,2009.

[10] Baowei Wang, Jianwei Su, Youdong Zhang, Biqiang Wang, Jian Shen, Qun Ding,Xingming Sun, "A Copyright Protection Method for Wireless Sensor Networks Based on Digital Watermarking", *International Journal of Hybrid Information Technology Vol.8, No.6 pp.257-268* .2015.

[11] Farid Lalem, Muath Alshaikh, Ahc`ene Bounceur, Reinhardt Euler, Lamri Laouamer,2 Laurent Nana, Anca Pascu1,"Data Authenticity and Integrity in Wireless Sensor Networks Based on a Watermarking Approach" ,unpublished.

[12] Djallel Eddine Boubiche1,Sabrina Boubiche1,Homero Toral-Cruz,Al-Sakib Khan Pathan,Azzedine Bilami ,Samir Athmani ,"SDAW: secure data aggregation watermarking-based scheme in homogeneous WSNs", *Telecommunication Systems Volume 62,Issue 2,pp 277-288,2016*

[13] Sun, X., Su, J., Wang, B., & Liu, Q., "Digital watermarking method for data integrity protection in wireless sensor networks". *International Journal of Security and its Applications 7(4) 407– 416.* 2013.

[14] Qun Ding, Baowei Wang, Xingming Sun, Jinwei Wang,Jian Shen, "A Reversible Watermarking Scheme Based on Difference Expansion for Wireless Sensor Networks", *International Journal of Grid Distribution Computing Vol.8, No.2 , pp.143-154* .2015.