
Muhammad K. Khan,1,2 Muhammad Shiraz,2 Qaisar Shaheen,3 Shariq Aziz Butt,4 Rizwan Akhtar,5 Muazzam A. Khan,6 and Wang Changda1

1School of Computer Science and Communication Engineering, Jiangsu University, Zhenjiang, China
2Department of Computer Science, Federal Urdu University of Arts, Science and Technology, Islamabad, Pakistan
3Department of Computer Science & Information Technology, Superior College, Lahore, Pakistan
4Department of Computer Science, University of Lahore, Pakistan
5Department of IT and Computer Science, Pak-Austria Fachhochschule Institute of Applied Sciences and Technology, Haripur, Pakistan
6Department of Computer Science, Quaid-i-Azam University, Islamabad, Pakistan

Correspondence should be addressed to Qaisar Shaheen; qaisar.shaheen2002@gmail.com and Wang Changda; changda@ujs.edu.cn

Received 28 February 2020; Revised 26 March 2021; Accepted 24 April 2021; Published 8 May 2021

1. Introduction

Wireless sensor network (WSN) comprises of large number of tiny and moderately low-priced computational nodes that transmit the valuable collected information to the base station (BS) for applicable processing. Figure 1 shows a typical structure of WSN. From last decade, wireless sensor networks (WSNs) have mesmerized extensive attention from academic research as well as from academic research. The major factor behind most of the research efforts in the field of WSNs is because it helps to resolve most of the real life problems like security, human safety, health care, and defense sector [1–8]. In WSNs, a number of research directions are available which includes corporeal design, routing strategy, procedures for power management, security issues, and sensing capability of sensor nodes [3]. Lifetime of the sensor nodes is the utmost issue for WSNs, because sensor nodes have very limited power resources [4–6]. Routing
protocol plays a major role in the lifetime of the sensor nodes. Routing in WSN is not similar to other wireless networks because of various unique properties of sensor node like energy constraints, processing accomplishments, transmission of collected information from multiple nodes to a single base station, improbability of global address and random deployment of sensor nodes, etc. In order to accommodate these types of properties, different types of routing protocols were developed. The ultimate goal of these routing protocols is to achieve energy efficiency and maximize the overall network lifetime.

Sensor node life is dependent on the life time of battery which provide power to sensor node, on which the life span of the whole network is dependent. An energy source (power entity) supplies energy to the memory unit, sensing unit, and transceiver. The memory unit is used for the storage of application related data and also contains device identification information, sensing unit contains sensors to capture data from their environment, and the transceiver is responsible for the transmission and reception of data. Quick energy dissipation in sensor nodes which make them lifeless is the major problem in the field of WSNs [9]. By experiments, it is observed that inefficient routing algorithm causes quick dissipation of battery. Hence, it is major requirement to design and use energy efficient routing algorithms in WSNs which will definitely increase overall life time of WSNs [1, 10–12].

There are three main types of routing protocols specifically proposed for WSNs: (1) flat, (2) location-based, and (3) hierarchical. In flat routing, the multihop strategy is used, and each sensor node performs the same tasks [13]. The sensing mission is performed by the collaboration of sensor nodes with each other. In location-based routing, position information of sensor nodes is used to spread the data to a specific region, instead of transmission to whole network [14]. In hierarchical routing, network area is divided into clusters, and sensor nodes with higher energy are responsible for the processing and transmission of information. Hierarchical type of routing protocols provides best results in terms of energy efficiency [2].

Figure 1: Structure of WSN.

Figure 2: These attributes are discussed and analyzed in Section 3. The base of almost all hierarchical energy efficient routing protocols is LEACH. The enhancements made in LEACH are achieved by making changes in the technique of routing, and the target of almost all enhancements is to optimize energy utilization and maximization of overall lifetime of sensor network [15].

