Node Clustering for Wireless Sensor Networks

SANIA BHATTI*, IMRAN ALI QURESHI**, AND SHEERAZ MEMON***

RECEIVED ON 21.09.2011 ACCEPTED ON 01.12.2011

ABSTRACT

Recent years have witnessed considerable growth in the development and deployment of clustering methods which are not only used to maintain network resources but also increases the reliability of the WSNs (Wireless Sensor Network) and the facts manifest by the wide range of clustering solutions. Node clustering by selecting key parameters to tackle the dynamic behaviour of resource constraint WSN is a challenging issue. This paper highlights the recent progress which has been carried out pertaining to the development of clustering solutions for the WSNs. The paper presents classification of node clustering methods and their comparison based on the objectives, clustering criteria and methodology. In addition, the potential open issues which need to be considered for future work are high lighted.

Key Words: **Clustering, Sensor Network, Static, Dynamic.**

1. INTRODUCTION

The recent advancements in MEMS (Micro-Electro-Mechanical Systems) technology, have made it possible that WSNs are able to fulfil the demand of low cost unsupervised applications. A WSN typically consists of SNs (Sensor Nodes) from few tens to thousands working together to collect data about an environment from time to time and then forwarding the data to a BS (Base Station). Each SN of the network consists of four main components: an array of sensors for obtaining information about the observed area, wireless transceivers, a processor for performing calculations and network protocol related functions and power supply. Energy is a precious resource for WSNs because SNs are expected to function until energy depletion. Therefore WSN protocols must ensure energy efficiency during every state of SNs, as every state consumes different amount of energy [1]. One way to preserve energy is to increase sleeping time of SNs in the absence of an event or assign node duty cycle. In node duty cycle SNs periodically turns on and turns off their radios. Ideally, the SNs should switch between active and sleep modes depending on the network activity to conserve energy [2]. Another approach is to organise a WSN into a set of interconnected clusters, thereby achieving better scalability, energy efficiency and resource allocation. Node clustering and data aggregation are able to decrease communication overheads for both one-hop and multi-hop communications in WSNs.

The node clustering arranges SNs to clusters and elects CH (Cluster Head) which satisfies the following constraints:

(i) The CHs allow the member clusters to communicate with it directly.

* Assistant Professor, Department of Software Engineering, Mehran University of Engineering & Technology, Jamshoro.
** Ph.D. Scholar, School of Information & Electronics, Beijing Institute of Technology, Beijing,P.R China.
*** Assistant Professor, Department of Computer Systems Engineering, Mehran University of Engineering & Technology, Jamshoro.

(ii) The CHs can send aggregated data to BS using one-hop or multi-hop communication links [2]. This also decreases the bandwidth requirement and provides scalability and robustness for the network.

Node clustering is a nontrivial technique for managing scarce network resources. Two significant benefits could be achieved through clustering sensor nodes: first, energy consumption by the network is reduced and secondly, communication complexity is considerably decreased [3]. The most widely assumed model of a cluster-based WSN is depicted in Fig. 1 [4]. The objectives of this paper are:

- To provide an improved study of the current state of art and practice, covering and focusing on the latest development and innovative aspects in the area.
- To produce classification of node clustering protocols based on either clustering process or CH selection process.
- To highlight research and technical challenges faced by node clustering protocols.

Several recent papers have presented node clustering in WSNs [7-32] and these studies showed that clustering

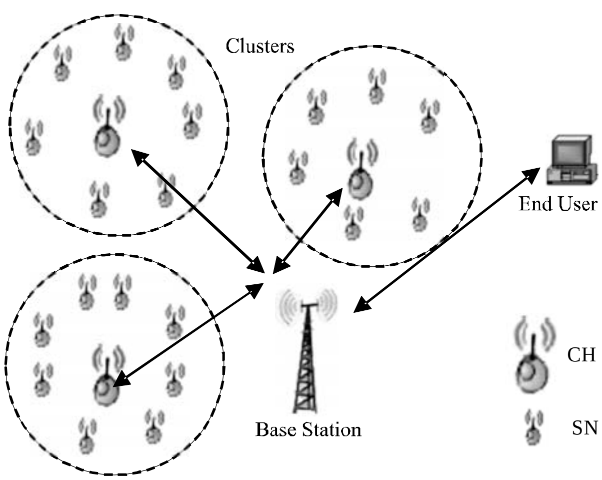

FIGURE 1 CLUSTER-BASED SENSOR NETWORK ARCHITECTURe

can significantly improve WSN performance. Thus this paper presents the current state of research over the existing node clustering surveys [2-5]. The paper is arranged in the following manner: Section 2 provides a classification of recent clustering schemes; Section 3 draws comparisons of these clustering schemes in tabular form and Section 4 states open issues and finally Section 5 presents conclusions.

2. CLASSIFICATION OF NODE CLUSTERING SCHEMES

Recently, a lot of research has been dedicated to node clustering protocols for WSN. Research on node clustering in WSN has focused on developing static/ dynamic and centralized/distributed algorithms. Static clustering has a few drawbacks regardless of its simplicity, for example, static membership is not robust from fault-tolerance point of view and it prevents SNs in different clusters from sharing information. In contrast, clusters are formed dynamically in dynamic clustering depending on the occurrence of certain events, for instance, as a SN with enough battery and computational power detects an event, it comes forward to act as a CH. The CH invites nearby SNs and makes them members of its cluster. Since SNs do not statistically form a cluster, they may belong to altered clusters at diverse timings. As only one cluster is active at a time, redundant data and interference is reduced [6]. Another way of categorization is the consideration of implementation method. The clustering algorithm can be executed by a central authority or in a distributed manner at local SNs. The drawback of centralized approaches is the requirement of global information of network decision making parameters which results in communication overhead. Because of this reason, distributed algorithms are appropriate for large scale networks. Generally, in these algorithms, SNs make decisions to join a cluster or become a CH based on its one-hop neighbour information.

