

Clustering in Wireless Sensor Networks- A Survey

Sukhkirandeep Kaur

National Institute of Technology, Srinagar, 190006, India
E-mail: kirangill0189@gmail.com

Roohie Naaz Mir

National Institution of Technology, Srinagar, 190006, India
E-mail: Naaz310@nitsri.net

Abstract—Increased demand of Wireless Sensor Networks (WSN) in various applications has made it a hot research area. Several challenges imposed which include energy conservation, scalability, limited network resources etc. with energy conservation being the most important. Clustering improves the energy efficiency by making high power nodes as cluster heads (CHs) which reduces the chance of energy depletion of nodes. Scalability, fault tolerance, data aggregation, energy efficiency are some of the main objectives of clustering. This paper discusses various challenges associated with clustering and different methods or techniques developed to overcome these challenges. Various clustering approaches have been summarized and few prominent Quality of service (QoS) based clustering routing protocols for WSN have been identified. Comparison of these approaches and protocols is discussed based on some parameters.

Index Terms—Wireless sensor networks (WSN), Quality of service (QoS), Clustering, Routing protocols.

I. INTRODUCTION

The ability of Wireless Sensor Networks (WSN) to detect and observe any physical phenomenon has made them very popular in today's era. The low power devices called sensor nodes collaborate to perform any given task and constitute a network called a sensor network. The nodes are equipped with sensing and processing capabilities and they sense any physical phenomenon such as temperature, pressure etc. and send this data to a base station (BS) or sink using single hop or multi hop communication. Multi hop communication is preferred over single hop as it reduces energy consumption. WSN is quite different from traditional adhoc and wired networks in terms of low energy, limited resources and their applications. The protocols and algorithms developed for these networks cannot be directly applied to WSN. Researchers need to consider energy and resource constrained nature of WSN while designing protocols.

Sensor networks work in a wide range of applications like military, surveillance, health monitoring etc. Depending upon the application, they can be deployed in deterministic or a random manner. Some applications are used in hostile environments and nodes are deployed in these environments in an uncontrolled manner and it is not possible to change the battery of nodes in these environments. So in case of random deployment nodes have to reconfigure themselves and large number of nodes may be deployed to overcome battery issues. In other applications where deterministic deployment is possible, position and placement of nodes is determined prior to deployment in a controlled manner. Node placement is a very important step and proper node placement can result in more energy efficient solutions. Different strategies for node placement are discussed in [1]. After deployment, coverage and connectivity are important issues to be considered, so proper coverage depending upon the application should be maximized and connectivity should be maintained. Different methods and issues related to coverage are discussed in [2, 3]. As a sensor network is used in hostile environment, energy conservation becomes an important issue and scalability problem also arises. Use of clustering overcomes these issues by incorporating network aggregation and processing which further reduces amount of data to be sent and overcomes several challenges related to resource constrained nature of WSN.

WSN suffers from many challenges that include resource constraints, operation in harsh environments, data redundancy, scalability, energy conservation, wireless channel, lack of global identification, node failures, in-network processing, fault tolerance etc. Efficient use of resources can overcome the problem of resource constraints. In harsh environments it is not possible to change the battery of nodes. Energy conservation becomes important issue that can be overcome by adaptive network operation. Large amount of data is involved in WSN as nodes sense same physical phenomenon in same region that leads to huge redundancy that can be minimized by in-network processing, data aggregation and data fusion techniques. This in-network processing requires efficient hardware and algorithms designed considering constrained nature

of WSNs. Design of different processing and compression algorithms for WSN is an important area of research in multimedia sensor networks. Sensor networks consist of large number of nodes where it is not possible to assign the global identification (GID). Positioning system consisting of GPS can be used to provide location of nodes but this further increases cost, so data centric algorithms need to be used. Scalability is also of important consideration in large scale network application and as number of nodes increases; network should maintain its performance.

Low power, low data rate, and low-cost wireless sensor communication is handled by different technologies. Choice of particular technology depends upon application. IEEE 802.15.4 is used for applications requiring short range communication. Zigbee, 6LoWPAN, Wireless HART are different standards of 802.15.4 for WSN. Zigbee [51] is typically used in low data rate applications that require long battery life and secure networking that provides reliability and interoperability. 6LoWPAN [50] refers to IPv6 over Low power Wireless Personal Area Networks and enables IPv6 communication over IEEE 802.15.4 links and assures interoperability with other IP devices. Wireless HART operates in 2.4GHz ISM band using IEEE 802.15.4 standard radios and is used for industrial and control applications and provides reliability, security, energy efficiency and also provides compatibility with existing devices. Bluetooth low energy

(BLE) overcomes the limitations of Bluetooth for WSN in terms of reduced power consumption, long battery life and reduced cost but maintains a similar communication range.

It finds its applications in healthcare, sports and fitness, security, and home entertainment, etc. For residential, light commercial environments and home automation applications, Z-WAVE is the promising technology. For home automation, it can be used for fire alarms, controlling light, detecting gas and flood [2]. WAVENIS is ultra-low power, long range communication technology developed by coronis systems. It has long battery life and provides reliable transmissions. It is used in environmental, agriculture, industry, building, biological, chemical and healthcare applications. EnOcean is an emerging WSN technology that is promoted by EnOcean Alliance [53]. It uses energy harvesting wireless technology that draws energy from surroundings and uses that energy electrically. It is best used in buildings and industrial automation.

The rest of the paper is organized as follows. Section II explains the clustering process in WSN followed by literature survey. Different clustering algorithms categorized on basis clustering parameters are discussed in Section IV along with the comparative analysis of different algorithms. Section V discussed various Clustering based QoS routing protocols and a comparison has been made based on these parameters.

Table 1. Comparison of different technologies

	Frequency	Data rate	Range	Power consumption	Applications
ZIGBEE	868 MHz 2.4 GHz	20 Kbps 250 Kbps	10-100 m	Low	Wireless light switches, Traffic management, Industrial equipments.
6LoWPAN	2.4 GHz	250/40/20 Kbps	1-100 m	Low	Smart home, vehicle monitoring , traffic monitoring, tele-health
Wireless HART	2400-2483.5 MHz	250 Kbps	1-100 m	Low	Industry automation
Bluetooth low energy (BLE)	2.4 GHz	1Mbps	200m	Ultra low	Healthcare, sports and fitness, security, and home entertainment
Z-WAVE	1 GHz	40 Kbps	30 m	Low	Residential, light commercial environments and home automation applications
WAVENIS	868,915,433 MHz	100 Kbps	1-4 km	Ultra low	Environmental, agriculture, industry, building, biological, chemical and healthcare applications.

II. CLUSTERING IN WSN

In clustering; sensor nodes are divided into a virtual group known as cluster and nodes in different clusters execute different functions. Clustering is defined as grouping nodes into clusters based on some criteria and selecting the most efficient node as CH from each cluster.

Clustering architecture has several advantages over other approaches, it enhances scalability, reduces amount of data to be transmitted by aggregating and summarizing the data, decreases the number of relay nodes, achieves load balancing, improves energy consumption and increases network lifetime etc. These advantages can prove clustering to be efficient approach for routing in WSN.

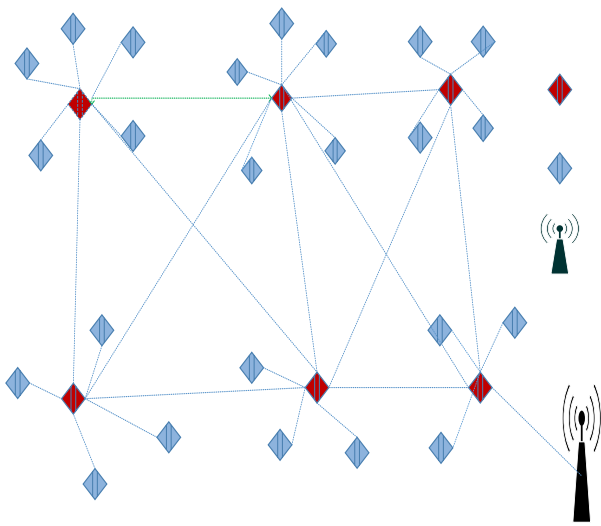


Fig.1. Clustering in WSN.

