A Comparative Analysis of Data Aggregation Based Routing Protocols Hierarchies and Comparing Secure Data Aggregation Protocols in Wireless Sensor Networks

Akila V[1], *Sheela* T[2], *Vempaty* Prashanthi[1], *Kiran* Kumar P[3], *P N S Sowmya* Bharadwaj[1]

[1]Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad
[2]Sri Sai Ram Engineering College, Chennai
[3]Vignan Institute of Technology and Science, Hyderabad

Abstract: Sensor nodes deployed in Wireless Sensor Networks (WSNs) are battery powered. It operates at low voltage. The main challenge in WSNs is to reduce energy consumption so that it increases the lifetime of the sensor nodes. The goal is achieved by using data aggregation technique and also by selecting the appropriate routing strategy. Sensor nodes are commonly arranged in unmanned location; the security of data plays another important role. In order to provide reliable data to the destination, there are various approaches designed for security. In this paper, we analyzed the various existing approaches for data routing, data aggregation and data security.

Keywords: Wireless Sensor Networks, Data Aggregation, Data Security, Routing Protocols

1. INTRODUCTION

Wireless Sensor Networks (WSNs) consist of limited power, worthwhile, multifunctional sensors deployed in dense environment. The physical infrastructure of the environment is not pre-configured and these sensors are deployed randomly in unsafe environment. The major goal of WSNs is to provide high Quality of Service (QoS) at very low cost with safe transformation of information. The variety of circumstance the WSNs can be applied vary from critical application such as military, health care to non-critical application such as habitat monitoring. Power, computational capacities and memory are very limited in these sensors. Sensing hardware, transceiver, processor is embedded in sensors. The sink node receives data that are collected from the various sensors. The energy consumption depends on the network topology.

2. DATA AGGREGATION

In the literature, various computations and different changes of the pre-arranged associations are used for picture assessment, course of action, and division. Different procedures have been taken a stab at other clinical data puts together, both on X-beam pictures of brain tumors and tumors from various human body segments. These papers were not considered further, as the accentuation was on the documents using a comparable X-beam picture database. The decision is to use all of the open planes could increase the database. As this could all-around impact the course of action yield by overfitting, pre-getting ready is required preceding dealing with the photos into the neural association.

Regardless, one of the known advantages of Convolutional Neural Network (CNN) associations is that the pre-planning and the part planning should not be performed. Taking everything into account, the best result in the composing using the separated picture parts as information sources are presented by Tripathi and Sack, with 94.64% precision. For commitment to the classifiers, they use features eliminated from the image's divided frontal cortex. In addition, they attempted their strategy using a 5-wrinkle cross-endorsement technique.

A relationship with the same state-of-the-art methods shows that our association got better results. The best outcome for 10-overlay cross-endorsement was refined for the record-wise procedure additionally, for the extended dataset, and the exactness was 96.56%. Taking everything into account, in the composition, no paper shows attempted hypothesis, through the subject-wise k-overlay method, for this informational image index. For the subject-wise procedure, we gained a precision of 88.48% for the expanded dataset. The typical test execution was under 15ms for each image. These results show that our association has a good theory capacity and incredible execution speed. It might be used as a practical decision help instrument for radiologists in clinical

3. DATA AGGREGATION HIERARCHIES

The various hierarchies used in data aggregation to transmit data are Tree, Cluster, Multipath and Hybrid based approaches. In Tree-based approaches, the routing tree is constructed based on the level. The level is calculated based on shortest distance to the sink node, hop count from the node, and also the residual energy level of the sensor nodes. In Cluster-based approaches, special nodes are elected as cluster head which is used to combine the data and forward the data to the sink node. In multi-path routing, the data is flowing along the different path to reach the destination. Data aggregation is performed along with the different path in the network. The Hybrid approaches combine the asset of tree-based approach and multi-path approach.

3.1. Tree-based Approaches

The Tree-based approaches are formed based upon the hierarchical organization of the nodes in the network and connected to Base Station (BS) as shown in Fig.1. The spanning tree routed at the sink is constructed first and the tree is the rooted at the sink. The levels are generated from the root node to the leaf's node. The aggregation is performed only in the intermediate node and not in the leaf's node. The main asset of the tree-based approach is no error due to approximation. The communication error is the major disadvantages of tree topology. The following are the protocol related to tree-based approaches.