Existing research efforts on node clustering can be categorized in two ways [2]:

(i) The parameter(s) used for electing CHs, for example, SN identifier, SN degree, residual battery energy, average distance between neighbours, etc.

(ii) The implementation way of a clustering algorithm or cluster formation procedure, which can be further divided into iterative and probabilistic clustering protocols.

Here we are considering the classification in terms of implementation way of the clustering approaches. It is critical to select CH during cluster formation procedure. This might involve a single metric or multiple parameters from neighbour SNs. This classification is considered to address the challenges faced by large scale WSNs.

2.1 Iterative Clustering Protocols

Iterative clustering enables a SN to either wait for a particular event to occur or messages from particular SNs to choose their role before making a decision. Specifically, SNs involve information from their one hop neighbour during clustering. Some of its examples are discussed below:

(1) *GESC:* The GESC (GEodegic Sensor Clustering) [7] protocol is dynamic in nature because each SN can employ its one-hop or two-hop neighbourhood while clustering. The cluster formation is based on a novel metric, i.e., computation of SN implication with regard to the number of messages passing through the SN. The network clustering is fast due to linearity in the amount of SNs and linearity in the amount of edges of the network neighbourhood in finding SN implication.

(2) *Static Clustering Based on One-Hop Distance:* In [8] optimal one-hop distance is used to form static clusters. The one-hop distance is worked out using device electronics instead of WSN topology to facilitate energy consumption reduction. The clustering algorithm in optimal one-hop communication is combined with the LEACH (Low Energy Adaptive Clustering Hierarchy) [9] algorithm, in which clustering is broken into rounds. At first, a CH in each cluster is selected at the border line, which works as local in-charge of its members and sends aggregated data to the BS or upper layer CH. The current CH would decide the candidate CH depending on the received signal strength. Conceptually, a layered architecture is used in which each layer is a cluster with a CH; responsible to send aggregated data to the upper layer CH or the BS with one-hop distance. Cluster size is controlled with transverse range, so that few clusters are created. However, optimal one-hop distance and transversal range are not clearly defined.

(3) *EECT:* An EECT (Energy-Efficient Clustering Technique) based on a virtual hexagon cluster formation is proposed in [10]. Based on the distance among common SNs and the virtual hexagon centre, sub-circles are formed inside virtual hexagon. The sink manages the virtual hexagon's centre location information and the optimal cluster radius and broadcasts this information to all the SNs. The SNs join the nearest virtual hexagon cluster after determining their distance with all the virtual hexagons. EECT circumvents the repeated CH election. It lacks experimental verification of the optimal number of CHs and the optimal cluster radius.

(4) *CMATO:* A distributed fault-tolerant mechanism called CMATO (Cluster Member Based Fault Tolerant Mechanism) is presented in [11]. According to authors, the CMATO is capable to recover from faults in a cluster-based arrangement by overhearing the transmissions of neighbour CHs. It is flexible to work with existing clustering algorithms. In CMATO, local fault detection and

fault recovery procedures are adapted to recover SNs after the failures of various CHs and the failures of communication links within the cluster. CMATO is shown to detect faults by implementing it over LEACH and HEED (Hybrid Energy Efficient Distributed Clustering) clustering algorithms, via J-sim simulations. However, discussion about the communication overhead due to inclusion of fault recovery procedure in LEACH and HEED is missing.

(5) *Clustering by Soude and Mehat:* The clustering algorithm presented in [12] uses a notification protocol to facilitate WSN topology discovery and an algorithm to split the WSN into clusters. Clusters are formed via cut algorithm and a spy method is used to decrease the number of messages during clustering. The notification protocol works in two steps: in first step SN inform their presence to the BS and in second step BS assign them the cluster using Set cluster message. The proposed algorithm comes under the iterative clustering protocols because the BS involves the discovery of neighbours during notification protocol. The proposed algorithm takes a graph as parameter which can represent size of the cluster, cluster depth, energy level and sensor position.

(6) *Clustering based on KSOM-NN:* An interesting study in which clustering is based on the KSOM-NN (Kohonen Self Organization Map Neural Network) is discussed in [13]. It defines how cluster behaviour can be studied with respect to the parameters related to the specific application so as to improve the operational efficiency of the network. The KSOM-NN has self-organizing properties and is an unsupervised training neural network. This protocol uses computational intelligence to form clusters using different parameters in a WSN. In this clustering, parameters of each SN such as memory available, number of hopes from SN to CH and energy levels could be collected for decision making at the BS. The results show the number of clusters formed with the avoidance of energy consumption.

(7) *ACAWT*: ACAWT (Adaptive Clustering Algorithm via Waiting Timer) [14] presents centralized and distributed models for choosing a new CH for an existing cluster. It is useful for situations where both centralized and distributed models are required. A random waiting timer and neighbour information are used for CH reselection process. The three phases of ACAWT are clustering, reselecting a CH and restructuring the clusters. The clustering phase involves CAWT. During CH reselection, in centralized model a new CH is decided based on the energy and neighbour information from its cluster members and in distributed model sub-clusters are formed by applying ACAWT. To clarify the performance of clustering algorithm, modified average model and energy consumption models are presented. The weakness of ACAWT is the avoidance of failure scenarios.