To reduce redundancy and to enhance scalability, clustering is the best approach to follow. In flat architecture of homogeneous nodes, all nodes will have same functionality and the entire load will be imposed on the nodes nearby the sink. Clustering alleviates this problem by considering CH as a powerful node which will be responsible for processing and transmitting data to BS. Clustering overcomes many problems related to WSN, it reduces redundancy, improves network lifetime and enhances scalability. Energy consumption can be reduced in intra-cluster communication by putting nodes to sleep mode while they are idle. Moreover cluster formed near sink should have less number of nodes in their cluster to save energy from intra cluster communication as they have to handle most of the traffic. Clustering helps in reduction of large amount of data by in-processing the data at CH. In-processing involves compression and aggregation/fusion, done at CH so that the amount of data to be transmitted to BS is reduced.

As amount of data is reduced and only CH is involved in transmission of data, energy consumption is reduced.

As amount of data is reduced and only CH is involved in transmission of data, energy consumption is reduced. If we evenly distribute CH and dynamically change CH,

load balancing can be achieved. Clustering also improves network lifetime and helps in achieving QoS. CH selection process should be carefully chosen so that it generates well balanced clusters. If we consider static clustering, clusters will drain off their energy faster if there are homogeneous nodes in a cluster. Dynamic clustering overcomes this problem by re-electing CHs but this leads to more energy consumption. Total number of CHs selected also effects network performance as large number of CHs will lead to more energy consumption. If fewer CHs are chosen keeping energy efficiency in mind they end up in retransmission and packet dropping due to poor connectivity. Number of exchanged messages should also be kept low. As WSN is application specific, clustering process should be selected based on application requirements. Different parameters which are involved in clustering process are discussed in next section.

III. RELATED WORK

Villalba et al. in [47] presents a detailed description of routing protocols in WSN. Different design constraints for routing in WSN are explained. This paper presents description of various Optimization Techniques for Routing in Wireless Sensor Networks that includes attribute based, energy efficiency, data aggregation, multi path communication, Quality of Service etc and provides survey of different routing protocols based on these techniques. Different Routing Protocols proposed by Spanish Universities has also been explained.

A survey of different hierarchical routing protocols has been presented by Joshi et al. in [48]. Authors provide detailed description of different routing protocols that includes LEACH, TL-LEACH, TEEN, APTEEN, EECS, HEED, PEAGSIS, CCS.

Boyinbode et al. presented a survey on Clustering Algorithms for Wireless Sensor Networks in [49]. Authors presented various challenges associated with clustering algorithms and explained clustering process that involves selecting CH and forming clusters. Different clustering algorithms has been discussed and compared on basis of CH selection and Cluster formation for non CH nodes.

A Survey on Clustering Routing Protocols in Wireless Sensor Networks has been presented by Xuxun Liu in [43]. Clustering attributes have been categorized into cluster characteristics, cluster-head characteristics, clustering process and entire proceeding of the algorithm. A more comprehensive and critical survey of prominent clustering routing protocols for WSNs compared with previous work is presented.

Different clustering routing algorithms for WSN in detail based on the classification of different algorithm-stages, and their characteristics with advantages and disadvantages has been discussed.

Finally, we summarize the previous work done in Table 2, which highlights the main contributions of each author along with the year of the survey.

Table 2. Summary of Survey on Clustering and Routing in WSN.

Year	Author	Title	Main contribution
2005	Akkaya et al.	A survey on routing protocols for wireless sensor networks[44]	1)Discussion of basic concepts related to the routing process 2) Analysis of data centric, hierarchical and location based protocols 3) Analysis of QoS aware protocols.
2009	Mamalis et al.	Clustering in Wireless Sensor Networks[46]	1)Objectives and Design Challenges of Clustering in WSNs 2)Classification of Clustering Algorithms 3)Probabilistic Clustering Approaches 4)No probabilistic Clustering Approaches
2009	Villalba et al..	Routing Protocols in Wireless Sensor Networks[47]	1)Classification of Routing Protocols in Wireless Sensor Network 2)Optimization Techniques for Routing in Wireless Sensor Network 3)Application of the Optimization Techniques: Routing Protocol 4)Routing Protocols proposed by Spanish Universities
2010	Boyinbode et al.	A survey on clustering algorithms for wireless sensor networks [49]	1)Main challenges for WSN clustering 2)Analysis of popular clustering algorithms for WSNs 3)Comparison of popular WSN clustering algorithms based on many attributes
2010	Bhuyan et al.	Quality of Service (QoS) Provisions in Wireless Sensor Networks and Related Challenges	1)Challenges for QoS Support in WSNs 2)Existing approaches for QoS Support in WSNs 3) survey on QoS Aware Routing Protocols
2011	Joshiet al.	A survey of hierarchical routing protocols in wireless sensor network [48]	1)Discussion of familiar clustering routing algorithms in WSNs 2)Comparison of familiar WSN clustering routing protocols based on energy conservation and network lifetime
2012	Goyal et al.	Routing Protocols in Wireless Sensor Networks: A Survey[45]	Survey of data centric, hierarchical and location based protocols listing their description, advantages and disadvantages
2012	Liu et al.	A Survey on Clustering Routing Protocols in Wireless Sensor Networks[43]	1)Summary of Previous Surveys on Clustering Routing Protocols in WSNs 2).Taxonomy of Clustering Schemes 3)Analysis of Prominent clustering routing protocols in WSNs 4)Comparison of Different clustering routing protocols in WSNs
2013	Jadidoleslamy et al.	An Introduction to various basic concepts of clustering techniques on wireless sensor networks[50]	1) An overview of WSNs explaining its features and characteristics. 2) Different Types of WSNs' Architectures. 3) Clustering and classification of clustering algorithms.

IV. CLUSTERING ALGORITHMS

Cluster count, cluster size and cluster density are the important initial parameters to be considered while designing a clustering algorithm. Selection of a clustering algorithm depends on number of parameters viz type of node, control manner, algorithm type, cluster formation and CH selection, Network Architecture, etc. The details of these parameters are:

A. Type of Node

Clustering process can be formed from two types of nodes i.e. homogeneous and heterogeneous. Homogeneous nodes have same functionality in the network. For example in forest monitoring application, temperature sensing nodes are deployed throughout the network. But the problem in homogeneous nodes is that they all have same energy level and the nodes near the sink deplete their energy faster. Heterogeneous nodes have different capabilities; some nodes having high energy levels are considered as powerful nodes and are selected as CH's. These nodes are responsible for performing operations such as data aggregation/fusion

and they send this data to sink. Energy is balanced among nodes and more intensive operations are assigned to these powerful nodes. LEACH [10] consists of homogeneous nodes and to conserve energy, role of CH is rotated randomly and periodically over all the nodes. As multi-hop communication is used in WSN, nodes near the BS will deplete their energy faster compared to other nodes. This problem can be overcome by reducing cluster sizes near sink or using heterogeneous nodes. Heterogeneous nodes may act as CH with powerful capabilities to be used for in-processing, compression etc. CHs are chosen among nodes based on some criterion such as nodes having more energy to be chosen as CH. Threshold is set and if criterion falls below that threshold value, CH election process is initiated again. Considering heterogeneous nodes helps to achieve energy efficiency and improve network lifetime and the overhead of re-clustering is reduced.

B. Control Manner

WSN clustering methods can be classified as centralized, distributed or hybrid. Control of cluster formation can be transferred to single authority in centralized algorithm and to all nodes in distributed

algorithms. Centralized clustering is performed by one single authority, sink or BS but in case of distributed each and every node has capability to select a CH.

