International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)
Volume 3, Issue 12, December 2014

A Literature Survey on Security and Clustering in Wireless Sensor Networks

K. Gowtham, Mr. S. Dhanasekar

***Abstract*-- Wireless sensor networks are often used for monitoring sensitive data. Therefore security is a significant issue in WSNs. We point out the constraints, security requirements and attacks in WSNs. Clustering is the concept which increases the network scalability and decreases the energy consumption in WSNs. This article presents importance of clustering and clustering algorithm in WSNs. We first outline the basics of wireless sensor networks and general notion of cluster that is how the clusters are formed and its communication. Also we highlight merits, demerits and issues of clustering protocols in wireless sensor networks.**

***Index Terms*-- Wireless Sensor Networks, Security, Clustering Algorithms.**

I. INTRODUCTION

Wireless sensor network is a network which can be made up of autonomous sensors and these sensors are spatially distributed where they are used to sense environmental conditions within the network. These sensed data are passed to the main location through the network. Such networks are bidirectional and enables sensor activity control. Nowadays the WSNs are used in many applications like process management, healthcare monitoring, industrial monitoring, environmental and earth sensing, combat field surveillance and so on.

The WSN consists of several sensor nodes, where each sensor node is connected with another. Each sensor node includes several parts as in Fig. 1.

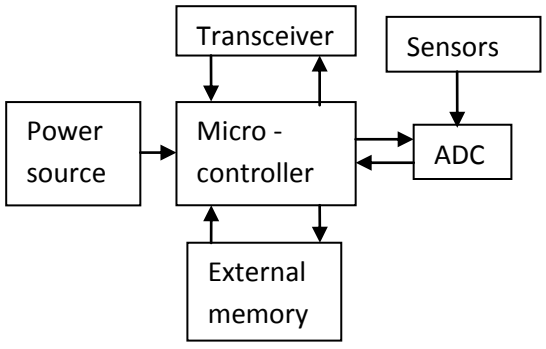

Fig. 1. Sensor node architecture

Sensors sense the data from the environment and those sensed data are converted into digital format by using analog to digital converter. A microcontroller is an electronic circuit, acts as an interface between power source and sensors. A transceiver is used to transmit and receive the data to and from the other nodes in the network and an external memory which is used to store the data. Sensor nodes in the wireless sensor network are grouped into one called cluster. A wireless sensor network consists of several clusters; a sensor node (leaf sensor node) joins a cluster based on receiving signal strength. Every cluster has a leader sensor node called cluster head (CH) sensor node. The CH sensor node has a highest capability than the leaf sensor node. Select a CH sensor node in a cluster either by the sensor nodes in that cluster or previously assigned by the network designer and the membership of that CH sensor node may be determined or variable. The CH sensor node aggregates the data from the leaf sensor nodes and these aggregated data are passed to a base station or a command center. Clustering technique has many advantages like network scalability support, maximize the life time of the network, minimize the size of routing table, conservation of communication bandwidth, stabilize the topology of network, and reduce the energy consumption.

A sensor node consists of five layers. They are physical, data link, network, transport and application layer. Physical layer tends to data encryption, signal deflection, modulation, frequency selection and generation. Data link layer tends to data multiplexing, data frame detection, medium access, error control, point to point and multipoint connectivity. Network layer tends to assign address and packet forwarding. Transport layer ensures reliable transmission of packets. Application layer is responsible for data requisition, data provision and interaction with end user.

Mostly, sensors are used to monitor sensitive data like enemy movement on the protected area. Therefore security is the major challenging task in wireless sensor network applications. The following constraints [8], [11] make security as a challenging

ISSN: 2278 – 909X All Rights Reserved © 2014 IJARECE

one. They are low computation capability, limited memory, limited energy resources, use of insecure channel, sensitivity to physical capture, frequent changes in topology of network, sensor nodes are thickly deployed, sensor nodes are liable to failures, sensor nodes are not having id globally and number of sensor nodes can be varied in size. In WSNs, secure communication is done by the use of keys and efficient key distribution is provided by the security scheme.

II. CONSTRAINTS IN WSN

Energy constraints: Energy consumption for the sensor nodes transducer, communication with other nodes and computation in microcontroller. Every bit transmitted in the WSN consumes energy more than the energy required for executing thousand instructions. Since the energy consumption is higher for the communication than the computation.

Computation constraints: Processors in the sensor nodes have low computation capability than wired networks. Therefore complex cryptographic algorithms not used for the computation.

Memory constraints: Sensor node's memory has RAM and flash memory. Purpose of RAM is to store sensor data, application programs and computations. Purpose of flash memory is to store downloaded application code. After the OS and application code execution, there is no space for executing complex cryptographic algorithms.