(8) *UCCP:* The authors in [15] employ a UCCP (Multi Criterion Optimization algorithm) to satisfy multiple conditions simultaneously. A uniform cross layer design is adopted to maintain quality of service guarantees. UCCP (Unified Clustering and Communication Protocol) work in rounds and every round consists of two parts namely, a self organization phase and a data transmission phase. In self organization phase, cluster topology is managed through communication between SNs. In data transmission phase, a reservation based TDMA (Time Division Multiple Access) is followed in order to avoid packet collision, idle listening and contention free

transmission slots. However, practical applicability of TDMA requires synchronization between the SNs.

(9) *MCLB:* For balancing load and energy efficiency, MCLB (Multi-Hop Clustering algorithm for Load Balancing) is proposed in [16]. MCLB works in set up phase and steady phase similar to LEACH. In set up phase, first temporary clusters are formed based on coverage area. Then two layers of multi-hop communication are made. The top layer comprises of temporary CHs and bottom layer consists of SNs. In steady phase data is transmitted to the CH through intracluster communication and to the BS through intercluster communication. Simulation comparison of MCLB with LEACH proves its energy efficiency and increased network life time; however, the exact figures related with energy efficiency are not discussed. In addition, communication overhead incurred in both phases is not computed.

(10) EEPSC: A static clustering protocol named EEPSC (Energy Efficient Protocol with Static Clustering) is proposed in [17] which removes the overhead of dynamic clustering by forming the cluster only once during the network operation. It is a modified description of LEACH which employs temporary CHs with a new setup, responsible node selection and steady state phase. Simulations show the better performance of EEPSC than LEACH regarding amount of data messages reached at the BS and network life time; however, energy consumption comparison is not discussed, which is one of the main concerns.

(11) *ADCA:* An ADCA (Adaptive Distributed Clustering Algorithm) for energy consumption reduction is presented in [18]. The proposed protocol works in two phases namely, cluster formation and adaptive sleep duty cycle phase. The clustering is based on data generation rate and the similarity between data series. In sleep duty cycle phase, after comparison of sending rates of nodes with a threshold level, a sleep duty cycle is chosen for a fixed period of time based on their rates. After collecting the data from its members, if the CHs notify major change in data rate, it will report this to the sink along with the data. The sink then carries out re clustering after analyzing data. This clustering scheme is encountered as iterative clustering scheme because sink node receives data from each SN and then applies adaptive clustering algorithm. Authors presented the results of energy consumption by ADCA; however the effect on the network lifetime is not discussed.

(12) *PEAP:* Authors in [19] present PEAP (Power aware Energy Adaptive Protocol) with hierarchical clustering for WSN. PEAP model is based on confidence value associated with broadcast from CHs. Confidence value of a CH is a function of some parameters such as distance between the CH and the node, the CH current battery power and the number of nodes already were a member of this CH. PEAP uses CSMA (Carrier Sense Multiple Access) as the MAC layer protocol. The efficiency of PEAP is presented via energy consumption and network lifetime comparisons with LEACH; however exact figures are not discussed.

(13) *Clustering by Ebadi, et al. :* A clustering algorithm for selection of two CHs for each cluster is proposed in [20]. In this algorithm, the function of one CH is data collection, data aggregation and data transmission to the BS or high level CH and the function of second CH is to receive data from low level CH and forward them to high level CH. The CH selection is made based on the remaining energy of SN, number of

neighbours and the received signal strength. Simulation results confirm that the proposed clustering method improves the WSN lifetime more than 28% compared with LEACH. However, the energy consumption comparison is not discussed, which is important issue in WSNs.

(14) *EECF:* An EECF (Energy-Efficient Cluster Formation) protocol is proposed in [21], which organizes the network into clusters based on three-way message exchange between every SN and its one-hop neighbors. In each cluster one SN is selected as CH which also acts as relay node because it routes the received data from peer CHs towards sink. The performance of EECF is analyzed using network lifetime, ratio of elected CHs only and total energy cost incurred by EECF is not discussed.

(15) *One of the Significant Issues is the Consideration of Optimal Number of CHs during Clustering.* Among all iterative clustering protocols only a single study [10] has taken into account optimal number of CHs during its clustering procedure. One of the benefits of clustering is reduction of communication overhead for both single hop and multi hop networks. However, only some of the iterative clustering protocols [7,12,17-18,21] have emphasized on this overhead in their results.

2.2 Probabilistic Clustering Protocols

In probabilistic protocols each SN decides its role independently in the cluster based network whilst maintaining low communication overhead. In these protocols energy level of SNs is used as the primary parameter for selection of CHs. A few examples of this approach are: LEACH, NAC and HEED.

(1) *LEACH:* LEACH is an application specific clustering protocol [9]. In LEACH, SNs are organized in clusters and the CHs are not fixed. The SNs employ randomized rotation to choose the CHs depending on the amount of energy left to evenly distribute energy load among altered SNs at different times in the network. The CHs then announce that they are now new CHs. Each SN decides which cluster it should join depending on the minimum energy used to communicate with a CH. The CHs assign TDMA schedules to SNs in their clusters. The SNs send data to the CHs according to TDMA schedules and the CHs are responsible to send aggregated data to the BS. The increase in the number of SNs necessitates reassignment TDMA schedule resulting in communication overhead. In LEACH since CHs are randomly chosen so it is possible that in some areas of network CH may not exist. A limitation of this scheme is that it assumes that each SN has enough power to communicate with the BS, which is not satisfied in large WSNs.