The primary attention of this research is to focus on mechanism of existing routing protocols and find their shortcomings which may lead us to work on it to have efficient and optimum routing protocol for the next generation networks. Among various identified issues in WSN routing, energy efficiency is more important due to its linkage with entire computation process of the sensor nodes. It is the task of routing protocol to establish connection between sensor nodes and CH, sensor nodes share hello messages on continuous basis to stay connected with the active nodes. The distance between communicating nodes also affects the energy efficiency.

The rest of the paper is organized as follows: Section 2 introduces different LEACH based hierarchical routing protocols. Performance results and evaluation of hierarchical routing protocols are presented in Section 3. Routing issues and challenges are discussed in Section 4. Finally, Section 5 ends up with the conclusion.

2. Materials and Methods

In this review paper, a detailed literature study of hierarchical routing protocols for WSNs is performed, and different technologies of routing protocols are analyzed. The weaknesses and strengths of reviewed hierarchical routing protocols are highlighted. Then, performed comparative analysis on the basis of multiple factors discussed in Results and Discussion.

3. Significance of Study

The WSN network is the very broad and trendy domain of research for both industry and research community. It has
so many parameters to study and protocols to adopt for communication. In this study, we highlighted the protocols of routing in the WSN network as the data routing is the most important aspect of any network. In this, we have conducted a comprehensive survey about the routing protocols with respect of performance, network life, reliability, and so on. In this paper, we thoroughly explain these protocol’s features, performance levels, and issues with the routing. This paper is aimed at continuing the research work in the routing protocols by describing the each protocol for the industry professionals and research community. This paper may help to find the issues with the protocols and continue introduce of new works in WSN network.

4. Review on Hierarchical Routing Protocols for WSNs

Cluster based hierarchical routing protocols divide network into multiple clusters by defining hierarchy of sensor nodes. Every cluster consists of one cluster head (CH) and multiple sensor nodes. Each sensor node collects information and forwards it to its CH. CHs perform aggregation on the collected information and forward aggregated data to the next hop or main station which is called as BS, on the basis of predefined routing algorithm.

LEACH routing protocol provides base for the development of most of the hierarchical routing protocols in WSNs. LEACH is self-organizing and adaptive clustering routing algorithm [1]. In it, the whole network area is divided into multiple clusters. Every cluster consists of multiple no. of sensor nodes, one sensor node will be designated as CH, and all other sensor nodes existed in that cluster become member nodes. Member nodes transfer data to CH, and CH forward collected data to BS after performing aggregation process [1, 16, 17].CHs perform as an intermediate node between member nodes and BS. Due to additional duties, CHs dissipate more energy as compared to normal nodes. Figure 3 illustrates the basic communication hierarchy of LEACH routing protocol.

Setup phase and steady-state phase are two basic phases of every round of LEACH routing protocol. In the setup phase, clusters are structured, and CHs have been selected. Data transmission and aggregation is done in steady state phase [16]. To achieve minimum overhead, setup phase is shorter as compared to steady state phase. After formation of clusters, every sensor node takes decision whether to serve as a CH or not, for the existing round. The proposed ratio of CHs in the network (resolve a priori) and the number of times so for the node has been selected as a CH are the decision criteria for the sensor node to serve as a CH. For example, a random number from 0 to 1 is selected by a node “z.” The sensor node will be designated as CH only if the threshold $T(z)$ is more than a predefined value. For the calculation of $T(z)$ [1, 10, 16] below formula (Equation (1)) is used.

$$T(z) = f(z) = \begin{cases} P & \text{if } z \in G, \\ 1 - P \times (r \mod \frac{1}{P}) & \text{otherwise,} \end{cases}$$

where $z$ is the total sensor nodes in network area, $P$ is the preferred percentage of CH, $r$ is the current round, and $G$ is the set of node energy needed for communication.