Communication constraints: Range of communication is based on the receiving signal strength. Signal strength depends on the environment.

III. SECURITY REQUIREMENTS IN WSN

The purpose of security services is to provide security for the resources from unauthorized users. The security requirements are authentication, authorization, availability, integrity, confidentiality, freshness, non repudiation, forward secrecy and backward secrecy [11], [13].

Authentication: Ensures the originality of communication from node to another node.

Authorization: Ensures the information that is provided only by the authorized sensor nodes.

Availability: Ensure that the network services are available even if there is an attack.

Integrity: Ensures that there is no modification in the message.

Confidentiality: Ensure that the message is understood only by the desired recipient.

Freshness: Denotes recent data and ensures that the message is not replay by an attacker.

Non repudiation: Denotes that the sensor node cannot deny the sending of message.

Forward secrecy: If the node leaves from the network, then it does not read the future messages.

Backward secrecy: If the new node joins with the network, then that node is able to read messages that are transmitted already.

Threat model: Assumed that if an attacker comes to know the security mechanisms used in the WSN then he can able to compromise a node. Once the node is compromised attacker can get the key within that node [11], [12].

Classification of attacks

The sensor network attacks are classified as follows [13]:

Outsider attacks: Attack sensor network from outside nodes that is not belong to a WSN.

Insider attacks: Sensor nodes within the network attack the sensor network by misbehave.

Passive attacks: Monitor the packets communicated with in the WSN.

Active attacks: Modify the original message or create a false message that is transmitted.

Mote class attacks: Attack sensor network by using some nodes with same capability.

Laptop class attacks: Attack sensor network by using devices like laptop. Since these devices are more powerful than the sensor nodes in WSN.

The security scheme in WSN is evaluated based on the metrics. Such as security, resiliency, energy efficiency, flexibility, scalability, fault tolerance, self healing and assurance.

IV. ATTACKS IN WSN

Based on the security requirements, categorize attacks in WSN [8], [11], [12]. Such as attacks on

ISSN: 2278 – 909X All Rights Reserved © 2014 IJARECE

network availability, attacks on secrecy and authentication and stealthy attacks against service integrity.

A. Physical layer attacks

Jamming: It is an attack which interfere radio frequencies used among sensor nodes in the network. If the jamming source is having high power then the entire network gets disrupt. An attacker can able to disrupt the entire network by using less powered jamming source also. Protection techniques against the jamming attacks are spread spectrum and code spreading. Frequency hopping spread spectrum use pseudo random sequence that is known by the sender and receiver to switch the carrier frequency channel for transmitting signals. Code spreading technique is not suitable for WSNs since it needs high energy and complex design.

Tampering: An attack in which extracting sensitive data from sensor node by alter or replace to create compromised node and that data may be a cryptographic key. Tamper proofing technique protects from this attack but it needs additional cost.

B. Data link layer attacks

Collisions: Two nodes simultaneously try to access same frequency then the collision will occur. Changes may occur in data portion of the packet due to collision. It causes to discard those packets since checksum mismatch at the receiver side. Error correction codes defense collisions and it requires additional communication and computation overhead.

Exhaustion: Resource exhaustion due to continuous collision that is trying to transmit corrupted packet repeatedly causes energy depletion. Exhaustion prevented by limits the MAC admission control to ignore excessive requests. Another prevention technique is to allocate time slot for the nodes to transmit by using time division multiplexing.

Unfairness: An attacker use link layer attacks intermittently such as exhaustion, collision causes unfairness in the network. Reduce the unfairness by using small frames in the communication since it needs lesser time to transmit.

C. Network layer attacks

Spoof, alter, or replay: Routing information is spoofed, altered, or replayed when two nodes exchange this information through the disruptions such as partition the network, create fake error message, short or extend source routes, increase end to end latency, attract or repel network traffic, and generate routing loops. Spoof and alter methods prevented by adding message authentication code with message packets. Defend the replayed information by adding counter or timestamps in the message packets.

Selective forwarding: The sensor node in the multi-hop network forward only particular packets and drop others. This attack is known as black hole attack. One countermeasure for this attack is to use multiple paths to forward packets. Another countermeasure is to detect the black hole node or take an alternative route.

Sinkhole: Adversary's sensor node or attacker compromised node acts as a sinkhole. Neighbor nodes choose sinkhole node as the next node to route their data packets. It leads to selective forwarding attack.

Sybil: In this attack, a sensor node has more than one identity in the network. It affects distributed storage, fault tolerance schemes and maintenance of network topology. In distributed storage scheme, achieve the level of redundancy based on three replicas. If the adversary's compromised node pretends like two of three nodes then the storage scheme concludes but it is not real.