(2) *EDACH:* In [22] an EDACH (Energy Driven Adaptive Clustering Hierarchy) scheme is proposed which is an improvement over LEACH. It selects a proxy node to carry out the duty of a CH having insufficient energy. It selects more CHs in the area distant from the BS. Each round of EDACH consists of set-up phase and self organized data gathering and transmission phase. EDACH is useful for large area WSNs. The performance evaluation of EDACH shows that it maximizes network lifetime by selecting proxy node and CHs using distance to the BS. However, energy consumed by EDACH is not discussed.

(3) *NAC:* In NAC (Neighbour-Aware Cluster Head) [23], clustering is based on the current energy level of a SN. Three energy levels are defined for each SN. A SN can become CH if its energy level is greater than threshold_1 and a SN can become part of a cluster if its energy level is between

threshold_1 and threshold_2. A SN will join a CH if CH energy level is greater than threshold_1. If a CH energy is between threshold_2 and threshold_3 it will request its member SNs to search for another CH. A CH will send supervisor withdrawal message to its member nodes, when its energy decreases below threshold_3. In NAC, LINT (Local Information No Topology) and LILT (Local Information Link-State Topology) are used to control the number of member nodes in a cluster. And the synchronization between a CH and its members is maintained via RBS (Reference Broadcast Synchronization) algorithm.

(4) *HEED:* The HEED protocol [24] makes use of the residual energy as the main parameter to probabilistically select a CH. In order to balance load among the CHs, HEED entails a secondary parameter such as a SN degree or average distance to neighbours, besides residual energy, which makes it a hybrid protocol. HEED ends after invariable number of steps regardless of the network diameter resulting in a balanced allocation of the elected set of CHs across the network. It improves network life time and decreases message overhead.

(5) *DEBC:* DEBC (Distributed Energy Balanced Clustering) [25] is a modified form of LEACH [9]. It also works in two phases with multiple rounds. It dynamically forms clusters, selects CHs and chooses cluster senders based on remaining energy. In DEBC, the function of CHs is to create and send TDMA schedule to the SNs and the function of CH senders is to send aggregated data to the BS using single-hop or multi-hop communication. In order to transmit data in bidirectional way in a cluster, DEBC arranges the SNs into a bidirectional ring topology to form a cluster. For a specific set of rounds, the CHs' role remains fixed and cluster senders' role is rotated to different SNs. It outperforms LEACH in terms of energy consumption and network lifetime; however practical implementation of TDMA structure requires much effort.

(6) *Clustering by Kim et al.:* The scheme proposed in [26] improves LEACH protocol by prolonging network lifetime with the consideration of remaining energy during CH selection procedure. In the proposed scheme each node self selects the CH by a novel probability function which is associated with energy possession rate, individual round and the count the node itself had been selected as the CH. In set up phase, a fixed proportion of SNs stochastically select them selves as CHs. In steady state phase, each SN collects data and sends data packets to their CHs using CSMA/CA. Simulation comparisons show that the proposed scheme achieves an obvious improvement in the network lifetime compared with existing schemes.

(7) *RRCH:* A RRCH (Round Robin Cluster Head) algorithm is presented in [27]. To avoid repetitious cluster set up process RRCH forms clusters in one setup phase. After the detection of anomalous SN, RRCH broadcast this information to the entire cluster through frame modification, and then each SN deletes the anomalous node from its schedule. The RRCH improves the energy effectiveness compared to LEACH algorithm by avoiding the cyclic set-up processes involved in dynamic clustering. However, the improvement in the network lifetime by RRCH is not depicted.

(8) *Clustering by Yang and Sikdar:* The study [28] presents a framework for calculating optimal probability with which a node is selected to become a CH with the intention to reduce total energy expenditure by the network. In this paper a sleep wakeup based, decentralized MAC

protocol is used in LEACH protocol instead of TDMA. This protocol is scalable with the increase of number of SNs. Optimal probability of CH selection is calculated for both small and large network scenarios. In case of small network, all the CHs are assumed at the same distance from the sink and in case of large network; distances of different areas of the network and the CHs from the sink are assumed to be different. The calculated optimal probability for CH selection gives better network lifetime as compared to LEACH.

(9) *DSC:* The authors of [29] propose two cases of DSC (Dynamic/Static Clustering) protocol. The basic functionality of dynamic case is like LEACH-Centralized which works in two phases. In setup phase, BS is responsible to select CHs for each cluster based on the energy levels and positions of the SNs. In the steady state phase, CHs use TDMA schedule and member nodes transmit data to the CH only in the allocated time slots. The static case of DSC consists of only steady phase, in which a new cluster selection phase is initiated after a particular number of rounds. Thus, static case has less number of cluster formation phases than dynamic case which results in reduction of transmission overhead. The better performance of DSC than LEACH-C in terms of energy efficiency, network lifetime and communication overhead is revealed. The limitation of DSC is the assumption that all SNs can communicate with far away BS which is not possible in wide area WSNs.

(10) *FT-DSC:* An improvement over DSC is discussed in [30] in terms of energy efficiency and provision of fault tolerance called FT-DSC (Fault Tolerant Dynamic Static Clustering). In FT-DSC CHs can detect the failure of member nodes and BS can detect the failure of CHs. The CHs or BS subscribe to the member nodes for the occurrence of event of interest which results in reduction of energy consumption. It is found that FT-DSC has better performance than DSC in terms of control packets, number of rounds and energy consumption.