CHs use MAC protocol to broadcast status messages, i.e., carrier sense multiple access (CSMA). In the selection of CHs, Received Signal Strength Indication (RSSI) is computed by member nodes. CHs form Time Division Multiple Access (TDMA) schedules for associated member nodes. Specific time slots have been allotted to each member node, and every
member node sends data according to their allotted time slot. Rest of the time member nodes stay in sleeping state. This feature in LEACH leads to less dissipation and increase the battery life of each single node.

Comparatively LEACH hierarchical routing protocol consumes less energy and improves overall lifetime of the network than other nonhierarchical routing protocols. In spite of significant improvement in energy efficiency, still there are areas which may be improved. For instance, firstly, LEAH routing protocol do not perform well for large-scale networks. The reason behind that is its single hop strategy. Secondly, CH’s selection mechanism is not efficient; there are possibilities that a node with less residual energy may be selected as CH and a node with more residual energy may not be selected as CH. Thirdly, LEACH cannot ensure proper distribution of CHs in the network field, and fourthly, extraoverhead can be reduced by replacement of dynamic clustering with static clustering techniques.

MODLEACH [18] improves LEACH routing protocol’s CH replacement mechanism. In it, a programmed threshold level of energy is used. In every round, CH’s energy is compared with threshold level, and until the energy level of CH is greater than the minimum threshold level, the CH performs his duties and may not be replaced with new CH. By adopting improved CH’s replacement mechanism, MODLEACH significantly reduced energy consumption.

In addition to optimization of CHs selection technique, two different amplification levels for communication signals have been introduced. On the basis of distance between communication devices and type of communicating devices, different amplification levels have been used as shown by below formulas (Equations (2)– (5)). For intracluster communication, less amplification level is used as compared to cluster to BS communication. This strategy not only saves energy consumption but also reduce packet drops and also reduce interference with other signals.

\[
\text{Amplification Energy (Cluster to BS) } d \geq d_0 (E_{fs}) = \frac{10pJ}{m^2},
\]

(2)

\[
\text{Amplification Energy (Cluster to BS) } d \leq d_0 (E_{mp}) = \frac{0.0013pJ/\text{bit}}{m^2},
\]

(3)

\[
\text{Amplification Energy (Intra Cluster Comm.) } d \geq d_1 (E_{fs1}) = \frac{E_{fs}}{10},
\]

(4)

\[
\text{Amplification Energy (Intra Cluster Comm.) } d \leq d_1 (E_{mp1}) = \frac{E_{mp}}{10}.
\]

(5)

MODLEACHHT [18] and MODLEACHST [18] are two variants MODLEACH [18]. MODLEAHHT uses hard threshold level for communication process. The hard threshold level is the attribute outright value which is used to trigger the sensor node to turn on its transmitter and connect with CH to start communication process. On the other hand, soft threshold level is a little variation in the value of recognized attribute which triggers sensor nodes to power on the probe and start communication process [19]. Both MODLEACHHT and MODLEACHST use reactive approach and significantly improve energy consumption as compared to MODLEACH.

MODLEACH [18], MODLEACHST [18], and MODLEACHHT [18] adopt efficient CH replacement mechanism, but routing protocol needs to be improved. Single hop
routing technique is used which do not work well for large-scale networks. It is observed that in these routing protocols, the CHs located at far end, dissipate more energy and become dead very soon.

Three Levels Hierarchical Clustering LEACH Protocol (TLHCLP) [20] assumes that the BS is situated at middle of the network field. The communication strategy is defined on the basis of predefined radius around the BS. Figure 4 shows basic communication architecture of TLHCLP. In it, cluster member nodes are at level 1, CHs located outside the range of predetermined radius are at level 2, and CHs located within the range of predetermined radius are at level 3. Nodes at level 1 collect information from the environment and send data to CH. CHs at level 2 send collected information to CHs at level 3, which is located within the range of predefined radius. CHs at level 3 communicate directly with BS and send the collected information to BS.