B.1 Centralized Method

A centralized method evolves the sink or CH to acquire global information and the decisions regarding any change in network is taken by CH or sink. The clustering process is controlled by the central authority, base station (BS) or sink. Information regarding nodes is sent to BS and based on that information clusters are formed. This central authority is high energy device which performs energy intensive tasks. Some of algorithms belonging to this category are discussed below:

LEACH-C [5] is similar to LEACH [10] in set-up and steady-phase but differs in choosing a CH. In LEACH-C, during steady phase all the nodes send their energy information to BS and based on that information CH is selected. Load balancing is also taken into account in the clustering process. After finding cluster heads, sink broadcast cluster ID's. Nodes compare their ID with the ID obtained, if it does not match then node determine its TDMA slot and go to sleep mode until its next transmission. This approach leads to more energy consumption as most of the energy is consumed on Base Station.

Energy Efficiency Load-Balancing clustering (EELBC) [6] provides both energy efficiency and load balancing. Sensor nodes are assumed to know their position via gateway by broadcasting their IDs and location to the gateways and based on that, the distance from node to gateway is calculated. Sink is responsible for overall operation and assign sensor nodes to their respective gateways. Min-heap is used on the number of allotted sensor nodes to the gateway. Root node of min heap is gateway and the nearest node to gateway is assigned first. Min-heap is adjusted so that minimum loaded gateway is at root. Comparison has been carried out with Gateway Load Balancing Protocol (GLBP) and Load Balancing Cluster (LBC) and EELBC shows better results than GLBP and LBC in terms of Energy consumption, Load balancing and execution time.

Centralized approaches are not well suited for a large scale network. Moreover, nodes near the central authority will deplete their energy faster as huge amount of data transfer takes place between nodes and sink. Unequal clustering should be preferred in this approach; size of clusters near the BS should be small so that energy can be conserved.

Other protocols include "A Regional Centralized Clustering Routing Algorithm for Wireless Sensor Networks" [7], "Multiple Sink Network design problem in large scale wireless sensor networks (MSNDP)" [8] and "Energy efficient Dynamic Clustering (EEDC)" [11].

B.2 Distributed Algorithms

Distributed algorithms aims at assigning each node in the network capability to select a CH based on several parameters. Nodes exchange information among

themselves and based on that information, the cluster head selection decision is taken. This decision can be based on several parameters like a weighted metric is used which may include residual energy, SNR or several other parameters. Nodes exchange these parameters and one having maximum value is selected as CH. Distributed approach is considered more efficient than centralized as amount of information transfers that takes place between sink and nodes are reduced and as nodes themselves take decision regarding clustering approach, scalability is enhanced.

Low Energy Adaptive Clustering Hierarchy (LEACH) [10] is the first clustering protocol developed for WSNs. It is a hierarchical protocol where communication is done at one hop and CH is responsible for data aggregation and routing process. Algorithm proceeds in two steps i.e. setup phase and steady phase. During setup phase, node decides whether to become a cluster head or not based on two conditions viz. how many times a node has been selected as CH and number of CHs. For CH selection each node chooses a random number between 0 and 1. A node is chosen as CH if number is less than the following threshold:

$$T(n) = \begin{cases} \frac{r}{1 - P \cdot \text{mod}(\frac{1}{P})} & \text{if } n \in G \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

P is the desired percentage of CHs, r is the current round, and G is the set of nodes that have not been elected CHs in the last 1/P rounds. After a node is elected a CH it broadcasts its status to other nodes. Nodes determine to which CH they should attach themselves based on distance criteria. CH waits for join request messages and after nodes join, it sends a TDMA schedule to Cluster members and data transmission takes place.

Energy efficient heterogeneous clustered scheme for wireless sensor networks is used for heterogeneous nodes where sensor nodes are divided into two categories, one having high energy and other with limited energy called normal nodes. As in LEACH based on probability function any node can become CH which will lead to faster depletion of energy, in EEHC nodes having more energy are selected as CH's. Decision is based on weight which is obtained by dividing initial energy of each node by initial energy of normal node. Weighted probabilities are found and based on that CH's are elected. Compared to LEACH it improves the lifetime and performance of the network [12].

Linked Cluster Algorithm (LCA) [13] developed for a wired network is also applicable for WSN. To acquire information about all nodes in the network, 2n time slots are required where n is number of nodes in the network. Nodes exchange information with each other and based on that information, CH is chosen if it has highest ID number or if node does not find any CH within its range. Communication is done using TDMA scheduling to avoid collisions. LCA was designed for small scale networks and delay was within acceptable range but if used for large networks it imposes more delay.

Power-efficient and adaptive clustering hierarchy protocol for wireless sensor networks (PEACH) is a distributed protocol that consists of homogeneous nodes having various transmission power levels. It supports both location aware and location unaware protocols for WSNs. Cluster formation process is different from other protocols as it does not include overhead of announcing CHs, joining clusters, scheduling intra-cluster communication [43]. Cluster set is chosen from node set and chosen CH waits to gather multiple packets from other nodes in cluster set. After, CH transmits the aggregated packet to next hop. Compared to other protocols it has improved energy consumption and better network lifetime.

Other examples of distributed protocols include an “Energy-efficient multi-level clustering algorithm for large-scale wireless sensor networks (EEMC)” [14], “Time controlled Clustering algorithm (TCCA)” [15], “Hybrid Energy Efficient Distributed Clustering” [16], “Distributed Weight-based Energy-efficient Hierarchical Clustering protocol (DWEHC)” [17], “Algorithm for cluster establishment (ACE)” [18].

B.3 Hybrid Algorithms

Hybrid algorithms use combination of distributed and centralized approaches. Some of the hybrid algorithms developed for clustering in WSN is discussed below:

Grid-clustering Routing Protocol for Wireless Sensor Networks (GROUP) is a hybrid protocol that includes grid based structure build dynamically and randomly by sink and CH selection process is done by small number of sensor nodes. This protocol is designed for query based applications, like forest monitoring. Queries generated by sink are forwarded to nodes via CH's and the node which has data matching to the query route back to sink via CH's. Protocol works in four phases, Cluster Grid construction, Query forwarding, Data forwarding and Failure recovery. In cluster grid construction one sink that is located in centre of Grid is selected as Primary Sink (PS). PS elects one node as Grid Seed (GS) by initiating the clustering process by sending GS-electing command within its range. All nodes which receive this command compare their distance with PS and if found appropriate send GS-reply packet to PS. PS then based on residual energy select one node as CH having maximum value of residual energy. GS is located at one crossing point of cluster grid structure and the node which is located near this point is elected as CH. CH election is initiated by GS by broadcasting CH-election packet to its neighbor and based on distance criteria one node is selected as CH. Query forwarding is done by broadcast and unicast communication. Location unaware queries are forwarded to the closest neighbor of sink which then broadcasts it and location aware queries are forwarded using unicast communication.

After receiving query, if it matches with sensor data it is forwarded to sink via CH's. If any CH fails which can impact the query and data forwarding process, sensor

node broadcast CH-search process to elect new CH. Compared with LEACH this protocol gives better results in terms of end-to-end delay, packet delivery ratio and energy consumption. This protocol works well for large scale wireless sensor networks [19].

Random Competition based Clustering (RCC) [20] was earlier developed for mobile ad-hoc networks but later on applied to WSN. This is based on First Declaration Wins rule in which node broadcast itself as CH. If it is the first node to broadcast, neighbor nodes join that node and declare it as their CH. But conflict arises when number of nodes broadcast at same time. To avoid this situation, a timer is employed or node having lower ID is chosen during conflict.