Wormholes: An adversary uses a link to replay network messages between two parts of the network. This link is called as wormhole and it is created by single node or pair of nodes to forward messages. Wormhole attack is prevented by packet leashing mechanism.

Hello flood: Most of the algorithms assume that if the hello packet received from the sender node then consider that node is within the range. An adversary has large transmission power and send hello packet to all the network nodes from outside the network. Hello packet received nodes believe that the sender is within the network. So the adversary can easily steal sensitive data.

Acknowledgement spoofing: An adversary node intended to deceive the acknowledgement of eavesdropped packets for giving the false information to the neighborhood nodes.

D. Transport layer attacks

Flooding: It leads to memory exhaustion when the protocol needs to maintain state at both end of connection. Attacker makes new connection request repeatedly for emptying the resources.

ISSN: 2278 – 909X All Rights Reserved © 2014 IJARECE

Countermeasure for this attack is to solve puzzle by the client for making connection with the server.

De-synchronization: An attacker disrupts the existing connection by sending spoofed messages repeatedly to the destination host since it reduces the host's ability to exchange data. Countermeasure is to authenticate all the packets exchanged during communication.

V. SECURITY

Clustering in WSNs achieve network scalability and management, which increases lifetime of sensor node and reduces bandwidth consumption. Low Energy Adaptive Clustering Hierarchy protocol balances the energy consumption in cluster based WSNs. This protocol rotates cluster head nodes randomly among the sensor nodes in the network to balance the energy consumption. Therefore adding security to this protocol is defiance. This type of protocols affect from orphan node problem since it uses symmetric key management for security [3]. If a sensor node does not share a pairwise key with other sensor nodes in its preloaded key ring then the orphan node problem occurs. Due to this problem sensor node cannot joins with any cluster. So that the sensor node elected itself as a cluster head and orphan node problem increases the number of cluster heads in the network. Therefore this problem increases transmission overhead and energy consumption.

Instead of using symmetric key management system for security apply the asymmetric key management system to solve the orphan node problem in the network. Asymmetric key management system offers digital signature service. This service provides digital certificate which is used to bind the identification of signer and public key. The identity based digital signature scheme [2], [13], [13] draw an entity's public key from its identity details such as name or identity number based on the hardness of factoring integers from identity based cryptography.

VI. CLUSTERING TECHNIQUES

A. Network model

Depending on different applications of wireless sensor networks consider various architectures, design goals and constraints. Some of the architectural parameters and their implications are given below [9].

Network elements: Fundamentally WSNs be made up of sensor nodes, base station and sensed events. Suppose that sensor nodes are fixed in most of the network architectures. Sensor node is considered compulsory to support the mobility of base station or CH sensor nodes on some occasions. Mobility of sensor node changes the node membership dynamically. Therefore clustering technique is very challenging. Sensed events may be intermittent or continual according to the applications. Forest fire detection is the example of intermittent events, allowing the network to create traffic only when reporting and not necessary the network to be in active always. In the case of continual events, the network generates traffic consequently since reporting periodically and these events stable clusters. Events fluctuate in intermittent events favor adaptive clustering strategy.

Data aggregation/fusion: Multiple nodes might produce redundant data, similar packets since number of transmissions would increase and it causes to increase energy consumption and network traffic. To achieve energy efficiency and traffic optimization use data aggregation functions such as suppression, min, max and average. Each sensor node does these functions either partially or fully, by permitting sensor nodes to conduct. All of these aggregation functions are applied to specialized nodes in some network architectures. Through signal processing techniques possible to do data aggregation and it is mentioned as data fusion where a node combines signals from multiple nodes by using beam forming technique and generates more accurate signal by minimizing the noise. CHs perform data aggregation/fusion means those CHs require limited number of sensors per cluster to reduce overburden. Notice that on some occasions essential to appoint backup CHs for a cluster or rotate the role of CH among sensor nodes influence the scheme of cluster.

Placement of nodes and constraints: The deployment of nodes may be deterministic or self-organizing. The sensor nodes are placed manually and pre-determined paths are used for data routing in deterministic node deployment. For this reason no need for clustering or it is preset. Although, in self-organizing node deployment use an infrastructure like an ad hoc manner. In that case the sensor nodes are randomly scattered, the BS or the CH position is crucial. If the nodes are scattered then optimal clustering comes to be major issue to save energy. CHs are selected from placed sensor nodes in networks of homogeneous sensor nodes and in order to avoid quick depletion of their energy those CHs are tasked carefully. Also have to consider issues

ISSN: 2278 – 909X All Rights Reserved © 2014 IJARECE

like the corresponding CHs nearness to the BS and communication range. This inter-CH connectivity issues affect the scheme of cluster. In heterogeneous networks, the sensor nodes have different functionalities since the clustering process may have more constraints and to preserve resources of those nodes have to either avoid such special nodes or limit the subset nodes of CHs.