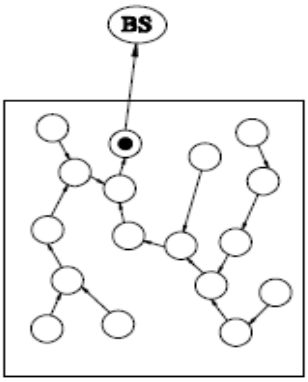

Fig.1 A Tree-based Routing Topology

Zeenat Rehena et al. [1] - They have proposed the power efficient method for transmitting the data called selective transmission. In selective transmission, the data are transmitted in the completely different direction to the sink node. There are three phases in this scheme. In distance discovery phase, the startup packets are transmitted to all the nodes and its hop count is calculated for each and every node. In negotiation phase, REQ message to the sending node and receive the shortest hop distance. In data transmission phase, the data are immediately transmitted to the requesting node. After increasing the number of nodes, the energy consumption of this concept comparatively reduced to already existing methods.

H. O. Tan et al. [2] - Sensor node has data to deliver to the base station by maximizing the lifetime of the sensor node. The location of the sensor nodes is not movable. The nodes aggregate the data before transmitting the data to destination. Life time is defined as the number of operation it takes for the failure of single node. Usually receiving the data from the other nodes is costly operation compare to the sending the data. The node which has less energy has given the high preference to send the data to the BS and also they are not allowed to receive the data from other node. The simulations are performed with the various techniques like LEACH, PEGASIS, PEDAP. It shows that the simulation of PEDAP reduces the number of death of the nodes compare to other LEACH and PEGASIS. The last node lifetime is very important in this approach. The main routing technique used is Prim's minimum spanning tree algorithm.

H.S. Annapurna et al. [3] - The various reasons for reliability failures are link failure, hardware failure, software error and the changes happen in the environment. The sensor node deployed in the Wireless Sensor Network ready to face these problems. The data sensed by the sensor nodes is partitioned into number of small data and these small data are transmitted into different path to reach the destination. Multiple copies of same data get transmitted to reach the destination to avoid the problem of communication failure. Each sensor node is preloaded with the key and also receive key for each and every session and perform XOR operation with the existing key and the received key. This process leads to the generation of new key each and every time. The divided data undergone mask designing technique and transmit the data and at the receiver side the same technique in reverse order to retrieve the original data. This concept reduces the communication link failure problem by introducing the concept of data redundancy. The data confidentiality is obtained using end to end encryption which is performed by using simple boolean AND or OR operation.

Wei Xing et al. [4] - They have proposed an energy efficient technique based on mobile sensor networks. The sensor node deployed consist of two phase first one is called routing tree construction and another one is called sensed data transmission phase. There are three phases in routing tree construction phase. They are starup, diffusion and schedule creation phase. This transmission of data in this scheme give guaranteed output compare to already existing system. The lifetime and energy saving of the sensor node is significantly increased compare to the already existing schemes. In the sensed data transmission phase, the middle node can combine and transmit the data to next hop. The average connectivity scheme of this network increased compare to other in the network.

Dimitris Tsitsipis et al. [5] -They have proposed the concept which uses the extra space in the data packet. The arrived data packets are merged with the already existing packet till the maximum size of the packet reached. The arrived packets check the existing packets empty space. If the space is not enough to fit for the arrived packet, then arrived data packet is split into small segments. The segmented packet fit in the existing packets. The merge timer is assigned for each and every packet. The packets get transmitted when the merge timer is completed or the packet get filled and the data packets get forwarded to the next node. The analysis is performed for different merge timer and the packet loss is calculated for different algorithms. The result shows that our algorithm depicts very less packet loss result compare to already existing algorithms.

3.2. Cluster-based Approaches

The Cluster-based approaches are formed based upon the hierarchical organization of the nodes in the network connected to Base Station (BS) as shown in Fig.2. Sensor nodes located in the network are partitioned into number of groups called clusters. Each cluster node has one cluster head. Always the cluster head node has more power to execute than other nodes. The pros and cons of this approaches are identical to tree-based approaches. The following are the protocols related to Cluster-based Approaches.