Energy Efficient Hybrid multi-hop Clustering algorithm in wireless sensor networks (EEHMCS) [21] follows centralized approach for all CH set-up decisions that are executed at the BS. Other decisions such as cluster formation relay node selection and data transmission are taken by sensors following distributed approach. If the BS lies within sensing range, then single hop communication is used, otherwise multi-hop communication is used to reach to the BS. After CH is selected, it broadcasts an advertisement message to all its neighbors and nodes join particular cluster based on the signal power received. A TDMA schedule is created for nodes and data is transmitted. This protocol uses both centralized and distributed approach which increases its efficiency. Compared with LEACH, A-LEACH, LEACH-C, EEHC and EEHMCS show better performance in terms of network lifetime.

Other protocols belonging to this category are “Hybrid Clustering Technique using Quantitative and Qualitative data for Wireless Sensor Networks” [24], “A Hybrid Cluster based target tracking protocol for Wireless Sensor Networks(HCTT)” [25], “HCMTT: Hybrid Clustering for Multi-Target Tracking in Wireless Sensor Networks” [42].

C. Algorithm Type

Algorithm type describes the method of clustering process. It can be probabilistic or non-probabilistic. In probabilistic, CH selection process takes place based on some probability value assigned to sensor nodes. Non-probabilistic approaches select CH based on the information received from other nodes. The details of some of algorithms based on algorithm type are discussed below:

C.1 Probabilistic (Random or Hybrid) Clustering Algorithms

In case of probabilistic algorithm, a probability value is assigned before clustering process takes place. Each node assigns its role by itself and based on some probability criterion a CH is selected. Other criteria like selecting a CH based on some metric like residual energy, communication cost is also considered in these methods. Some use combination of metrics known as combined

weight criteria for efficient selection of CH. These criteria lead us to faster execution time and reduced volume of exchanged messages. Probabilistic algorithms or protocols aim at achieving high energy efficiency and prolonged network lifetime. A survey of different probabilistic protocols is presented below:

Hybrid Energy-Efficient Distributed Clustering (HEED) is a clustering protocol that overcomes the drawbacks of LEACH. It considers residual energy of nodes for selection of CH which overcomes the possibility of selecting low energy node as CH. Along with residual energy it also considers intra cluster communication cost which result in even distribution of nodes in the network. Probability that two nearby nodes become a CH is very small in HEED. For choosing a CH probability is calculated as

$$CH_{\text{prob}} = C_{\text{prob}} * \frac{E_{\text{residual}}}{E_{\text{max}}} \quad (2)$$

E_{residual} is the current energy in the sensor, and E_{max} is the maximum energy, which corresponds to a fully charged battery. Value of CH_{prob} is not allowed to fall below a certain threshold that is selected to be inversely proportional to E_{max} . Cluster head selection process goes through number of iterations. First if node does not find a CH nearby it selects itself as CH. It remains in this phase until it finds low cost CH. Secondly if node probability become 1, it permanently become a CH. Nodes based on minimum distance join CH and data transmission from CH to base station takes place in a multi-hop fashion[26].

Improvement over HEED algorithm has been shown in [9]. This algorithm differs from HEED as in this case core algorithm is applied to eliminate some nodes. Numbers of iterations are performed and if node does not hear from tentative or final CH it volunteers to become CH with probability CH_{prob} . Node has a list of its neighboring tentative CHs, if node itself is least cost it become a CH, otherwise, it selects a CH based on minimum cost. Now the non-core nodes which do not become volunteers are excluded and then HEED is applied to rest of nodes. Final CHs are selected from participating nodes and remaining nodes become non CHs, these non CHs should have one neighboring CH in its range otherwise it becomes an uncovered node. Again algorithm is applied to these uncovered nodes to select additional CHs. This algorithm improves HEED by eliminating large number of nodes in CH selection process. Performance of this protocol is evaluated by checking performance of number of CHs vs. transmission range and number of rounds. This protocol shows better results than HEED in sparse networks where number of iterations is limited.

Energy-Efficient Hierarchical Clustering Algorithm (EEHCA) [27] uses a CH selection method different from others. All nodes receive hello message and based on signal strength of message, nodes calculate their distance to sink. Cluster radius is determined and based on that sink forms clusters, coordinates of nodes are received and their distance is evaluated. Two nearest nodes are

selected as Primary CH and secondary CH, after receiving a message from the sink node compare its ID with the ID in message, if matched they discover themselves as CH and send invitation messages within their own radius so that nodes within their range associate with that CH. TDMA schedule is determined by CH and data is sent using multi-hop communication. Compared to LEACH and HEED, EEHCA performs better in terms of network lifetime.

Some other examples of probabilistic algorithm are LEACH [10], "A Configurable Time-Controlled Clustering Algorithm for Wireless Sensor Networks" [44], and "Energy Residue Aware (ERA)", "Clustering Algorithm for Leach-based Wireless Sensor Networks" [28].

C.2 Non Probabilistic Clustering

In non-probabilistic approach, clustering process takes place by exchanging messages between nodes. Based on information received from other nodes and nodes connectivity, degree decision is taken. This approach requires more extensive exchange of messages. Combined metric approach is used as combination of metrics like combining residual energy, distance to sink and transmission power achieve more generalized solutions. Based on application, selection criteria are defined. Non Probabilistic Clustering can further be differentiated based on different approaches like a) Node Proximity and Graph-based clustering b) Weight based clustering c) Biologically inspired clustering approaches.

a) Node Proximity and Graph-based Clustering

Node proximity and graph based approach deals with building a tree, tree formation is done followed by cluster formation. Any node in a tree initiates a process by evaluating its distance to hop and informs other nodes by broadcasting a signal. Breadth First approach is followed during tree discovery phase. Neighbor node finds a shorter route through parent node having minimum number of hops to root. In graph based clustering network can then be modeled by a graph $G = (V, E)$ where each node in V represents a sensor and one node act as BS and E represents edge that links nodes via wireless means.

Highest Connectivity Cluster Algorithm (HCC) selects CH based on maximum number of one hop neighbors i.e. one having highest connectivity. Resources are assigned to CH which is shared among the members. Increased number of nodes leads to dropped throughput. The creation of one-hop cluster and clock synchronization requirement limits the practical usage of the algorithm [29].

Energy-Aware Clustering scheme with transmission power control for sensor networks (EACLE) [30] consists of three components; EACLE clustering, EACLE Routing and EACLE transmission timing control. In EACLE clustering, node waits for a prespecified time and if it does not hear from anyone it declares itself as CH.

Node sends two clustering packets with power high and power low. Power low signal for normal nodes and power high for participating nodes, after receiving signal it adds sender ID to its neighboring list. If both lists contain same ID they are already connected otherwise node waits for some time and become a CH. A Mesh tree is created for routing. Packet is broadcast from sink node, and route fix packets are used for routing which identifies next hop and the previous hop from where data is received. CHs are classified into relay nodes, end nodes and adjoin nodes. Relay node transfer data from source to destination. When a node receives route packet from a node not intended to be its source it changes its parent to the same node and end node is which does not receive any flooding packet during a specified time interval. Transmission timing control is achieved by using packet high for inter-cluster communication and packet low for intra-cluster communication. CH receives and binds the data packet and makes large packets. In round two, relay nodes send data only when it exceeds buffer length. A Graph-Center Based Scheme for Energy-Efficient Data Collection in Wireless Sensor Networks approach is based on concept of center of graph. Network is modeled as a graph and link (u, v) belongs to E only if u, v are in range of each other. Isolated sensors are not considered in this approach to maintain connectivity. Eccentricities and centers are calculated and node at center is chosen to be BS as this node is equidistant to all other nodes. CH's are selected that collects data from sensor nodes and transmit it to central node (BS). As network scales more number of BS can be added in the network. This approach requires more enhancement as it does not include issue of energy balancing and a simple communication model is used in this approach [48].