B. Objectives for network clustering

Load balancing criteria's: Objective for the case where CHs do data processing and duties of intra-cluster management are that distribution of sensor nodes among clusters evenly since it reduces data delay. To achieve the expected performance goals CHs has to balance the load among clusters and to extend the lifetime of the network have to set an equal-sized clusters since it prevents energy consumption of a subset of a cluster heads.

Fault tolerance techniques: The WSNs usually operated at the harsh environments and thus the nodes can easily get damaged. Therefore it is necessary to tolerate the failure of CHs for avoiding the loss of significant data in such applications. Re-clustering the network is one of the solutions to recover the failure of CHs, but this solution is not suitable for the on-going operation. For this reason use contemporary fault-tolerance technique that is assigning backup CHs to recover. Whenever the variations occur in the normal network operation the backup CHs have to play their role. The leaf nodes join the cluster based on the receiving signal strength and when CHs does not include nodes because of the long radio range then the neighboring CHs can adapt those disjointed nodes. Solve the issues in fault-tolerance and load balancing in a cluster by means of rotating the role of CHs among sensor nodes in that cluster.

Maximize connectivity and minimize delay: Inter-CH connectivity plays an important role while selecting the CH among sensor nodes in a cluster. The goal of connectivity is to ensure the availability of path between CHs to base-station. When assigning the CH role to some sensors, the objective of connectivity makes one of many variations in K-dominating set problem that is K-hop clustering. Intra-cluster connectivity is responsible for data latency. Factoring the delay usually based on maximum number of hops 'K' in a path.

CH deployment: Specifically, the resource-rich nodes can be selected as CHs such as laptop computers, robots and mobile vehicle. The designer wanted to employ the smaller number of nodes since their size, cost, vulnerability and complexity of deployment.

Increase network longevity: The lifetime of the network depends on the sensor nodes energy. The resource-rich CH has to minimize the energy consumption for intra-cluster communication and deploy the CH near to most of the sensors in its cluster if possible. In the case where CHs are normal sensors and has to increase their lifetime by reduce the work load. Adaptive clustering, combined clustering and route setup are the concerns to increase the lifetime of the network. Fig. 2 shows intra and inter-cluster communication.

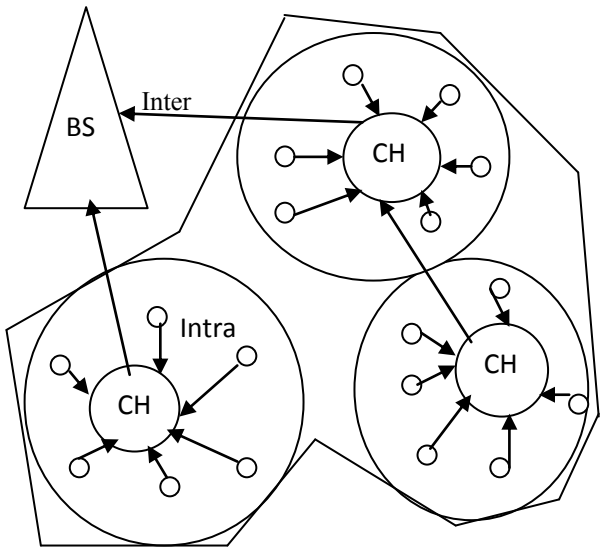

Fig. 2. Data communication in cluster

C. Clustering attributes

The clustering algorithms are categorized based on the following attributes [5], [7]:

Cluster count: The number of CHs and clusters are predetermined in some approaches. Selecting CHs randomly gives variation in number of clusters.

Stability: In the adaptive clustering scheme, changes made in the cluster count and node's membership. In the fixed clustering scheme, cluster count and node's membership remains the same.

Intra-cluster topology: There is a direct communication between sensor and its nominated CH in some clustering scheme. Although, there is a need for multi-hop connectivity between sensor and CH in sometimes; particularly the case where

ISSN: 2278 – 909X All Rights Reserved © 2014 IJARECE

sensor's range of communication and/or CH count is limited.

Inter-CH connectivity: If the CH does not have long haul communication capabilities then the connectivity has to be provided between CHs and base-station. Clustering scheme ensures the feasibility of creating an inter-CH route in that case.