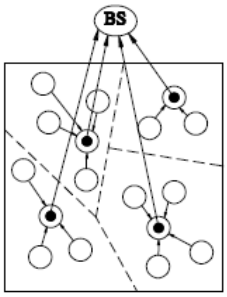

Fig.2 Cluster-based Routing Topology

K. Kishan Chand et al. [6] – They have proposed a technique to save the energy of the node and reduces the number of broadcasting packets to the nodes. It consists of three phases. The cluster head is selected for each and every group for certain period of time. After that new cluster heads are selected and old cluster heads retain as normal node. The cluster head changing process is repeated so that all the nodes in the group get chance to become the cluster head. The communication lifetime of all the nodes get increased compare to already existing works. Throughput also decreased with increase of time. The energy spent for each and every node decreased with time and the lifetime of the node decreases with time. This can be applicable to proactive networks. It can be enhanced for multi path data delivery.

Jan-Feng et al. [7] – They have proposed to achieve the network survival time. There are two phases. One is cluster building phase and another is the data transfer phase. The group of sensors connected to form clusters under cluster head. These cluster heads are connected to form super cluster head. The super cluster heads are combined and sent data to the sink node. There form a lot of layer is called multi-layer clustering. The redundant data is eliminated so that data transfer time considerably reduced. The current package and last package are compared and the compared value is higher than threshold. The value of the current package is stored in it. If it is less than threshold, then discards the current package.

Basavaraj S et al. [8] - Power Consumption and Reliability are the two important factors to be considers in this technique. Clusters are formed near the co-ordinator node and this node select cluster head depend on the remaining energy and distance from the co-ordinator node. The only operation performed in the cluster head is the aggregation and transmit the data to the co-ordinator node. This node measures the loss ratio by comparing with the threshold loss ratio of the packet. The loss ratio is calculated depend on the number of forwarding nodes cost value in the network. The cost value is calculated depend on the energy remains in the neighbouring node and the distance of the neighbouring node. The simulation result of EERDAT is compared with various technique.

Mohan B A et al. [9] - The selection of cluster head is depending on mobility, density, degree of energy. This cluster has the role to play as aggregator and as the encryption of the data. In first phase sensor nodes are distributed randomly and the calculation of degree energy, density and weight based on the application. With these details cluster head is calculated, the cluster head transmits the data to another cluster head or straight to the sink by selected energy balanced route. The secure transmission of the data happened by applying RSA encryption and decryption algorithm. This help to transmit secure data between the sensor node and the cluster head and also between the cluster head to another cluster head or base station. The simulation is performed for EECSSDA and compare with LEACH and LEACH-MF. The parameters like energy consumption, throughput and packet drop decreased compare to LEACH and LEACH-MF. The delivery ratio of EECSSDA is increased considerably compare to other methods like LEACH and LEACH – MF. The integrity and confidentiality of the data is obtained by using the secure RSA transmission.

A.F. Salami et al. [10] - They have proposed to enhance the energy consumption. The discovery packets passed to all the nodes from the sink node. Then the hop count is determined based on the discovery packets from the sink node. The random number generated between 0 and 1. High energy nodes are selected as aggregators. The greater and smaller energy nodes are selected based

on their energy level. The analysis is performed based on network lifetime when we increase the number of rounds the number of active node is more compare to already existing work. If we increase the number of nodes, then the total energy consumed is diminished significantly. If we increase the number of rounds, then the throughput has increased compare to already existing work. The data forwarding algorithm use their information to forward the data packets to sink. Multi-hop transmission is used in data forwarding algorithm.

Gaukhar Yestemirova et al. [24] - Multiple sinks avoid the problem of transferring the data using cordless network used by single sink. In present system, they used a single sink which became the problem because of which network loss, lack of communication sink crash could take place and it takes lot of time to recover that sink. They proposed a system using multiple sinks instead of single sink. The idea behind using multiple sinks is that if one sink goes down the data can be accessed using the other sinks. Also, if someone tries to attack on any of the sink and tries to modify the data, it can easily be identified. The main goal is to formulize the data accumulations problem, provide two algorithms and simulate algorithms and show effectiveness. They developed a tree which is a NP-Complete that connected various sinks namely virtual sinks and super virtual sinks connected by a backbone using two algorithms Minimum Spanning Tree and Shortest Path Tree to solve the problem which had two phases as building backbone and tree phase and scheduling phase. They concluded that both algorithms worked well in solving the problem. They minimized number of packets and small number of packet transmissions. It can be extended to include node-failure tolerant and conduct experiments on real testbeds.