Some other algorithms belonging to this category are GROUP [19], "A Formal Graph-Based Model Applied to Cluster Communication in Wireless Sensor Networks" [49].

b) Weight based Clustering

These involve combination of metrics like residual energy, transmission range, SNR, etc. known as weighted metrics. Node having high cost is selected as CH. Combining these metrics improves performance and increases energy efficiency like selecting residual energy leads to a CH with more energy than member nodes and will not lead to depletion of energy. Weight based schemes increase throughput by electing better CH's in a proficient manner. Some weight based metrics use centralized and some use distributed approach. Some of the algorithms are discussed below:

Distributed Weight-based Energy-efficient Hierarchical Clustering protocol (DWEHC) is a distributed clustering algorithm similar to HEED. It also considers residual energy as in HEED but differs in cluster sizes and uses location awareness in intra cluster topology. Each node calculates its weight as follows:

$$W_{\text{weight}_s} = \frac{E_{\text{residual}_s}}{E_{\text{initial}_s}} * \sum_u \frac{R-d}{6R} \quad (3)$$

E_{residual_s} is residual energy in node s , E_{initial_s} is initial energy of all nodes, R is the cluster range and d is the distance from node s to neighboring node u . It forms a multi-level hierarchy. A node having highest weight becomes CH and all other nodes become member nodes. Member nodes operate at level one and communicate directly with CH. Nodes evaluate whether it reaches to CH in one hop or it can traverse multiple hops that can result in energy efficiency. If it can go through multi hops it become h-level member otherwise stays at 1 level. Member nodes should stay in particular range of CH and send their data to parent node which performs data aggregation and sending the data to CH using TDMA scheme. CH transmits data to BS in single hop [17]. Simulation results for HEED and GC (Generic Clustering based on residual energy) show better performance in terms of throughput and network lifetime [31].

Weight Based Clustering in Heterogeneous wireless sensor Networks (WBCHN) [45] protocol selects CH based on residual energy, number of alive neighbors of sensor node and its distance to BS. Residual energy is calculated based on average energy of nodes. If energy falls below certain threshold, CH selection process starts again. After end of each round, nodes send alive messages that determines number of alive neighbors. Compared to HEED and GC, WBCHN utilizes better heterogeneity in energy and has improved stability period.

In Distributed Weight-based Clustering Algorithm for WSNs, a sensor node selects itself as a cluster head based on the its' own resource, such as residual energy level, degree and local resource, such as neighbors' energy level, the sum of the distances between itself and each neighbor and the distance to the BS [47]. Based on received signal strength, node finds its distance to BS and finds neighboring nodes within its transmission range. Based on information received from other nodes, weighted metric of above parameters is computed and nodes with minimum weight are selected as CH. CDMA based communication is used in this protocol. Compared with LEACH and LEACH-C in terms of energy consumption during CH formation, this algorithm consumes more energy than LEACH but less energy than LEACH-C.

Some other weight based protocols include, "QoS Based and Energy Aware and Multi-path Hierarchical (QEMH)" [35], "Energy Aware QoS routing protocol (EAQoS)" [36], "Stable Election Protocol (SEP)", "Multi-weight Based Clustering Algorithm (MWBCA)" [46].

c) Biologically Inspired Clustering Approaches

These approaches are based on design paradigms developed from biologically inspired phenomena. Several biologically inspired approaches have been proposed.

One of them is Ant Colony optimization. In this approach ants deposit their pheromone when they search for food.

Table 3. Comparison of Various Clustering Techniques

Algorithm	Control manner	Energy Efficiency	Data Aggregation	Load Balancing	Scalability	CH election	CH Rotation
LEACH-C	Centralized	Low	No	Yes	Poor	Probabilistic	Yes
EELBC	Centralized	Medium	No	Yes	Poor	Probabilistic	Yes
EEDC	Centralized	High	Yes	Yes	Poor	Probabilistic	Yes
LEACH	Distributed	Low	No	No	No	Probabilistic	Yes
LCA	Distributed	Low	No	No	No	Id-based	Yes
EEMC	Distributed	High	Yes	No	Yes	Probabilistic	Yes
PEACH	Distributed	Medium	Yes	Yes	Yes	Probabilistic	Yes
DWEHC	Distributed	High	Yes	No	Yes	Weight based	Yes
ACE	Distributed	Low	No	Yes	Yes	Connectivity	No
EEHMC	Hybrid	High	No	Yes	No	Weight based	No
HEED	Distributed	Yes	Yes	Yes	Yes	Probabilistic	Yes
EEHCA	Distributed	High	Yes	Yes	Yes	Probabilistic	Yes
GROUP	Hybrid	Yes	Yes	No	Yes	Proximity	No
WBCHN	Distributed	Yes	No	No	Yes	Weight based	Yes
EACLE	Distributed	Medium	Yes	Yes	No	Connectivity	No
WBHCN	Distributed	High	Yes	No	Yes	Weight based	Yes
MWBCA	Distributed	Medium	No	No	Yes	Weight based	Yes

They deposit pheromone on their way and if they found optimal path they again deposit it on their way back.

Other ants follow strong scent of the pheromone and get food. Several techniques or algorithms have been developed based on this approach. Ant based control (ABC) was the traditional algorithm developed for telephone network and AntNet and Mobile Ants Based Routing (MABR) developed for wired networks.

Another approach is based on foraging process of honey bees. Dance bees share the distance, direction, and quality of a flower site with other foragers. Packers, scouts, foragers, and swarm agents are used. As bee stores food, similarly packers receive and store data packets. Scouts discover routes from source to destination. Foragers are recruits by scouts using metaphors dance. Then they receive data packets from packers and transmit them to destination Packers.

Swarm optimization techniques consist of Genetic Algorithms that belong to class of evolutionary algorithms (EA), which provide solutions to optimization problems using techniques inspired by natural evolution, such as inheritance, mutation, selection, and crossover. In [32] Genetic algorithm is applied to weighted clustering algorithm (WCA). In mapping, process information about CH and nodes is obtained from WCA. Then load distribution is done among CH and number of CHs is minimized. Using genetic operations, such as encoding, crossover, mutation, the one with the lowest fitness value is replaced as the best chromosome in the population. A biologically-inspired clustering protocol for wireless sensor networks [33], and a biologically inspired clustering algorithm dependent on spatial data in sensor networks has been proposed in [34]. Table 3 provides comparison of different clustering techniques based on performance metrics.

V. QUALITY OF SERVICE BASED CLUSTERING PROTOCOLS

Achieving QoS in WSN is more challenging due to limitations of WSN in term of energy consumption, resource constraints, unbalanced traffic, limited node energy etc. Due to different nature of WSN from other networks, QoS requirements in WSNs also differ from traditional networks. In traditional data network, QoS is defined by certain parameters such as packet loss, delay, jitter, bandwidth etc. However, the QoS requirements in WSNs such as data accuracy, aggregation delay, coverage, fault tolerance and network lifetime etc. are application specific and they are different from the traditional end-to-end QoS requirements due to the difference in application domains and network properties [22]. The QoS models used in internet and adhoc networks like Differentiated Services (DiffServ) and Integrated Services (IntServ) cannot be used in WSN due to their different nature. In case of WSN, QoS is provided by layered and cross layered approach. Layered approach deals with specific layer and cross layer requires simultaneous interaction of different layers and these approaches provide required QoS to users and applications but at the cost of reduced energy. QoS can be differentiated based on application and network, where application specific QoS depends on particular application and it varies depending upon application requirements.

Different application specific QoS parameters considered are system lifetime, response time, delay, data reliability etc. Network QoS specifies how underlying communication network can deliver QoS constrained data while efficiently utilizing network resources. Basically network QoS deals with the delivery of data and there are different data delivery models i.e. Continuous, Query Driven and Event Driven. Depending upon the

application requirements, QoS parameters are chosen and network is designed such that it satisfies QoS requirements.