Mobility: Cluster has to continuously maintain when a CH acts as a mobile and membership of a sensor changes dynamically. In the case where fixed CH contributes to give stable clusters and help managing inter and intra cluster network. To make a network performance good CHs can move for a particular distance and relocate itself.

Node capability: Subset of nodes that are deployed considered as CHs which has computation and communication resources highly.

CH role: In a cluster CH perform data aggregation/fusion or acts as a relay. On some occasions CH acts as sink or base station depending upon identified phenomena.

Approaches: Clustering has to be carried out in a distributed way, when CHs are usual sensor nodes. In some cases, a centralized authority controls membership of cluster and partitions the offline nodes. When CHs have more resources, hybrid scheme can be established. In the case when an individual CH making its own cluster, inter CHs coordination is carried out in a distributed way.

Node grouping: Several objectives have been followed for making cluster. For example network connectivity, load balancing, etc.

Selection of CH: CHs can be selected randomly from the deployed nodes or pre-assigned.

Complexity: The rate of convergence and complexity of the algorithm may be constant or variable based on number of CHs and/or sensors.

VII. CLUSTERING ALGORITHMS

Based on the convergence rate, the clustering algorithms are divided into two [4], [5], [7]:

A. Algorithms with variable convergence time

Noteworthy factor is the time in the clustering algorithms convergence. Convergence time is increased corresponding to number of nodes in the network. Networks having maximum number of nodes possess highest convergence time. Generally, variable convergence time algorithms control the cluster properties efficiently than constant convergence time algorithms.

Linked cluster algorithm: To maximize the network connectivity LCA is introduced. The goal is to create network topology which can handle nodes mobility. Through clustering, CHs form a backbone network which is used to connect the cluster members while on the move. Assumptions in LCA are synchronized nodes and accessing the medium based on time. Assign the slot in the frame by a node and that matches its ID. In the first round, every node in the network broadcasts its ID and makes attention on transmission of remaining nodes. Node broadcasts the neighbors by next round and thus every node find out its first hop and second hop neighbors. If a node has highest ID among its neighbor nodes then it becomes a CH. Since LCA yields more number of clusters, selection of CH approach refined. The idea is to select a node randomly as the first CH and assign neighbors to that CH to form first cluster. The lowest ID node in the cluster selected as a second CH and assigns the nodes that are not reachable to the first CH to form a second cluster. Repeat this procedure to form next cluster and so on [6], [10].

Adaptive clustering: Cluster the network reduces delay in the data delivery and unique code is assigned to the cluster. A single hop intra cluster topology is created like LCA. A CH manages code selection to communicate with the neighboring CHs. This algorithm control size of cluster by reuse of channels and reduce delay in data delivery by avoid inter cluster routing. For intra cluster communication TDMA is used like LCA. Adaptive clustering algorithm supports multimedia applications efficiently.

Random competition based clustering: This algorithm mainly concentrate on stabilization of cluster. It uses first declaration wins rule. By this rule, any node can manage other nodes within that radio coverage if that node claims to be a CH. Neighboring nodes join the cluster based on CH claim packet and this packet is broadcasting periodically to maintain cluster stability. Neighbor nodes can send CH claim packet concurrently while on-going claim is unknown. Concurrent broadcast create conflict and it arises when there is a time delay between broadcasting and receiving CH claim packet. This algorithm uses random timer and node ID for arbitration to avoid such problem. Each node in the network reset random time value before it sends CH claim packet and it stops sending CH claim packet

ISSN: 2278 – 909X All Rights Reserved © 2014 IJARECE

while it receives any other CH claim during random time. Algorithm uses node ID if the conflict not solved and select a node as CH with lower ID that is nearer to existing CH. Node mobility still creates problem in selecting CH.

GS^3: This algorithm is used to create cellular hexagonal structure for wireless sensor network by dividing the area into cells with equal radius and this structure is used for grouping the nodes. GS stands for Geometric Size. Notice that geographical boundary is important for cluster based wireless sensor network. Consider the circle that includes all nodes in the cluster and that circle's radius is a measure for geometric size. A cluster with huge radius consumes high energy, communication within cluster have high reliability, reduces spatial reuse of signals. Assume large and little nodes are there in the system. Large nodes start cluster creation process and small nodes act as interface. One of the large nodes selected as a head and forms a cluster by becoming unselected neighboring little nodes as their members. After selecting the cell head, it is relocated to the center of that cell. This process is continued until all the cells established in specified network area. GS^3confirmed the number of CHs and its placement in the system, convergence requires one-way diffusion under perturbation areas that is disturbed areas, long intra cluster communication is possible and it is suitable for dynamic networks.