3.3. Chain-based Approaches

Each node transmits its data towards its close neighbour [25]. The chain continues till it reaches the sink node. Aggregation takes place each and every node in the transmission. In case of cluster-based networks, there might be excessive energy wasted in transmitting towards cluster head connected to Base Station (BS) as shown in Fig.3. The following are the protocol related to Chain-based Approaches.

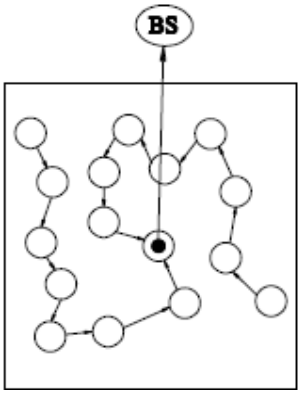

Fig.3 Chain-based Routing Topology

Jisoo Shin et al. [11] – They have proposed a chain based routing and the main concept of this routing is to save energy. There are two phases in it. One is the throwing schedule and another phase is chain establishment. The Sink determines the number of throwing to be assigned to any node. Throwing energy is common for all the nodes. It is the work of base station to calculate the cumulative forwarding energy and order them in the decreasing order of cumulative forwarding energy. Levels in the chain are calculated based on the forwarding energy. The node with very less energy is loaded in the level-1 nodes, the next number of exhausted nodes is in level-2 and the rest are in level 3 nodes. The chain is formed by using Kruskal's Minimum Spanning Tree. It takes up the shortest link from the node and it is added one by one till the new chain is formed.

Poonam Lohan et al. [12] – It is a chain-based routing algorithm. The sensor field is partitioned into small virtual square grid depend on their position information. The node with the greatest energy in each virtual square grid has the active state. This active state node sends the data directly to the sink. The unused nodes are put in the sleep state. The communication distance between the nodes decreased considerably due to chain-based concept. In vertical direction chain is formed between the nodes and allows longer lifetime of the nodes. It also reduces the efficiency of the energy. The performance is based on the percentage of dead nodes when increasing the number of rounds. It automatically reduces the percentage of dead nodes.

3.4. Structure-free Approaches

It spends some extra energy in building a structure. It is most suitable for dynamic environment. It performs data aggregation using the local information of that particular node. The periodic schemes are not applicable in that approaches.

Mohammad Hossein Yeganeh et al. [13] - They have proposed a scheme that involve both spatial and temporal aggregation of the data. Aggregation ratio is calculated and the node with high aggregation ratio

sends the data to the node which is spatially convergent. The calculation of fitness value groups similar type of data at a node. The next hop is selected from the nodes which has higher priority from the group are forwarded to that high priority node. If the priority is not assigned to that node, then the priority is selected based on real time routing policy. If there is no high priority packet exist in the forwarding path, then back pressure rerouting algorithm is assigned to that node. Due to that packet dropping concept is avoided. The remaining hop count value or slack time is the ratio of number of hops from the present node to the sink node to the next hop forwarding node. If the next hop is sink node, then it uses the entire slack time to reach the destination.

The various parameters based on increasing order from top to bottom are Structure (ST), Energy Consumption (EC), Communication Cost (CC), Delay (DY), Life Time (LT), Aggregation Function (AF), Reliability (RB), Scalability (SB), Throughput (TP), Survival Rate (SR), Delivery Ratio (DR), Drop (DO) and Fault Tolerance (FT). In structure (ST) row parameter indicates Tree(T), Cluster(C), Chain (CH) and Structure Free (SF). In AF, S indicates sum function.

Table 1: Comparison of Data Aggregation Routing Protocols based on its Hierarchies

	1	2	6	7	8	9	3	10	4	11	12	5	13
ST	T	T	C	C	C	C	T	C	T	CH	CH	T	SF
EC	Y	Y	Y	Y	Y	Y	-	Y	Y	Y	Y	Y	Y
CC	-	-	-	-	-	-	Y	-	-	-	-	-	-
DY	-	-	-	-	Y	Y	-	-	-	-	-	Y	-
LT	Y	Y	Y	Y	Y	Y	-	Y	Y	Y	Y	Y	-
AF	-	-	-	-	-	-	S	-	-	-	-	-	-
RB	-	-	-	-	-	Y	Y	Y	-	-	-	-	-
SB	-	-	-	-	-	-	-	Y	-	-	-	-	-
TP	-	-	Y	-	-	Y	-	Y	-	-	-	-	-
SR	-	-	Y	Y	-	-	-	-	-	-	-	-	-
DR	-	-	-	-	Y	Y	-	-	-	-	-	-	-
DO	-	-	-	-	Y	Y	-	-	-	-	-	Y	-
FT	-	-	-	-	-	Y	Y	-	-	-	-	-	-

In Table 1 shows the comparison of data aggregation routing protocols based on its hierarchies with number of parameters in first column and paper reference number in first row.'Y' indicates the presence of parameter for that reference paper number.'-' indicates that parameter not specified for that reference paper number.