Quality of Service is mainly required in real time applications and these applications employ several challenges at network and MAC layer, they consist of real and non-real time data so traffic differentiation is very important to achieve better performance. In some cases real time data utilize most of the bandwidth leaving non real time with a very few resources. So efficient utilization of bandwidth should be provided to both types of data with priority given to real time data. Depending upon the application, criticality of packet is identified and priority is assigned so that QoS mechanisms should differentiate packet importance and set up priority structure. In [23] authors proposed a protocol in which least cost delay constrained path is found for real time traffic and throughput is maximized for non-real time traffic. A Queuing model is employed which separates real and non-real time, and performance is evaluated based on end to end delay. Simulation results show that protocol performs well with respect to average lifetime of node, average delay per packet and network throughput. But this protocol simulation does not include comparison with any other protocol. Queuing model is employed in many protocols in literature but if we differentiate this traffic by using queuing model and traffic is sent on single path that will lead to huge delay and decrease the performance. For that reason multi-path routing is the most efficient method. In [39], authors proposed QoS and Energy aware multi-path routing algorithm for real-time applications in wireless sensor networks (QEMPAR). It is a QoS aware protocol that increases the lifetime of network by considering energy consumption as an important parameter. Node disjoint paths are found and packets are transmitted over these paths by breaking them into smaller parts and attaching a sequence. Protocol shows better performance in terms of end to end delay and average energy consumption but does not consider traffic differentiation. Multimedia applications have different delay requirements. Hard real time delay guarantees become difficult to achieve because of the limited resource and energy issues of WSN. Multimedia data requires high bandwidth and low delay and these applications require predefined QoS to achieve required performance.

Another multi-path routing algorithm for improving QoS in wireless sensor network Clustering based Multipath routing algorithm to improving QoS in wireless sensor networks (CMQ) [37]. CMQ is a multi-path clustering algorithm which provides better energy consumption and required QoS in terms of minimum end to end delay and maximizing packet delivery ratio. As discussed in above sections the advantages of clustering and incorporating clustering approaches in QoS based routing can enhance scalability, improve network performance and help in load balancing. Some of the QoS based routing protocols that incorporate clustering are discussed below:

Energy Aware and Multi-path Hierarchical protocol (QEMH) is energy aware clustering based protocol that considers residual energy and node's distance to sink for choosing cluster heads. Current Energy (CE) is obtained by following equation

$$CE = \frac{(E_{node-res_j})}{(1 - (\frac{dis(j)}{100})^2)} \quad (4)$$

$(E_{node-res_j})$ is the current residual energy of node j and distance (j) is calculated as

$$Dis(j) = (\sum(|D_{db}(j) - D_{db}(i)|) * t_p * k) \quad (5)$$

Here, D_{db} is node distance to sink. We also assume that number of bits, $k = 1$ and Transmission power, $t_p = 1$. Node having maximum value of and parameters is chosen as cluster head. Based on link cost metric which is a combination of energy factor, available buffer and link performance factor, cluster head chooses next hop and node disjoint paths are chosen that increases robustness, throughput and balance load among nodes. Queuing model is used that differentiate real and non-real time traffic with r paths chosen for real time and n for non-real time. Queuing model is presented in Fig.2.

Alternative paths are found in this way and path having lower delay is chosen for real time traffic to avoid delay. To improve reliability correction codes are added using XoR based correction method.

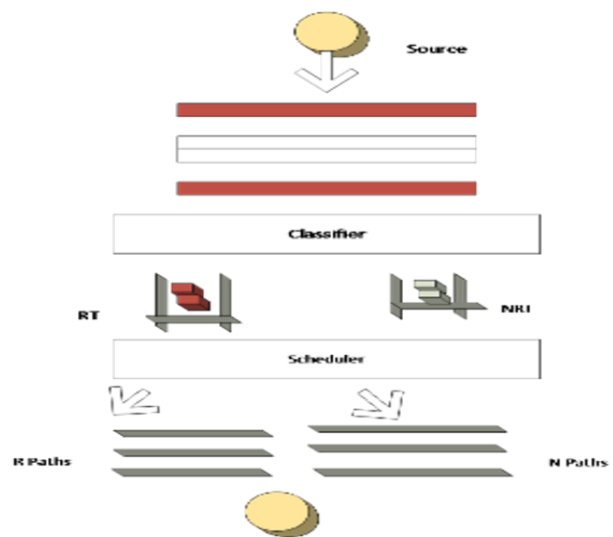


Fig.2. Queuing model of QEMH [35]

Comparison is done with multi constrained QoS multi-path routing (MCMP)[40] and Energy-Aware Routing Protocol (EAP) [41] protocols and it achieves more energy saving, low average end to end delay, more lifetime and high packet delivery ratio [35].

QoS Enhanced Base Station Controlled Dynamic Clustering Protocol (QBCDCP) consists of homogeneous nodes and achieves QoS by including delay and bandwidth details while selecting a route. Considering some limitations of Leach protocol enhancements has been made to the CH selection threshold equation of LEACH. It increases threshold for some node that has not been selected as CH for previous rounds. Based on following equation decision is taken

$$T(n)_{new} = \frac{P}{1 - P(\text{rmod} \frac{1}{P})} \left[\frac{E_{n_current}}{E_{n_max}} + \left(r_s \text{div} \frac{1}{P} \right) \left(1 - \frac{E_{n_current}}{E_{n_max}} \right) \right] \quad (6)$$

Where, $E_{n_current}$ is the current energy, maximum energy is given by E_{n_max} , r_s represents rounds in which node has not been selected as CH. TDMA is used as MAC protocol where one slot is fixed for control messages and other slots segmented to send and receive messages. Base Station sets up a QoS based route Q between itself and a CH based on a metric that includes bandwidth delay and transmission energy. Left out energy is considered by BS while selecting CH so that CHs never deplete their energy. Slot assignment or Communication scheduling is carried out after selecting QoS route by BS. Whole process takes place from base station [38]

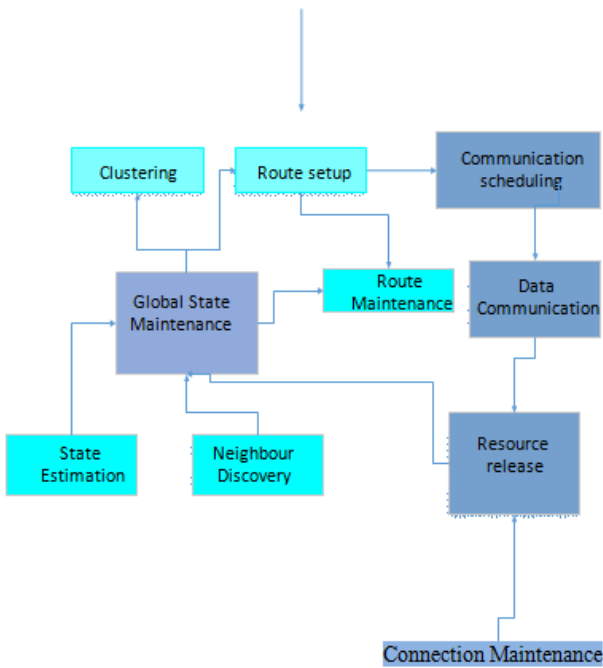


Fig.3. QoS Enhance Base Station Controlled Dynamic Clustering Protocol [38]

Energy Aware QoS routing protocol (EAQoS) is QoS aware protocol for real time traffic that finds a least cost delay constrained path for real time traffic and maximizes throughput for non-real time traffic. A queuing model is employed which separates real and non-real time traffic. Gateway allocates bandwidth ratio (r value) to both type

of traffic and service rate is calculated. Candidate paths are found and using Dijkstra algorithm, least cost paths are found. Cost function is calculated for each link which includes communication cost ($c_0 * (\text{dist}_{ij})^l$), Energy stock ($c_1 * f(\text{energy}_{ij})$), energy consumption rate ($\frac{c_2}{T_j}$), Sensing cost (c_4), Relaying cost (c_3), Maximum connections per delay (c_5), error rate ($c_6 * f(e_{ij})$).

$$\text{cost}_{ij} = \sum_{k=0}^6 CF_k = c_0 * (\text{dist}_{ij})^l + c_1 * f(\text{energy}_{ij}) + \frac{c_2}{T_j} + c_3 + c_4 + c_5 + c_6 * f(e_{ij}) \quad (7)$$

Based on above function, cost is found and then Dijkstra algorithm is run to find least cost paths. Bandwidth (r) values are calculated for found paths and if value is not between 0 and 1 then extended Dijkstra are run to find more paths.