In TLHCLP [20], in contrast to LEACH, every CH has no need to directly communicate with BS. By using the concept of predefined radius, the CHs those have less distance from BS only communicate with BS. As the distance between communicating nodes has been reduced, minimized amplification energy is needed, so the energy dissipation due to transmission has been minimized. The major limitation in TLHCLP [20] is that it does not ensure that all sensor nodes become member of clusters.

Fuzzy-Logic LEACH (FL-LEACH) [21] routing protocol employs fuzzy logic in the selection of CHs and finds the exact number of sensor nodes to be selected as CHs. The system of fuzzy logic is composed by a fuzzifier, a fuzzy inference system, a set of rules, and defuzzifier. Mamdani method is adopted by FL-LEACH for fuzzy inference, which has following four processes. (1) Fuzzification is performed by compelling the crisp values from the input variables and determining their membership values and fuzzy sets. (2) Rule evaluation is done by smearing the fuzzified inputs to the predecessor and resultant fragments of every fuzzy rule. (3) Aggregation is done at the output fuzzy to set the output variable. (4) Defuzzification is done for the calculation of the crisp value of the output. Block diagram of the FL-LEACH protocol is shown in Figure 5. FL-LEACH [21] used fuzzy logic on the basis of two input variables, which are the density of the network and the total number of nodes in the WSN. Equation (6) shows the computation of CHs ($P_{FL}$) percentage.

$$P_{FL} = \frac{\int x \cdot \mu_{FL} \cdot dx}{\int \mu_{FL} \cdot dx},$$

where $x$ is the universe discourse and $\mu_{FL}$ represents the aggregated output for membership. Although FL-LEACH [21] protocols provide best distribution of CHs and enhance the overall network lifetime, but it increases the complexity level of the routing protocol. CH selection criteria are also not efficient; threshold level can be used for the selection of CHs, which may increase the overall lifetime of the network.

In LEACH-GPS [22], all the nodes in the network have added feature of Global Positioning System (GPS). Before startup of the steady state phase and set-up phase, each sensor node shares its location information with BS. BS divides the network area into a number of groups (like A1, A2, A3, ..., etc.) on the basis of location information received from sensor nodes. Sensor node location and probability “$p$” are the two selection criteria for the creation of groups. This
group formation activity does not consume much energy because it is done only once in the lifetime of the network. Each group contains a CH, which is randomly selected in each round to avoid uneven distribution of CH’s in the network. For data communication, multihop strategy is adopted. Each CH sends collected information to the nearest CH, and then, the last CH finally forwards it to the BS. It reduces energy consumption by reducing the distance of radio communication. Figure 6 represents a typical radio communication model. The major deficiency in this protocol is that the CH selection technique is inefficient; it consumes much energy in every round for the selection of CHs. In order to avoid unnecessary CH rotation in every round and save more energy, threshold level should be used for the selection of CHs.

Advanced Zonal Rectangular LEACH (AZR-LEACH) [24] uses static clustering technique and also improved CH selection technique. It logically partitions network into three sections, such as rectangular clusters, advanced clusters, and zones. In it, BS is situated at the middle of the network field. The network field is divided into fixed number of clusters which are nominated as rectangular clusters. The cluster situated around BS is nominated as advanced clusters, and member nodes of these clusters are called advance nodes. The CHs located in advance clusters forward data directly to BS. A group of rectangular clusters forms zone. Every zone contains minimum one advance cluster. The structural diagram of AZR-LEACH is shown in Figure 8.

AZR-LEACH [24] introduces efficient CH selection and reselection strategy. In it, every rectangular cluster, a CH with highest energy is selected as CH. After selection of CH, every node in rectangular cluster forwards their energy status to CH. When CH reaches to a threshold level, then the existing CH nominates its member node as a CH. This CH selection process reduces a lot of energy consumption. As the area of the network field is equally distributed into clusters and zones, the traffic load is equally distributed in the network. The shortcoming in AZR-LEACH [24] is that it is not scalable.