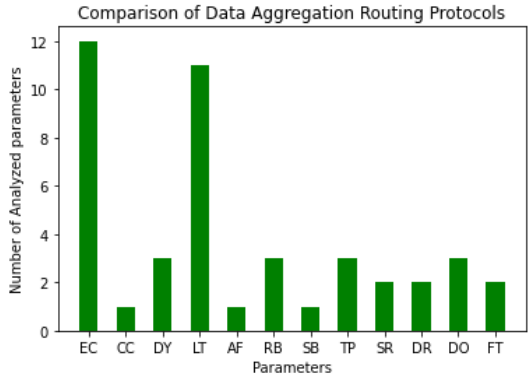

Fig.4 Comparison of Data Aggregation Routing Protocols based on its Hierarchies

Fig.4 shows the x-axis with number of parameters and y-axis with reference paper number. The Algorithm name present on the right-hand side shows the name of the algorithm needed for representation based on the increasing order of reference number from top to bottom. The various parameters based on increasing order from right to left are Energy Consumption (EC), Communication Cost (CC), Delay (DY), Life Time (LT), Reliability (RB), Scalability (SB), Throughput (TP), Survival Rate (SR), Delivery Ratio (DR), Drop (DO) and Fault Tolerance (FT). The number along with the parameters denotes the reference paper number.

4. COMPARISON OF SECURE DATA AGGREGATION

Privacy preserving data aggregation protocols is divided into various types based on the encryption technique. Data Confidentiality is an important factor that are related to protecting the personal information related to home, human body and sensed critical values [26]. The following are the secure data aggregation-based protocols.

Rabindra Bista et al. [14] - Data Confidentiality is mainly needed in the application used to record personal information such as heart beat and blood pressure. The sample contains five private seed data. It is used in round robin fashion. The sample data is split into two pieces and the two seeds in the sample are aggregated to maintain data confidentiality. Again, the same data is split into two pieces to maintain the integrity of the data. The sum aggregation function is used to aggregate the data and the accuracy is not preserved. The number of transmitted messages is less compared to already existing method. It uses the concept of duty cycling and number of operations performed is also reduced and it maintains the energy of the sensor node. In future, it can be extended to support various security features under collusive attacks.

Mohamed Ben Haj Frej et al. [15] - It provides the exact data to the same destination. It avoids sending the same message repeatedly. It reduces the battery life time and efficiency. Availability of the node is increased by using the multipath communication and also preserves the nodes life time. The location of the sensor node is traced as required. The three modules used in SDAM. In operation sub layer, it determines the MAC address to find out the proper nodes. In data aggregation module, it removes the repetition and reducing the size of the packet. In communication sub layer, it decides the number of packets needed to be aggregated and forwarded. Here two types of network are used first one is homogeneous and another one is heterogeneous network. Heterogeneous network concept mainly used since head node with more power and bandwidth. It detects that malicious node by using operation sub layer and also avoid data loss. The packet loss gets reduced and leads to retransmission of the packets. In cross layering overhead, the number of aggregations also increased and the time to aggregate get reduced. It gives

40% improvement in the cross layering overhead method.

Joyce Jose et al. [16] - It is based on Tree based data structure. Data Confidentiality can be achieved using homomorphic encryption algorithm. Message authentication using secret key and its corresponding identification pair of each node. It avoids duplication of packet and increase accuracy. It reduces number of transmissions and leads to the reduction in communication overhead. It reduces the number of decryptions, minimize the power and improve the energy efficiency. The reduction in the number of transmissions leads to the reduction in the number of collisions. The reduction of collisions leads to the packet loss reduction. It increases the accuracy of the packet. Leaf node divide the data into m pieces and m-1 pieces are transmitted. It reduces the count of transmission and increases the energy efficiency compare to already existing method.