(11) *EECPNL:* An EECPNL (Energy Efficient Clustering Scheme to Prolong Sensor Network Lifetime) is proposed in [31]. The scheme works in rounds like LEACH and each round consists of two parts namely set-up and data transfer phase. During set-up phase after the cluster formation, a head list is chosen for each cluster. Based on the residual energy, one member of every head list is selected as the active CH. In data transfer phase, active CH collects the data from the member nodes, aggregates it and forwards it to the BS. It outperforms the LEACH in terms of energy consumptions and network life time.

(12) *SDEEC:* B. Elbhiri et al. [32] have proposed a stochastic scheme to prolong the life time of heterogeneous WSNs named SDEEC (Stochastic Distributed Energy-Efficient Clustering). SDEEC uses DEEC scheme as its base, where SNs elect themselves as CH based on the initial and residual energy levels. It guesses the threshold value of network lifetime, which is used as the bases to calculate the reference energy that every SN consumes during each round. A balanced CH election procedure for all SNs in the network is adapted using their residual energy which results in reduction of intra-cluster transmission. SDEEC gives better performance than DEEC in terms of network lifetime, remaining energy in the network and number of messages received.

(13) *EDBC:* An energy efficient clustering communication protocol for WSN named EDBC (Energy and Distance Based Clustering) is

presented in [33]. For selection of CHs EDBC takes into account both the remaining energy of SNs and the distance of each SN from the BS. Hence the SNs having smaller amount energy than the other SNs and the SNs having more distance from the BS have the smallest opportunity to be selected as a cluster-head for current round. The whole network is divided into circular segments around the BS and the number of CHs in each segment is different. EDBC is one of the improvements over LEACH in terms of energy consumption and network life time.

(14) In majority of probabilistic clustering protocols, energy level of SNs is used for clustering and in few protocols distance to the BS is also considered. The protocols discussed in [22-23, 25,27,29,31] do not take into account the optimal CH selection probability which is one of the significant issues. One of the solutions is to employ existing investigation for example, in [28] a decentralized MAC protocol to control the sleep and wake up schedules of SNs in LEACH along with optimal CH selection framework is presented. A probability function for optimal selection of CHs is given in [26]. Further, only limited probabilistic clustering protocols [24,29,32] have mentioned incurred communication overhead in their results, which is one of the nontrivial issues of clustering. Fig. 2 presents the classification of node clustering protocols.

3. COMPARISON OF CLUSTERING PROTOCOLS

Table 1 presents the comparison of all the clustering protocols which are discussed in this paper with respect to following parameters.

CH Selection: Table 1 compares the clustering schemes based on the method of CH selection. For example, CH selection can be static or dynamic.

Objective: One of the ways to differentiate clustering protocols is their main objective. Different clustering objectives can be load balancing, maximizing network lifetime, fault-tolerance etc. Table 1 illustrates the main objective of each clustering scheme.

Cluster Criteria: It is critical to perform optimal clustering in terms of energy efficiency and reducing the clustering cost. Different strategies are applied in the clustering protocols which are discussed in this paper.

Methodology: Clustering can take place in a disseminated way without synchronization with centralized authority or with coordination with centralized authority. In few clustering schemes hybrid approach is also adapted.

Data Transmission: Some clustering schemes are based on direct communication called single-hop between a SN and its designated CH. However, in case of limited communication range of SNs multi-hop sensor-to-CH connectivity is needed.

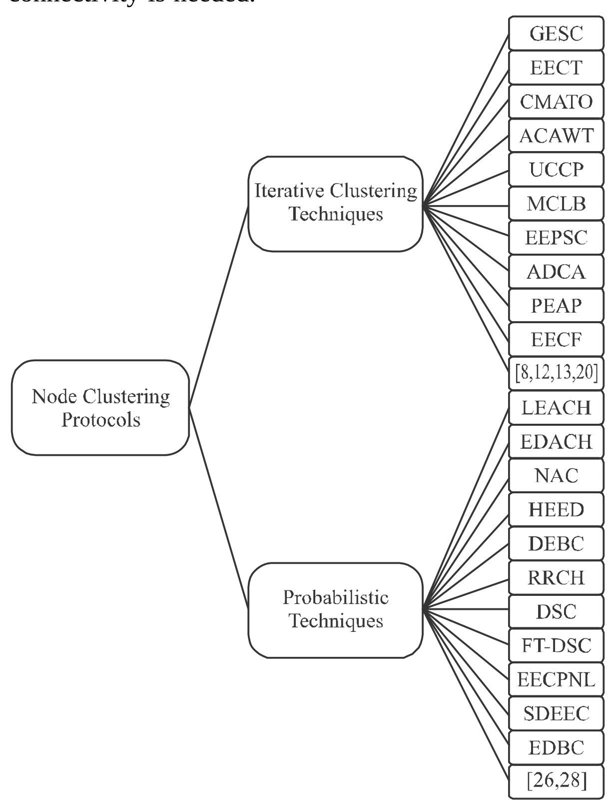

FIGURE 2 CLASSIFICATION OF NODE CLUSTERING PROTOCOLS

Table 2 compares different clustering schemes with LEACH in terms of network lifetime and energy consumption because these are two important issues in a network. LEACH is selected for comparison which is one of the nontrivial representatives of the cluster-based methods with load balancing. The constrained energy of SNs results in an inadequate network lifetime for SNs in a WSN. One of the ways to improve network lifetime is clustering which also reduces the energy usage.