If no r value is found, connection of the node to gateway is simply rejected. Performance of the protocol is calculated based on end to end delay that is sum of queuing and propagation delay. Simulation results show that protocol performs well with respect to average lifetime of node, average delay per packet and network throughput. Simulation does not include comparison with any other protocol [36].

Clustering based Multipath routing algorithm to improving QoS in wireless sensor network (CMQ) [37] is multi-path clustering algorithm which provides better energy consumption and required QoS in terms of minimum end to end delay and maximizing packet delivery ratio. It considers all nodes to have equal energy and can calculate packet sending and receiving probability. Protocol has three phases viz. cluster head announcement, cluster formation and data transmission.

Cluster head announcement phase process is similar to HEED but differs when nodes calculate their probability of CH election by:

$$CV_{CH} = \alpha \left(\frac{E_r}{E_i} \right) + \beta \left(\frac{N_{elt}}{N_n} \right) + \gamma (PSPS) \quad (8)$$

Where a , b and g are influence coefficients of energy, E_r is remaining energy of sensor node and E_i is initial energy of sensor. N_{elt} defines neighboring nodes with remaining energy less than threshold and N_n signifies number of nodes in the network. PSPS is the probability of successful packet sending of sensor node. PSPR is the probability of successful packet receiving of all neighbors of node. After CH's is announced using equation (9) and all nodes select CH which has highest value for the given equation

$$S_{CH} = \frac{E_{CH}}{(D_{n,CH})^2} \quad (9)$$

To setup a path, next hop is found using link suitability as a function of remaining energy, link quality between two nodes and SNR. Using these parameters next hop is found, node disjoint paths are constructed and traffic is

sent over these paths. Compared to Energy Efficient and QoS aware multipath routing protocol (EQSR) [36] it gives better results in terms of end to end delay, average energy consumption and average packet delivery ratio.

QoS based routing protocols developed so far are mainly location based; a little work has been done on clustering based QoS protocols. As we have discussed

clustering improves scalability, reduces redundancy; incorporating efficient clustering method will enhance the performance of QoS based routing protocols.

Comparison of these protocols based on several QoS parameters viz. energy efficiency, Clustering process, CH selection, QoS support, Reliability, Multi-path, Timeliness presented in Table 4.

Table 4. Comparison of QoS based Clustering Protocols

Protocol	Energy efficiency	Clustering Process	CH election	QoS Support	Reliability	Multi-path	Timeliness
QEMH[35]	High	Distributed	Weight based	Yes	Yes	Yes	Yes
QBCDCP[38]	Medium	Centralized	Weight based	Yes	No	No	Yes
EAQoS[36]	Medium	Distributed	Weight based	Yes	Yes	Yes	Yes
CMQ[37]	High	Distributed	Probabilistic	Yes	Yes	Yes	Yes

VI. CONCLUSION

To achieve scalability, increase network lifetime and improve energy efficiency, clustering is an important mechanism to follow in WSN. Significant research has been carried out to develop efficient clustering algorithms and protocols and work is still going on to increase the efficiency of clustering. In this paper we have presented the existing approaches of clustering process and discussed various algorithms associated with each category. These algorithms have been compared based on some chosen metrics. Finally Quality of Service based Clustering Protocols have been discussed and comparison is presented.

REFERENCES

- [1] M. Younis, K. Akkaya, "Strategies and techniques for node placement in wireless sensor networks: A survey", *Ad Hoc Networks* 6 (4), pp.621-655, 2008.
- [2] R. Mulligan, H.M. Ammari, "Coverage in Wireless Sensor Networks: A Survey", *Network Protocols and Algorithms*, ISSN 1943-3581, 2010.
- [3] J. Hill, M. Horton, R. Kling, L. Krishnamurthy, "The platforms enabling wireless sensor networks", *Communications of the ACM*, pp. 41 – 46, 2004.
- [4] H.Jadidoleslami, "An introduction to various basic concepts of clustering techniques on wireless sensor networks", *International journal of Mobile Network Communications & Telematics (IJMNCT)*, Vol. 3, February 2013.
- [5] R.M.B. Hani and A.A. Ijeh, "A survey on leach based energy aware protocols for wireless sensor networks", *Journal of communication*, vol. 8, 2013.
- [6] P. Kula, P. K. Jana, "Energy Efficient Load-Balanced Clustering Algorithm for Wireless Sensor Networks", *ICCCS-2012*, pp-771-777, 2012.
- [7] T. Gao, R. Jin, "A regional centralized clustering routing algorithm for wireless sensor networks", *IEEE*, 2008.
- [8] E.I. Oyman and C. Ersoy, "Multiple sink network design problem in large scale wireless sensor networks", in *Proceedings of the IEEE International Conference on Communications (ICC 2004)*, Paris, June 2004.
- [9] H. Huang, J.Whu, "A probabilistic clustering algorithm in wireless sensor networks", in: *Proceeding of IEEE 62nd Semi-annual Vehicular Technology Conference (VTC)*, September 2005.
- [10] W.Heinzelman, A.Chandrakasan, and H. Balakrishnan, "Energy-Efficient Communication Protocols for Wireless Microsensor Networks", *Proceedings of the 33rd Hawaiaian International Conference on Systems Science (HICSS)*, January 2000.
- [11] K. Padmanabhan, P. Kamalakkannan, "Energy-efficient Dynamic Clustering Protocol for Wireless Sensor Networks", *International Journal of Computer Applications*, Volume 38– No.11, January 2012.
- [12] D.Kumar, T.C. Aseri, R.B.Patel, "EEHC: Energy efficient heterogeneous clustered scheme for wireless sensor networks", *Computer Communications*, Vol.32, pp-662-667, 2009.
- [13] D.J. Baker and A. Ephremides, "The architectural organization of a mobile radio network via a distributed algorithm", *IEEE Transactions on Communications*, pp.1694–1701, 1981.
- [14] Y. Jin, L. Wang, Y. Kim, X. Yang, "EEMC: An energy-efficient multi-level clustering algorithm for large-scale wireless sensor networks", pp.542-562, 2008.
- [15] S. Selvakennedy and S. Sinnapan, "The Time-Controlled Clustering Algorithm for Optimized Data Dissemination in Wireless Sensor Networks", *IEEE Conference on Local Computer Networks*, pp.509-510, 2005.
- [16] O. Younis, S. Fahmy, "HEED: Hybrid Energy Efficient Distributed Clustering Approach for Ad hoc Sensor Networks", *IEEE Transactions on Mobile Computing*, 2004.
- [17] P.Ding, J. Holliday, A.Celik, "Distributed Energy-Efficient Hierarchical Clustering for Wireless Sensor Networks", *Distributed Computing in Sensor Systems*, 2005.
- [18] H.Chan, A. Perrig, "ACE: An emergent algorithm for highly uniform cluster formation", *European Workshop on Sensor networks (EWSN)*, January 2004.
- [19] L.Yu, N. Wang, W.Zhang, "GROUP: A Grid-Clustering Routing Protocol for Wireless Sensor Networks", *WiCOM*, pp. 1-5, 2006.
- [20] K. Xu and M.Gerla, "A heterogeneous routing protocol based on a new stable clustering Scheme", in *Proceeding of IEEE Military Communications Conference (MILCOM2002)*, Anaheim, CA, October 2002.