EEHC: The purpose of Energy Efficient Hierarchical Clustering algorithm is to increase the lifetime of network. CHs are responsible for collecting leaf sensor node's data within cluster and send the aggregated result to the base-station. There are two stages that is single level clustering and multi-level clustering. At the initial stage or single level clustering, every node informs itself as a CH with probability p to the nearness nodes within k hops range. These CHs considered as volunteer CHs and receiving neighbor nodes that is not itself CH considered as leaf node to the nearest cluster. A node that does not receive any message within given time interval t then it is considered as forced CH which is not located in a cluster. In second stage or multi-level clustering, repeat the clustering process at the level of CHs to create h levels of cluster hierarchy and assume that h is the highest level. Sensor nodes gather data from network environment and these data transmitted to first level CHs. Aggregated data at the first level CHs transmitted to the second level CHs and so on. The top level CHs transmit aggregated data to the base station. This algorithm is applicable for large network since it has time complexity of $O(k_1+k_2+....+k_h)$. The operations involved in the network consume energy based on the parameters p and k of the algorithm. By choosing optimal parameter values can reduce energy consumption [1], [6], [10].

B. Algorithms with constant convergence time

Algorithms that converged number of iterations, regardless of number of nodes are known as constant convergence time algorithms. These algorithms follow a localized strategy that algorithms executed by the nodes independently and cluster membership decisions are based on their own state and neighbor's state.

LEACH: Low Energy Adaptive Clustering Hierarchy algorithm creates cluster depending upon the receiving signal strength and CHs act as routers to reach base-station. Data aggregation and fusion operations performed within a cluster. LEACH uses distributed algorithm to create cluster, where nodes make decisions by their own. A node wants to be CH with probability p, first broadcasts its decision. Every non CH node joins cluster by selecting the CH with minimum energy. Load balancing is done by rotate the role of being CH periodically among the cluster nodes. Every node selects a random number T between 0 and 1 and if this number is less than the threshold value then node becomes CH for current round. There is a chance that a node with low energy selected as CH and then the entire cluster gets damaged while that node dies. Assumed that CH has high communication range to communicate directly with the base-station and this assumption is not realistic since the nodes in the cluster are regular sensors and due to the signal propagation problems base-station does not reach the nodes directly. Every node in the cluster can directly communicate with the CH and then with the base-station by using intra and inter cluster topology. Therefore, LEACH is not suitable for large networks [1], [3], [6], [13].

FLOC: Fast Local Clustering Service is a distributed technique which yields clusters with equal size and lowest over-lap. Nodes are classified as inner and outer nodes corresponding to their proximity to the CH by the radio model that is assumed. Inner nodes communicate with CH with small interference and outer nodes communication may be lost. To maximize the robustness of intra-cluster traffic FLOC aids inner node membership. Consider the transitions of node's state such as idle, candidate, CH, inner and outer. Until a node receives an invitation from any CH, node's state is idle and waiting for some time to get invitation. If the node does not receive any invitation, node's state changed to candidate state

ISSN: 2278 – 909X All Rights Reserved © 2014 IJARECE

that is the node becomes candidate CH and broadcasts a message of candidacy. After receive an invitation a recipient node *'r'* that is inner member of another cluster 'c_r' already will send reply message about the membership details to the candidate CH. Then the candidate CH will understand the conflict and join c_r as outer node. If no conflict then it becomes a CH and invite members. If an idle state node does not receive closer CH invitation then it joins as an outer node with the cluster. If an idle state node receives closer CH invitation later then the node changes its membership. FLOC supports constant convergence time and it does not depend on the network size. Outer nodes can switched as inner nodes in different cluster since it has self-healing capability. This algorithm can be executed by the new nodes and this node can form a new cluster by inviting outer nodes of neighboring clusters as its member or it can join with the existing cluster. The problem is not clear that how the data are scattered among clusters.

ACE: Algorithm for Cluster Establishment estimates a node's potential as a CH prior to that node becomes CH and it works in iterations for individual nodes. Spawning and migration are the two functions in ACE. If the node is decided to become a CH then it spawns new cluster. Selected CH sends invitation to its neighbors. Nodes that receive invitation become follower of that new CH. Migration is another function which selects best node for being CH. Every CH tests neighbor node's potential for being CH and if any of these neighbors has highest number of followers than it has then that CH decides to step down. A node is considered as best node to become CH if it has more followers and minimum overlap. Compare ACE with the Linked Cluster Algorithm, ACE covers the whole network in three iterations. The time complexity is *O(d),* where *d* is the density of node per cluster. Increase the regularity of cluster design by increase number of iterations. By multiplying the degree of overlap, provide attraction between remote clusters. Clusters are scattered corresponding to the node density to increase spatial coverage. An ACE supports integration of new nodes and heals structure damage easily.