Hongjuan Li et al. [17] It is based on Tree Structure. Accuracy is an important factor in data aggregation with security. The efficiency of the system is increased by using less bandwidth, less usage of energy and the data has to be protected with moderate overhead. The communication overhead is decreased by decreasing the number of message transmission. The accuracy of the data transmitted got raised with the raise in the time interval. The wait time interval for the data aggregation is increased means more data are aggregated. It leads to a smaller number of transmissions that reduces the chances of collisions. The chances of collisions reduced means than the messages are delivered within deadline. The data privacy is accomplished by using slicing and assembling technique. The data freshness is obtained by changing the encryption key for each and every session. In this technique less number of data transmission compare to already existing method.

Ekta Choudhari et al. [18] - It used cluster-based data structure. In that cluster head is the place where of the group of sensor transmit their sensed data. If that cluster head node got compromised, then false data can be injected in to the aggregated data. To avoid this problem cluster head, contain the iterative filtering algorithm to run frequently. We have calculated the skewed sample in order to calculate the robust variance compare to normal mean and variance for other methods. The cluster head nodes possess more energy which is needed to run the Iterative Filtering algorithm and to perform data aggregation. Three procedures are performed in the cluster head node. First is the estimation of noise level in the readings of the sensor. Second is the iterative filtering and the last is the eliminating colluders in the network. The test was conducted for biased and unbiased nodes. Root Mean Square Error is measured frequently and this value is minimized by using the filtering technique.

Djallel Eddine Boubiche et al. [19] - The basic security scheme used is encryption technique. The problem with this technique is that key generation in the base station and the generated key is transmitted which consume lot of energy. It requires large memory to store key. Light weight fragile watermarking technique without encryption is used to avoid the above problem. Watermark continuously generated for each and every packet and integrated the watermark with the data. Multi-hop transmission is used to transmit the data until it reaching the cluster head. The head perform aggregation of all the data and again apply light weight fragile watermark technique before transmitting to the cluster head. The communication overhead gets decreased compare to encryption-based protocols. When the data packet affected by interference problem then it is detected in the receiver node. It is called as false positive detections. It produced 17% more accurate result compare to already existing method.

P. Padmaja et al. [20] - The physical attacks rely upon the environment in which it is positioned. Message encryption and authentication are the two measures to avoid the interface attack. Software level attacks risk reduced using custom software development. The number of control packets transferred for verification in TESDA is small compare to SDAF. Overhead in transmission is high in SDAF than TESDA. It leads to the reduction of energy consumption, network life and increases the overhead.

Lifei Wei et al. [21] - The security is provided by using Identity-Based Multi-Signature (IBMS). It consists of two types of IBMS. They are IBMSQR -1 and IBMSQR -2. This is secure from message and identity attack. IBMS scheme consist of various phases and it has its own functions. Setup phase is the key generation system with both master keys like public key and secret key. Extra phase generates private key and distribute through secure channels. Sign phase is used to generate signature and the corresponding identity of the node. MSign phase is used to generate multi signature system. MVeriy is used to verify the multi signature system. Authenticating different sensors take place through MSign. The method based on cubic residence which is more applicable for marine applications related WSNs.

Ahmed Alghamdi et al. [22]- It provides secure data aggregation using elliptic curve and Hilbert curve-based method. Elliptic curve method is used for seed exchange between the sender and receiver and Hilbert curve method is used for data transmission between two nodes. The intruder cannot get the original value by using these methods. Three phases are proposed to achieve privacy preservation. In network construction phase, the HELLO messages are broadcasted to the nodes within the communication range from the sink node. The nodes with in the communication range of sink node set the sink node as parent node. Again, these nodes started broadcasting message to other nodes with in the communication range [28]. This process continues till all the nodes identify its parent. In data encryption phase, the seed data is exchanged between sender and receiver using elliptic key exchange method. Hilbert curve is used to transform the aggregated data into two-dimensional data. The key value contains the

direction and level of the node which is added to the two-dimensional data. In data aggregation phase, each leaf node sends it data to its parent node. Parent node analysis the direction of the child with its direction and the child node direction differs from parent direction then it changes the direction of child to parent node direction. It reduces the communication cost but still it has the overhead of computation cost. In our paper, data confidentiality is achieved by using encryption and decryption technique. In order to avoid the computation complexity, encryption and decryption are performed by using simple sum function like Homomorphic encryption [29]. It is based on end-to-end encryption technique. The aggregation function used in the parent node reduces the number of transmitted messages compare to existing approach. Automatically there is a less possibility of collisions. The communication cost and energy consumption are significantly reduced. New session key is generated for each and every session to preserve the data freshness. Data Integrity is achieved using Outlier Detection Algorithm in the sink node[30].