TABLE 1 COMPARISON OF CLUSTERING PROTOCOLS

Algorithm	CH Selection	Objective	Cluster Criteria	Methodology	Data Transmission
GESC [7]	Dynamic	Network longevity	SN implication	Distributed	Single-hop
[8]	Static	Reduce energy	One-hop distance	Distributed	Single-hop
LEACH [9]	Dynamic	Save energy	CH frequency	Distributed	Single-hop
EECT [10]	Static	Energy efficiency and load balancing	Distance between virtual hexagon's centre and SN	Centralized	Single-hop
CMATO [11]	Existing Clustering	Fault recovery	Pre deployment	Distributed	Multi-hop
[12]	Static	Reduce cost of cluster formation	Cut algorithm	Centralized	Multi-hop
[13]	Dynamic	Improve Operational efficiency	Kohonen Self Organization Map Neural Network	Centralized	Not defined
ACAWT [14]	Dynamic	Extend the lifetime of the network	Clustering algorithm via waiting timer and local criteria	Distributed and centralized	Multi-hop
UCCP [15]	Dynamic	Energy efficiency and prolong network lifetime	Multi-criterion optimization algorithm	Distributed	Two-hop
MCLB [16]	Dynamic	Load balancing and energy efficiency	Signal strength	Distributed	Multi-hop
EEPSC [17]	Static	Energy efficiency	Energy level	Distributed	Single-hop
ADCA [18]	Dynamic	Reduce power and minimize data loss	Data sending rate	Distributed	Not defined
PEAP [19]	Dynamic	Longer lifespan and reduce energy consumption	Confidence value	Distributed	Multi-hop
[20]	Static	Prolong network life time and energy efficiency	Remaining energy, distance to its CH, number of neighbours	Centralized	Single-hop/ Multi- hop
EECF [21]	Dynamic	Increase network lifetime	Residual energy and degree	Distributed	Two-hop
EDACH [22]	Dynamic	Enhance network lifetime	Distance to the base station	Hybrid	Single-hop
NAC [23]	Dynamic	Reduce energy	Energy level	Distributed	Multi-hop
HEED [24]	Dynamic	Save energy	Residual energy	Distributed	Single-hop/ multi-hop
DEBC [25]	Dynamic	Reduce energy	Energy level	Distributed	Single-hop/ multi-hop
[26]	Dynamic	Prolong network lifetime	Probability function	Distributed	Single-hop
RRCH [27]	Dynamic	Energy efficiency	Round robin	Hybrid	Single-hop
[28]	Dynamic	Optimal CH selection	Distance to the sink node	Distributed	Not defined
DSC [29]	Dynamic/static	Energy efficiency and network longevity	Energy level and position of SNs	Hybrid	Single-hop
FT-DSC [30]	Dynamic/static	Provision of fault tolerance	Energy level and position of SNs	Hybrid	Single-hop
EECPNL [31]	Static/dynamic	Prolong network lifespan	Residual energy	Distributed	Single hop
SDEEC [32]	Dynamic	Save energy and extend network lifetime	Initial and residual energy	Distributed	Multi-hop
EDBC [33]	Static	Reduce total energy consumption	Residual energy and distance to the BS	Centralized	Single-hop

4. CHALLENGES

Challenges making it complicated to implement node clustering in WSN applications include:

- CH selection, optimal cluster size, re- selection of CHs and cluster maintenance must be considered during clustering because in case of dynamic clustering these might be affected.
- During dynamic clustering determining optimal frequency for CH rotation in order to maximize network lifetime.

TABLE 2 NETWORK LIFETIME AND ENERGY CONSUMPTION COMPARISON WITH LEACH

Algorithm	Lifetime Characteristics	Energy Consumption
[8]	Prolong system lifetime as compared to LEACH.	The total energy consumption by the network is reduced than LEACH.
LEACH [9]	First node death occurs 8 times later than conventional methods, increasing network lifetime.	Energy reduction is achieved compared to conventional methods.
EECT[10]	LEACH, HEED and EECT can run for 150, 220, 280 rounds respectively before the death of first node.	Energy consumption comparison is not discussed.
[12]	Network lifetime comparison is not discussed.	Cost of clustering is reduced by a factor of 10 than LEACH-C, resulting in reduction of total energy consumption.
UCCP [15]	UCCP extends the network lifetime approximately 25% compared to LEACH	UCCP accomplishes smallest amount of energy consumption as compared to LEACH.
MCLB [16]	Simulation results reveal that number of dead nodes is more in LEACH than MCLB.	Energy consumption comparison confirms that MCLB consumes less energy than LEACH.
EEPSC [17]	The first node death in LEACH occurs after 220 seconds, whereas all SNs stay alive for 320 seconds in EEPSC; which is 45% additional than LEACH.	Energy consumption comparison is not discussed.
PEAP [19]	The network lifetime of PEAP in terms of number of dead SNs is considerably larger than LEACH.	The energy consumption of PEAP is smaller than LEACH.
[20]	Increase network lifetime more than 28 % in comparison with LEACH.	Energy consumption results are not given.
EDACH [22]	EDACH algorithm increases the WSN lifetime of LEACH by 80%.	Energy consumption comparison is not discussed.
HEED [24]	HEED can run for 320 more rounds than LEACH before the first node dies.	Ratio of energy used in clustering to total dissipated energy is approximately two times more in LEACH than HEED.
DEBC [25]	DEBC achieves 32% more number of rounds than LEACH-C.	LEACH-C consumes more energy than DEBC.
[26]	Obtains 74% and 58% improvement in the network lifetime in terms of first node death over LEACH and PEACH respectively.	The proposed scheme consumes less energy than LEACH.
RRCH [27]	Network lifetime comparison is not discussed.	Energy consumption is decreased than the amount of energy consumed by recurring set-up processes in dynamic clustering
[28]	Network lifetime comparison is given with varying number of nodes and probability values. It maximizes the lifetime of the network than LEACH.	Comparative energy consumption figures are not given.
DSC [29]	The first node death in DSC takes place afterwards than the first node death in LEACH-C.	DSC reduces communication overhead in setup phase which results in energy reduction as compared to LEACH-C.
EECPNL [31]	Network lifetime in terms of number of dead nodes is significantly improved than LEACH.	The average energy consumption by EECPNL is lower than LEACH.
EDBC [33]	It increases the network lifetime by 94% in terms of the first node death and more than 6% in terms of the half of the SNs alive compared with LEACH.	Energy saving up to 15% is obtained with EDBC compared to LEACH.