HEED: Hybrid Energy-Efficient Distributed clustering algorithm is a distributed one in which selects CH from the sensor nodes that are deployed by considering energy and cost of communication. Sensors with large residual energy selected as CH. There are three characteristics in HEED. First one is probability of becoming CHs is little within the transmission range of two nodes. It denotes that network of distributed CHs. Second one is energy consumption for all the nodes not equal. Third one is to make sure inter-CH connectivity the probability of selecting CH is adjustable. Every node is connected with only one cluster and it can communicate with CH directly in HEED. In fig. 3, H denotes cluster head and C denotes cluster. There are three phases in this algorithm.

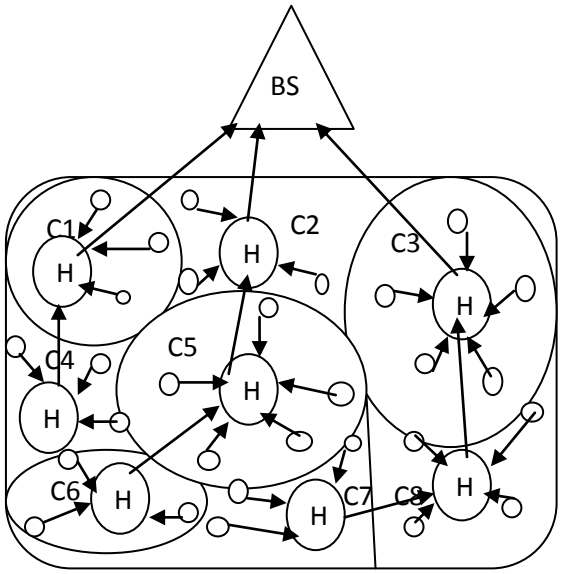

Fig. 3. Multi-hop with clustering

Initialization phase: Initialize the percentage of CHs in the network of all nodes. Percentage value, C_{prob}, used as boundary for initial CH announcements. Probability of node becoming CH, CH_{prob}, it is calculated as $CH_{prob}=E_{residual}/E_{max}*C_{prob}$, where $E_{residual}$ is the sensors current energy, E_{max} is the maximum energy depends on the battery. The threshold value, t_{min}, which is inversely proportional to E_{max} and it should not be $CH_{prob}<t_{min}$ in the network.

Repetition phase: When the sensor node finds its CH, it stops continuing iterations. If the node not finds CH then that node elect itself as a CH and broadcasts an announcement to the nearest nodes. Every sensor doubles the value of CH_{prob} and goes to next iteration. When the value of CH_{prob} reaches one, the node stops execution of this phase. For this reason, sensor announces with two types of CH status to its nearest nodes. They are tentative and final status. If the node's $CH_{prob}<1$ then it becomes tentative CH. If that node discovers lower cost CH then it changes the status to normal node at later iteration. In final status, if the node's $CH_{prob}=1$ then it becomes CH permanently.

Finalization phase: The node decides either selects lower cost CH or announce itself as CH. The

ISSN: 2278 – 909X All Rights Reserved © 2014 IJARECE

modified version of HEED re-executes the algorithm for orphaned nodes and would decrease number of CHs.

Attribute-based clustering: The goal is to spread data across the network efficiently. Data attributes hierarchy mapped to the network topology for creating the cluster. The base-station send request to the nodes for creating cluster. Based on those request the node decide to elect itself as a CH and wait for some random period depends on the node's energy. Nodes having high energy wait for long time. When the node selected itself as a CH it broadcasts invitation. If that node receives any CH claim packet from its neighbors during the wait time then it leaves CH invitation. Resend such packet when packets hop count is incremented by one. CH announcement sending node has different attributes and receiving node forms cluster based on these attributes. Increase the lifetime of network by rotating the role of CH among the cluster nodes. CH failure can be identified easily since the CH sends heartbeat message to members periodically. Assume CH is failure when the member do not receive heartbeat message and one of member takes the role of CH. Density of the network is high means recovery is done at the CHs level.