- [21] A. Patra, S. Chouhan, "Energy Efficient Hybrid multihop clustering algorithm in wireless sensor networks", *COMNETSAT*, pp.59-63, 2013.
- [22] Bhaskar Bhuyan, Hiren Kumar Deva Sarma, Nityananda Sarma, Avijit Kar, Rajib Mall, "Quality of Service (QoS) Provisions in Wireless Sensor Networks and Related Challenges", *Wireless Sensor Network Journal*, pp. 861-868, 2012.
- [23] J. Othman, B. Yahya, "Energy efficient and QoS based routing protocol for wireless sensor networks", *Journal of Parallel and Distributed Computing*, pp. 849-857, August, 2010.
- [24] J. Zhu, C.H.Lung, V. Srivastava, "Hybrid Clustering Technique using Quantitative and Qualitative data for Wireless Sensor Networks", *Elsevier, Ad-hoc networks*, volume 25, pp-38-53, 2015
- [25] Z.Wang, W. Lou, Z. Wang, J. Ma, and H. Chen "A Hybrid Cluster based target tracking protocol for Wireless Sensor Networks (HCTT)", *International Journal of Distributed Sensor Networks*, Article id 494863, 2013.
- [26] Ossama Younis, Sonia Fahmy, "HEED: A Hybrid, Energy-Efficient, Distributed Clustering Approach for Ad Hoc Sensor Networks", *IEEE Transactions on Mobile Computing*, vol.3, no. 4, pp. 366-379, October 2004.
- [27] H. Huang, J.Wu, "A Probabilistic Clustering Algorithm in Wireless Sensor Networks", *IEEE*, 2005.
- [28] S. Bandyopadhyay and E. J. Coyle, "An Energy Efficient Hierarchical Clustering Algorithm for Wireless Sensor Networks", *INFOCOM*, 2003.
- [29] S. Banerjee and S. Khuller, "A clustering scheme for hierarchical control in multi-hop wireless networks", in *Proceedings of 20th Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'01)*, Anchorage, AK, April 2001.
- [30] Z. Yanagihara, J. Taketsugu, "EACLE: Energy-Aware Clustering Scheme with Transmission Power Control for Sensor Networks", *Wireless Personal Communications*, pp. 401-415, 2007.
- [31] R. Tandon, B. Dey and S. Nandi, "Weight based clustering in wireless sensor networks", *Communications (NCC)*, 2013.
- [32] D. Turgut, S. K. Das, R. Elmasri, and B. Turgut, "Optimizing clustering algorithm in mobile ad hoc networks using genetic algorithmic approach," *IEEE Global Telecommunications Conference*, pp. 62-66, 2002.
- [33] S. Selvakennedy, S. Sinnappan, "A biologically-inspired clustering protocol for wireless sensor networks", *Journal Computer Communications*, pp. 2786-2801, 2007.
- [34] I. Wokoma, L. L.Shum, L.Sacks, I.Marshall, "A Biologically-Inspired Clustering Algorithm Dependent on Spatial Data in Sensor Networks", *proceedings of the Second European Workshop on Wireless Sensor Networks*, 2005.
- [35] M. Mazaheri, B. Homayounfar and S. Mazinani, "QoS Based and Energy Aware Multi-Path Hierarchical Routing Algorithm in WSNs", *Wireless Sensor Network*, pp. 31-39, 2012.
- [36] J. Othman, B. Yahya, "Energy efficient and QoS based routing protocol for wireless sensor networks", *Journal of Parallel and Distributed Computing*, pp.849-857, August, 2010.
- [37] A. R. Rezaie and M. Mirnia, "CMQ: Clustering based Multipath routing algorithm to improving QoS in wireless sensor networks", *IJCSI International Journal of Computer Science*, Vol. 9, Issue 3, No 1, May 2012.
- [38] A.O. Fapojuwo, A. C.Tinoco, "Energy Consumption and Message Delay Analysis of QoS Enhanced Base Station Controlled Dynamic Clustering Protocol for Wireless Sensor Networks", *IEEE transactions on wireless communications*, 2009.
- [39] S. Prakash T, K.B. Raja, "Real-Time Link Reliability Routing for QoS in Wireless Sensor Networks", pp. 05-16, 2014.
- [40] X. Huang and Y. Fang, "Multiconstrained QoS Mutlipath Routing in Wireless Sensor Networks", *Wireless Networks*, pp. 465-478,2008.
- [41] M. Liu, J.-N. Cao, G.-H. Chen and X.-M. Wang, "An Energy-Aware Routing Protocol in Wireless Sensor Networks," *Sensors*, pp. 445-462, 2009.
- [42] Faezeh Hajiaghajani, Marjan Naderan, Hossein Pedram, Mehdi Dehgha, "HCMTT: Hybrid Clustering for Multi-Target Tracking in Wireless Sensor Networks", 4th *International Workshop on Sensor Networks and Ambient Intelligence*, 2012.
- [43] Xuxun Liu, "A Survey on Clustering Routing Protocols in Wireless Sensor Networks", *Sensors*, 2012.
- [44] K. Akkaya and M. Younis, "A Survey on Routing Protocols for Wireless Sensor Networks", *Ad Hoc Networks*, pp. 325-349, 2005.
- [45] D. Goyal, M.R. Tripathy, "Routing protocols in wireless sensor networks: A Survey", *Advanced Computing & Communication Technologies (ACCT)*, pp.474 - 480, 2012.
- [46] Zhang/RFID and Sensor Networks, Chapter 12, "Clustering in Wireless Sensor Networks", pp.335, 2009.
- [47] Luis Javier, Garc ía Villalba, Ana Lucila Sandoval Orozco, Alicia Triviño Cabrera and Cláudia Jacy Barenco Abbas, "Routing Protocols in Wireless Sensor Network", *Sensors*, 2009.
- [48] A. Joshi, M. Priya Lakshmi, "A Survey of Hierarchical Routing Protocols in Wireless Sensor Network", 2011.
- [49] Boyinbode, O.; Le, H.; Mbogho, A.; Takizawa, M; Poliah, R., "A Survey on Clustering Algorithms for Wireless Sensor Networks", In *Proceedings of 2010 13th International Conference on Network-Based Information Systems*, Takayama, pp. 358-364,2010.
- [50] www.6lowpan.org/
- [51] <http://www.zigbee.org/>
- [52] <http://www.z-wave.com/>
- [53] <https://www.enocean.com/en/home>

Authors' Profiles



Sukhkirandeep Kaur is a research scholar in the Department of Computer Science & Engineering at NIT Srinagar, INDIA. She received her B.Tech in Computer Science & Engineering from Punjabi University, Patiala(India) in 2010 and M.Tech in Computer Science & Engineering from Lovely Professional University, Jalandhar(India) in 2012. Her research interests include Routing and Clustering in Wireless Sensor Networks.



Roohie N Mir is a professor in the Department of Computer Science & Engineering at NIT Srinagar, INDIA. She received B.E. (Hons) in Electrical Engineering from University of Kashmir (India) in 1985, M.E. in Computer Science & Engineering from IISc Bangalore (India) in

1990 and Ph D from University of Kashmir, (India) in 2005. She is a Fellow of IEI and IETE India, senior member of IEEE and a member of IACSIT and IAENG. She is the author of many scientific publications in international journals and

conferences. Her current research interests include reconfigurable computing and architecture, mobile and pervasive computing, security and routing in wireless adhoc and sensor networks.

How to cite this paper: Sukhkirandeep Kaur, Roohie Naaz Mir, "Clustering in Wireless Sensor Networks- A Survey", International Journal of Computer Network and Information Security(IJCNIS), Vol.8, No.6, pp.38-51, 2016.DOI: 10.5815/ijcnis.2016.06.05