- Scheduling intracluster and intercluster transmissions in favour of energy reduction.
- Calculating optimal number of clusters and cluster size during the cluster formation [11].
- Clustering protocols using TDMA necessitate synchronization to retain transmission schedule of SNs.
- Exploring hybrid static/dynamic clustering protocols and their feasible implementations.
- Investigating distribution of SNs in favour of static and dynamic clustering their energy consumption comparison.

5. CONCLUSION

The growing need of WSN in diverse applications has exposed many challenges to researchers. Due to energy limitations of SNs, significant attention has been paid to clustering algorithms which are one of the ways for energy consumption reduction. In this paper we attempted to present the comprehensive analysis of current state of node clustering schemes for WSNs. We summarized and compared their performance based on particular parameters. We also compared clustering schemes with LEACH, which is one of the nontrivial clustering protocol based on two main network constraints. The static and dynamic clustering protocols presented in this paper offer a promising improvement over conventional clustering; however there are issues which need to be explored. Most of the protocols have concentrated on energy minimization and prolonging network lifespan. However, they should be cooperated with the cost of intracluster and intercluster transmissions.

ACKNOWLEDGEMENTS

The authors would like to thank Mehran University of Engineering & Technology, Jamshoro, Pakistan, for providing access to online journals. Authors also appreciate Mr. Mohsin Ali Memon, Lecturer Department of Software Engineering, Mehran University of Engineering & Technology, Jamshoro, for discussions on this study.

REFERENCES

[1] Anastasi, G., Conti, M., Di Francesco, M., and Passarella, A., "Energy Conservation in Wireless Sensor Networks: A Survey", Ad Hoc Networks, Volume 7, No. 3, pp. 537-568, May, 2009.

[2] Younis, O., Krunz, M., and Ramasubramanian, S., "Node Clustering in Wireless Sensor Networks: Recent Developments and Deployment Challenges", IEEE Network, Volume 20, No. 3, pp. 20-25, 2006.

[3] Abbasi, A., and Younis, M., "A Survey on Clustering Algorithms for Wireless Sensor Networks", Computer Communications, Volume 30, pp. 2828-2841, 2007.

[4] Dechene, D., J., El Jardali, A., Luccini, M., and Sauer, A., "A Survey of Clustering Algorithms for Wireless Sensor Networks", Project Report, 2006.

[5] Kumarawadu, P., Dechene, D., J., Luccini, M., and Sauer, A., "Algorithms for Node Clustering in Wireless Sensor Networks: A Survey", Proceedings of 4th International Conference on Information and Automation for Sustainability, pp. 295-300, 12-14 December, 2008.

[6] Chen, W.-P., Hou, J.C., and Sha, L., "Dynamic Clustering for Acoustic Target Tracking in Wireless Sensor Networks", Proceeding of 11th IEEE International Conference on Network Protocols, pp. 284-294, Atlanta, Georgia, USA, November, 2003.

[7] Dimokas, N., Katsaros, D., and Manolopoulos, Y., "Node Clustering in Wireless Sensor Networks by Considering Structural Characteristics of the Network Graph", Proceedings of Fourth International Conference on Information Technology, pp. 122-127, 2007.

[8] Xiang, M., Shi, W., Zhang, X., Luo, Z., and Yang, X., "A New Clustering Algorithm Based on the Optimum One-Hop Distance in Wireless Sensor Networks", International Conference on Embedded Software and Systems, 29-31 July, 2008.

[9] Heinzelman, W.R., Chandrakasan, A., and Balakrishnan, H., "An Application-Specific Protocol Architecture for Wireless Microsensor Networks", Proceedings of IEEE Transactions on Wireless Communications, Volume 1, No. 4, pp. 660-670, October, 2002.

[10] Xin, G., Yong Xin, W., and Fang, L., "An Energy-Efficient Clustering Technique for Wireless Sensor Networks", Proceedings of International Conference on Networking, Architecture, and Storage, pp. 248-252, 2008.

[11] Lai, Y., and Chen, H., "Energy-Efficient Fault-Tolerant Mechanism for Clustered Wireless Sensor Networks", Proceedings of 16th International Conference on Computer Communications and Networks, pp. 272-277, 13-16 August, 2007.

[12] Soude, H., and Mehat, J., "Energy Efficient Clustering Algorithm for Wireless Sensor Networks", Proceedings of International Conference on Wireless and Mobile Communications, pp.7-7, 29-31 July, 2006.

[13] Veena, K.N., and Vijava Kumar, B.P., "Dynamic Clustering and Analysis for Sensor Networks", Proceedings of IEE International Conference on Information, Communications and Signal Processing, pp. 1-4, 2007.