Table 2: Comparison of Secure Data Aggregation in Wireless Sensor Networks

	14	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**
ST	T	C	T	T	C	C	C	C	T	T
DC	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
AC	Y	Y	Y	Y	Y	Y	N	N	N	Y
CC	Y	Y	Y	Y	N	Y	Y	N	N	Y
EC	Y	Y	Y	Y	N	Y	Y	N	N	Y
LT	Y	Y	Y	Y	N	Y	Y	N	N	Y
DY	Y	Y	N	N	N	Y	N	N	N	Y
AF	S	S	S	S	R	-	-	-	E	S
DI	N	Y	N	N	N	N	N	N	N	Y
AU	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
DF	N	Y	Y	Y	N	N	N	N	N	Y
ID	N	O	N	N	N	N	N	N	N	OD

In Table 2 shows the comparison of secure data aggregation in WSNs with parameters in first column and paper reference number in first row [27]. 'Y' indicates the presence of parameter for that reference paper number. 'N' indicates the absence of parameter for that reference paper number.'-' indicates parameter not specified for that reference paper number. The various parameters based on increasing order from top to bottom are Structure (ST), Data Confidentiality (DC), Accuracy (AC), Communication Cost (CC), Energy Consumption (EC), Life Time (LT), Delay (DY), Data Integrity (DI), Authentication (AU), Data Freshness (DF), and Integrity Detection Method (ID).In structure (ST) row parameter indicates Tree(T) and Cluster(C).In AF, the papers used various aggregation function like sum(S),root mean square(R) and Elliptic curve(E).In Integrity detection method ,methods used are Operation Sublayer(O) and Outlier Detection Algorithm(OD)

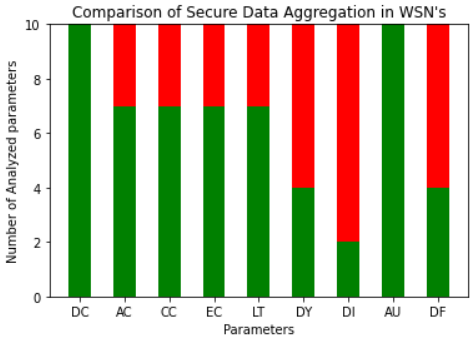

Fig.5 Comparison of Data Aggregation based on Security

In the above Fig.5 shows the x-axis with the number of parameters and y-axis with reference paper number. The Algorithm name present on the right hand side shows the name of the algorithm needed for representation based on the increasing order of reference number from top to bottom. The various parameters based on increasing order from right to left are Data Confidentiality (DC), Accuracy (AC), Communication Cost (CC), Energy Consumption (EC), Life Time (LT), Delay (DY), Data Integrity (DI), Authentication (AU), Data Freshness (DF), and Integrity Detection Method (ID).

5. CONCLUSION

The area of Wireless Sensor Networks is increasing day by day in all real time applications. The data aggregation based routing techniques help in decreasing energy consumption and also increases the life time of WSNs. In this paper, we surveyed the various routing strategies and its own set of advantages. We have compared the various existing routing techniques using various parameters like accuracy, delay, delivery ratio and life time. In the above comparison, EECSSDA [9] covers most of the parameters of the routing techniques. All the compared paper main consideration is to decrease energy consumption and increase lifetime. And also we compared various works related to secure data aggregation and its parameters like data confidentiality, data integrity, data authentication and data freshness. The comparison of secure data aggregation in WSNs is based on Tree and Cluster based topology. PSMD [23] is based on tree topology and SDAM [15] is based on cluster topology has covers all the parameters of secure data aggregation in WSNs.