Intraglobal LEACH (IBLEACH) [25] improves network lifetime by balancing the workload in the network.IBLEACH has introduced presteady state phase before the steady state phase. Three phases are involved in each round: setup, presteady, and steady state. Operation process of one
round of IBLEACH is shown in Figure 9. The operation of setup phase is the same as LEACH. In presteady phase, cluster workload (data aggregation + communication to BS) is calculated in the first frame, and then, a member of CH is selected which is capable of handling the aggregation process. If any member node did not have the capability of aggregation then for that cluster aggregation is done by CH, otherwise, every member node performs aggregation task on sensed data and forward aggregated data to CH. The steady state task is distributed into frames. In each frame, member node sends sensed data to aggregator on the basis of time slots allotted to them. The aggregator always keeps its receiver in listening state to receive data from the member nodes. As the cluster load has been shared with member nodes, the lifetime of the CH is increased; as a result, the overall lifetime of the network has been improved. The major shortcoming in IBLEACH [25] is that it has less scalability level. Furthermore, CH selection technique is inefficient, which consumes more energy in the selection of CH in every round. Overall lifetime of the network can be further improved by using threshold level for the selection of CHs, which is introduced by MODLEACH [18].
Centralized Low Energy Adaptive Clustering Hierarchy (LEACH-C) [2, 26] uses centralized clustering mechanism. In LESCH-C, routing protocol strategy for CHs has been improved. The process of this routing protocol consists of steady state phase and setup phase. The steady state phase has been improved, and no change is observed in setup phase as compared to LEACH. In it, every sensor node needs to be equipped with GPS module, because at start of every round, each sensor node sends his location information and energy level to BS. On the basis of received energy level and geographic location of sensor nodes, BS selects CH for every cluster which have higher remaining energy. After selection of the group of CHs, it is broadcasted to the entire network. On reception of CH’s broadcast, the sensor nodes compare their own unique id with the received CHs; if that id match, then the node considers himself as CH and if it did not get any matching id then that node tries to establish a link with its nearest CH. CHs use TDMA schedule [19, 27] to get data from member nodes. LEACH-C uses a deterministic threshold algorithm to gather energy level of sensor nodes and also to maintain record of nodes which has already been selected as CH. Due to centralized CH’s selection process, the overall load on CHs has been reduced, and energy dissipation has also been reduced. On the other hand, energy consumption in process of sharing GPS and energy information with BS has been increased, which degraded the performance of this routing protocol. LEACH-C [26] works well for small scale networks and under perform for larger networks.

Cell-LEACH [28] divides each cluster into seven equal subsections; these subsections are named as cells. Likewise, each cluster has a CH; each cell has a cell head. Sensor nodes in a cell communicate with cell head, and then, cell head is responsible to communicate with CHs. Aggregation is performed by CHs on collected data from cell heads, and aggregated data is forwarded to BS by using multihopping communication method. CH selection criteria have also been improved by using threshold level. Firstly, in the first round, as the energy level of all nodes is at the same level, so cell heads and CHs are selected in a random fashion. At the completion of first round, nodes may be at different energy level. When the energy of a cell head reaches to a threshold limit then cell head sends an alert message to neighbor nodes those contain information about minimum average energy necessary for the selection of cell head. At the reception of the alert message, every neighbor node compares its lingering energy with the average minimum energy necessary for the selection of cell heads and sends back the acknowledgement message to cell head which contains comparison information. At the reception of acknowledgment packets from all neighboring nodes, cell head decides the next cell head on the basis of maximum residual energy. For CH’s selection, the same strategy is adopted as in cell head selection. Figure 10 shows communication architecture of Cell-LEACH.