[14] Wen, C.Y., and Sethares, W.A., "Adaptive Decentralized Re-Clustering for Wireless Sensor Networks", Proceedings of IEEE International Conference on Systems, Man, and Cybernetics, pp. 2709-2716, Taipei, Taiwan, October, 2006.

[15] Aslam, N.W., Phillips, W., and Robertson, W., "A Unified Clustering and Communication Protocol for Wireless Sensor Networks", IAENG International Journal of Computer Science, Volume 35, No. 3, pp. 249-258, 2008.

[16] Israr, N., and Awan, I., "Multihop Clustering Algorithm for Load Balancing in Wireless Sensor Networks", International Journal of Simulation, Systems, Science and Technology, Volume 8, No. 1, pp. 13-25, 2007.

[17] Zahmati, A.S., Abolhassani, B., Beheshti Shirazi, A.A., and Shojaee Bakhtiari, A., "An Energy-Efficient Protocol with Static Clustering for Wireless Sensor Networks", World Academy of Science, Engineering and Technology, Volume 28, pp. 69-72, 2007.

[18] Manisekaran, S.V., and Venkatesan, R.,"An Adaptive Distributed Power Efficient Clustering Algorithm for Wireless Sensor Networks", American Journal of Scientific research, Volume 10, pp. 50-63, Euro Journals Publishing, July, 2010.

[19] Golsorkhtabar, M., Hosinzadeh, M., Heydari, M., and Rasouli, S., "New Power Aware Energy Adaptive Protocol with Hierarchical Clustering for Wireless Sensor Networks", International Journal of Computer and Network Security, Volume 2, No. 4, pp. 38-40, April, 2010.

[20] Ebadi, S., Ghasembaglou, M., Navin, A.H., and Mirnia, M.K., "Energy Balancing in Wireless Sensor Networks with Selecting Two Cluster-Heads in Hierarchical Clustering", Proceedings of International Conference on Computational Intelligence and Communication Networks, pp. 230-233, IEEE Computer Society, 2010.

[21] Chamam, A., and Pierre, S., "A Distributed Energy-Efficient Clustering Protocol for Wireless Sensor Networks", Computers and Electrical Engineering, Volume 36, pp. 303-312, 2010.

[22] Kim, K.T., and Youn, H.Y., "Energy-Driven Adaptive Clustering Hierarchy (EDACH) for Wireless Sensor Networks", Proceedings of the International Federation for Information Processing, pp. 1098-1107, 2005.

[23] Sheltami, T.R., "Neighbour-Aware Cluster Head for SNET", Proceedings of 22nd International Conference on Advanced Information Networking and Applications Workshops, pp. 561-566, 25-28 March, 2008.

[24] Younis, O., and Fahmy, S., "Distributed Clustering in Ad Hoc Sensor Networks: A Hybrid, Energy-Efficient Approach", Proceedings of IEEE Transactions on Mobile Communications Volume 3, No. 4, pp. 366-379, October-December, 2004.

[25] Bajaber, F., and Awan, I., "Distributed Energy Balanced Clustering for Wireless Sensor Networks", Proceedings of 10th International Conference on Computer and Information Technology, pp.385-392, IEEE CS Press, Bradford, UK, 29 June - 01 July, 2010.

[26] Kim, K.T., Yoo, H.K., Son, B.H., and Youn, H.Y., "An Energy Efficient Clustering Scheme for Self-Organizing Distributed Wireless Sensor Networks", Proceedings of 10th IEEE International Conference on Computer and Information Technology, pp.1468-1473, IEEE CS Press, Bradford, UK, 29 June - 01 July, 2010.

[27] Nam, D.H., and Min, H.K., "An Efficient Ad-Hoc Routing Using a Hybrid Clustering Method in a Wireless Sensor Network", Proceedings of Third IEEE International Conference on Wireless and Mobile Computing, Networking and Communications, pp. 60-65, 2007.

[28] Yang, H., and Sikdar, B., "Optimal Cluster Head Selection in the LEACH Architecture", Proceeding of IEEE International Performance, Computing, and Communications Conference, pp. 93-100, 2007.

[29] Bajaber, F., and Awan, I., "Dynamic/Static Clustering Protocol for Wireless Sensor Network", Proceedings of Second UKSIM European Symposium on Computer Modeling and Simulation, pp. 524-529, 2008.

[30] Karim, L., Nasser, N., and Sheltami, T., "A Fault Tolerant Dynamic Clustering Protocol of Wireless Sensor Networks", Proceedings of 28th IEEE conference on Global Telecommunications, pp. 1-6, 30 November - 4 December, 2009.

[31] Sehti, B.T. Pal, T., and Dasbit, S., "An Energy-Efficient Clustering Scheme to Prolong Sensor Network Lifetime", Proceedings of International Conference and Workshop on Emerging Trends in Technology, pp. 838-842, Mumbai, India, 2010.

[32] Elbhiri, B., Saadane, R., and Aboutajdine, D., "Stochastic Distributed Energy-Efficient Clustering (SDEEC) for Heterogeneous Wireless Sensor Networks", ICGST-CNIR Journal, Volume 9, No. 2, pp. 11-17, December, 2009.

[33] Saeidmanesh, M., Haji Mohammadi, M., and Movaghar, A., "Energy and Distance Based Clustering: An Energy Efficient Clustering Method for Wireless Sensor Networks", World Academy of Science, Engineering and Technology, Volume 55, pp. 555-559, 2009.