Table I. Comparison of clustering algorithms

Algorithm	Convergence time	Energy efficient	Cluster stability	Node mobility	Objective of clustering	CH selection	Scheme
LCA	Variable O(n)	No	Moderate	Possible	Connectivity	Random	Distributed
Adaptive clustering	Variable O(n)	N/A	Low	Yes	Bandwidth gain & QOS	Random	Distributed
RCC	Variable O(n)	N/A	Moderate	Yes	Stability & simplicity	Random	Hybrid
GS^3	Variable O(n)	N/A	Moderate	Possible	Scalability & fault tolerance	Pre-assigned	Distributed
EEHC	Variable O(k1+..+kh)	Yes	N/A	No	Save energy	Random	Distributed
LEACH	Constant O(1)	No	Moderate	Fixed BS	Save energy	Random	Distributed
FLOC	Constant O(1)	N/A	High	Possible	Scalability & fault tolerance	Random	Distributed
ACE	Constant O(d)	N/A	High	Possible	Scalability & load balancing	Random	Distributed
HEED	Constant O(1)	Yes	High	Stationary	Save energy	Random	Distributed
Attribute-based clustering	Constant O(1)	Yes	High	No	Bandwidth gain	Random	Distributed

ISSN: 2278 – 909X All Rights Reserved © 2014 IJARECE

VIII. CONCLUSION

Wireless sensor networks used in many applications like health care monitoring, air pollution monitoring, forest fire detection, landslide detection, water quality monitoring, natural disaster prevention, machine health monitoring, waste water monitoring, data logging, structural health monitoring, combat field surveillance, border protection, etc. Such applications require more number of sensors to managing the network and security in data transmission among sensors. Increase scalability in WSNs by using the clustering approach. In this paper, we classify the clustering algorithms and highlighted similarities and differences among those clustering algorithms.

REFERENCES

[1] H. Lu, J. Li, and G. Wang, "A Novel Energy Efficient Routing Algorithm for Hierarchically Clustered Wireless Sensor Networks", Proc. Fourth Int'l Conf. Frontier of Computer Science and Technology (FCST), 2009.

[2] H. Lu, J. Li, and H. Kameda, "A Secure Routing Protocol for Cluster-Based Wireless Sensor Networks Using ID-Based Digital Signature", Proc. IEEE GLOBECOM, 2010.

[3] P. Banerjee, D. Jacobson, and S. Lahiri, "Security and Performance Analysis of a Secure Clustering Protocol for Sensor Networks", Proc. IEEE Sixth Int'l Symp. Network Computing and Applications (NCA), 2007.

[4] A.A. Abbasi and M. Younis, "A Survey on Clustering Algorithms for Wireless Sensor Networks", Computer Comm., vol. 30, nos. 14/15, 2007.

[5] Basilis Mamalis, Damianos Gavalas, Charalampos Konstantopoulos, and Grammati Pantziou, "Clustering in Wireless Sensor Networks", Zhang/RFID and Sensor Networks, 2009.

[6] Vinay kumar, Sanjeev jain, and Sudarsan tiwari, "Energy Efficient Clustering Algorithms in WSNs: A Survey", IJCSI, Vol. 8, Issue 5, No 2, September 2011.

[7] Xuxun Liu, "A Survey on Clustering Routing Protocols in Wireless Sensor Networks", Sensors 2012.

[8] F. Akyildiz, W. Su, Y.Sankarasubramaniam, and E. Cayirci, "Wireless Sensor Networks: A Survey", Computer Networks 38 (2002).

[9] K. Akkaya and M. Younis, "A Survey on Routing Protocols for Wireless Sensor Networks", Elsevier Journal of Ad Hoc Networks 3 (3) (2005).

[10] S. Banyopadhyay and E. Coyale, "An Energy Efficient Hierarchical Clustering Algorithm for WSNs", in: Proceedings of the 22nd annual conference of the IEEE computer communications Societies (INFOCOM April 2003).

[11] Y. Wang, G. Attebury, and B. Ramamurthy, "A Survey of Security Issues in WSNs", IEEE communications surveys, 2nd quarter 2006, volume 8, no. 2.

[12] Chris Karlof and David Wagner, "Secure Routing in Wireless Sensor Networks: Attacks and Countermeasures", Ad Hoc Networks, vol. 1, nos. 2/3, 2003.

[13] Suraj Sharma and Sanjay Kumar Jena, "A Survey on Secure Hierarchical Routing Protocols in Wireless Sensor Networks", ICCS, 2011.

[14] Huang Lu, Jie Li, and Mohsen Guizani, "Secure and Efficient Data Transmission for Cluster-Based Wireless Sensor Networks", IEEE transactions on parallel and distributed systems, vol. 25, no. 3, march 2014.

K. Gowtham, PG student, Department of computer science and engineering, Info Institute of Engineering, Coimbatore, India,

Mr. S. Dhanasekar, Assistant professor, Department of computer science and engineering, Info Institute of Engineering, Coimbatore, India.

,

ISSN: 2278 – 909X All Rights Reserved © 2014 IJARECE