To reach BS, the shortest path is computed by CHs by maintaining a neighbor table, in which information of all other CHs is available. With the change in CHs, the routing path to reach BS is also updated accordingly. At the end of setup phase, BS broadcasts interest packets to CHs which portrays interest area of the network. CHs forward interest packets to cell heads, and then, cell heads forward it to their respective sensor nodes. Cell-LEACH [28] adopts a proactive approach in the routing of data packets. This approach consumes much energy, as all the time it has to act proactively. In order to achieve energy efficiency, the hybrid approach should be adopted.

In order to support mobility of sensor nodes within the network area, LEACH-Mobile [29] routing protocol is proposed. In it, CHs, as well as member nodes, can change their location, and BS is assumed to stay static. Mobile sensor nodes ensure their membership declaration within a cluster. Like LEACH routing protocol, each round comprises of two phases, steady state phase, and setup phase. It is assumed
that sensor nodes always contain data for the communication and nodes are mobile centric in nature. Each CH forwards data request packet to its member nodes at its applicable time slot and starts a timer for the response packet. If the timer expires, then the CH again forwards request packet to the member node and resets the timer. If again the timer will expire and no reply packet is received, then CH assumes that this particular member node has been moved out from its cluster area, then removes it from the member table, and TDMA schedule is updated. Conversely, the member node that has not received any request packets from CH in the allotted time slot assumes that the CH has been moved to some other location, and now, it is not the member of its cluster. The sensor node starts to broadcast join message packets to become a member of the new cluster. CHs those received join message reply with advertisement packet to the mentioned node. The node then decides to join a cluster on the basis of signal strength. Furthermore, the node sends the joining packet to the CH to declare its membership. After reception of join message, the CH updates its membership table, and also, TDMA schedule is updated. LEACH-Mobile [29] increased the quality of the network by achieving a higher level of data delivery, but the major tradeoff faced is the less power efficiency because packet loss ratio becomes very high if the CH continues in moving state earlier to the selection of the CHs for the next round.

As shown in Figure 11, the network area is divided into three logical stages (S1, S2, and S3), and a forwarder node is installed in the middle of the wireless sensor network field. The sensor nodes in stage S1 communicate with their CHs, and then, CHs forward the received data directly to BS. The sensor nodes in stage S2 communicate directly with the forwarder node, and then, forwarder node sends the received data to the base station. The sensor nodes in stage S3 communicate with their CHs; CHs forward the received data to the forwarder node, and then, forwarder node sends the accumulated received data to the BS. In EE-MRP, an efficient CH selection algorithm is adopted, and unnecessary frequency of reclustering is exterminated. Multiple amplification levels on the basis of distance between communicating nodes are used for the transmission of data between transmitter and receiver, which ensures efficient transmission for both intercluster transmission and intracluster transmission. It is observed that energy efficiency and performance has been improved convincingly.

Large-scale monitoring sensors are distributed around a large region, forming the supervisory control and data acquisition (SCADA) system, which collects monitoring input throughout the working field to centrally monitor and track a specific method. For large-scale Internet of Things applications, though, dependable and energy-efficient data gathering remains a problem. To extend lifespan with a trusted design, a trust-based energy intensive data gathering with unmanned aerial vehicle (TEEDC-UAV) protocol is presented. First, the TEEDC-UAV framework proposes an ant colony-based unmanned aerial vehicle (UAV) trajectory optimization technique that forms the maximum data connectors in the work environment with the shortest trajectory feasible. The proposed TEEDC-UAV framework can effectively find an optimal data gathering trajectory that makes the network’s
energy usage become even most efficient, as demonstrated by ample experimental results. The network lifespan has increased by 48.9% as opposed to the previous approaches. In the meantime, the trust model presented in this study will significantly boost the node safety level recognition performance that reached 91 percent when using only 8% of network life [33].