REFERENCES

[1] Zeenat Rehena, Sarbani Roy, Nandini Mukherjee (IEEE Communication Systems and Networks, 2011)
[2] H. O. Tan and I. Korpeoglu, SIGMOD Record, 32, 4, 66-71(2003)
[3] H.S. Annapurna, M. Siddappa, (IEEE International Conference on Emerging Research in Electronics, 2015)

[4] Wei Xing, Kuntai Li, Yongchao Wang, Wen Zhou, (IEEE Wireless Communications and Signal Processing ,2011)
[5] Dimitris Tsitsipis, Sofia Maria Dima, Angeliki Kritikakou, Christos Panagiotou and Stavros Koubias (International Conference on Industrial Technology, Athens,2012)
[6] K. Kishan Chand, P Vijaya Bharati and B. Seetha Ramanjaneyulu (IEEE Advances in Engineering, Science and Management, 2012)
[7] Jan-Feng Yan, Yuan-Liu Liu, (IEEE Electronics, Communications and Control,2011)
[8] Basavaraj S. Mathapati, Siddarama R. Patil (IEEE International Conference on Computing Sciences,2012)
[9] Mohan B A, K R Dayananda, Saroja Devi H (IEEE International Conference on Green Engineering and Technologies,2017)
[10] A. F. Salami, F. Anwar, H. Bello-Salau, A. M. Aibinu, (IEEE Mechatronics, 2011)
[11] Jisoo Shin and Changjin Suh (IEEE Communications and Networks, 2011)
[12] Poonam Lohan and Rajni Chauhan (Students Conference on Electrical, Electronics and Computer Science, Bhopal, 2012)
[13] Mohammad Hossein Yeganeh, Hamed Yousefi, Naser Alinaghipour, Ali Movaghar, (IEEE 17th International Conference on Embedded and Real-Time Computing Systems and Applications,2011)
[14] Rabindra Bista, Myoung-Seon Song, and Jae-Woo Chang, (IEEE International Conference on Networked Embedded Systems for Enterprise Applications,2010)
[15] Mohamed Ben Haj Frej, Khaled Elleithy, (IEEE 14th International Conference on Machine Learning and Applications, 2015)
[16] Joyce Jose, M Prince and Josna Jose, (IEEE International Conference on Emerging Trends in Computing, Communication and Nanotechnology, 2013)
[17] Hongjuan Li, Kai Lin. Kequi Li, Computer Communication, 34, 591-597 (2011)
[18] Ekta Choudhari, Ketan.D. Bodhe, Snehal M. Mundada (IEEE International Conference on Innovative Mechanisms for Industrial Applications,2017)
[19] Djallel Eddine Boubiche, Sabrina Boubiche, Homero Toral-Cruz, Al-Sakib Khan Pathan, Azzedine Bilami, Samir Athmani, Telecommunication Systems, 62, 2, 277-288 (2016)
[20] P. Padmaja, Dr.G.V. Marutheswar (IEEE 7th International Advance Computing Conference ,2017)
[21] Lifei Wei, Lei Zhang, Dongmei Huang, Kai Zhang, (IEEE 14th International Conference on Networking, Sensing and Control, 2017)
[22] Ahmed Alghamdi, Mesfar Alshamrani, Abdullah Alqahtani, Sultan Safar A. Al Ghamdi, Rami Harrathi, (IEEE International Symposium on Networks, Computers and Communications, 2016)
[23] Pallavi Sawale, D.J Pete (IEEE 9th International Conference on Intelligent Systems and Control ,2015)
[24] Gaukhar Yestemirova, Sain Saginbekov, (IEEE 32nd International Conference on Advanced Information Networking and Applications, 2018)
[25] Prasanna Lakshmi K, Reddy CRK. (Int Conf. on Networking and Information Technology,2010)
[26]AkilaV.,Sriharshini,K.,Sravani,P.,Sravanthi,D.,Gopi ,R.,Sheela,T., International Journal of Online and Biomedical Engineering,17(1),120–128 (2021)
[27]Hussaini,S.M.,Krishna,G.,Gupta,A.K.,Singh,S.K.,Jo urnal of Manufactoring Processes,18,151-158(2015)
[28]Sateesh,N.,SampathRao,P.,Ravishankar,D.V.,Satya narayana,K. (Materials Today: Proceedings.,2015)
[29] Hussaini, S.M., Singh, S.K., Gupta, A.K., Journal of Materials Research and Technology.,3(1),17-24,2014.
[30] Dhanalaxmi B, Apparao Naidu, G., Anuradha.K (Procedia Computer Science.,2015)