Ground vehicles and unmanned aerial vehicles (UAVs) are increasingly being reviewed as monitoring instruments which can gather data in the Internet of Things owing to the mobile nature (IoT). In data, gathering methods establishing and improving trust and safety conditions are fundamental and necessary specifications. Author proposed a new system called “Trusted Data Collections through Vehicles in the Smart Internet” (T-SIoTs scheme) that aims to create an integrity environment for social collecting through combining trust vehicles and unmanned aerial vehicles. To test and endorse the T-SIoTs scheme, extensive conceptual studies and analyses have been established. The T-SIoTs structure will increase the protection ratios of 46 percent to 55.60 percent as compared to earlier research. UAVs’ energy usage can be decreased through almost 46.93 percent using the routing algorithm [34].

With the technological increase of artificial intelligence (AI) technologies, social networks (SNs) will be using AI to recover valuable user information and enhance people’s lives. While AI has made significant progress, it still faces numerous data collection problems, including increased data retention and high power consumption. A matrix completion based Sample Point Choice joint Intellectual Unmanned Aerial Vehicle (UAVs) Traject Analysis (SPS-IUTO) framework for data processing is suggested to solve these challenges. The suggested algorithm is used to optimize the UAV’s motion route. The suggested technique can make substantial improvements in terms of power and incomplete features, as demonstrated by the experimental outcomes [35].

In several cases, real-time monitoring systems with green wireless sensor networks (WSNs) are important for preserving better energy efficiency. This article assesses a scenario wherein the green WSNs are used to track the state of the Internet of Things (IoT) that is among the more important forms of energy use in the field of information and communication technologies (ICT). More precisely, we concentrate on optimizing the cluster architecture to reduce aggregation converge cast latency and energy usage in green WSNs [36].
5. Results and Discussion: Comparative Analysis of Hierarchical Routing Protocols for WSN

The key differences among each of the routing protocol to show how each routing protocol is different from each other in terms of its enhancements are shown in Table 1. These protocols are analyzed on the basis of (1) Additional Features w.r.t. LEACH (AF), (2) Distributed/Centralized (D/C), (3) Homogeneous/Heterogeneous (HOM/HET), (4) Threshold (TH) Level used, (5) Path Establishment (PE), (6) Scalability Level (SL), and (7) Hop Count (HC).

(1) Additional Features w.r.t. LEACH (AF) attribute highlights the additional feature used in the enhanced LEACH routing protocol as compared to the basic LEACH routing protocol

(2) The Distributed/Centralized (D/C) scope of the routing protocol indicates the selection procedure of CHs. In case the algorithm works in distributed fashion, then the role of BS is limited to reception of transmitted data from CHs. In it, every node broadcasts hello packets to confirm its availability in the network area. On the basis of proximity information and remaining energy of nodes, the CHs and cluster area are autonomously determined in the sensor network area [37]. Distributed algorithm has four phases which are executed in setup frame. In $(n - 1)S_{rd} + \sum_{j=1}^{n-1}S_j$, where $S_{rd} = \sum_{j=1}^{4}S_j + S_{ss}$ and period of the state is represented by $S_{ss}$. In case the algorithm works in centralized fashion, CHs are selected by BS. For that purpose, Vector Quantization (VQ) technique is used by BS [38]. By using centralized technique, an even distribution of CHs is achieved all over the network. The average data signal received at BS is more than the distributed protocol. The major drawback in centralized algorithm is that it consumes more energy at the time of startup

(3) The Homogeneous/Heterogeneous (HOM/HET) attribute represents competencies nature of sensor nodes. The homogeneous network has nodes with same indistinguishable properties [39]. The primary goal of homogeneous networks is that all nodes expire at the same time and will stay alive throughout the network lifetime. LEACH [1, 2] adopts periodically random selection of CHs to perform the load balancing and equal energy consumption of each sensor node. In homogeneous networks, as any node can become CH, so it is indispensable for every node to have the capability of transmission at stretched distance and also have the capability of MAC coordination with routing knowledge. Since all nodes have the capability to act as CHs, catastrophe of any node may not affect the whole network seriously [40]. Heterogeneous (nonhomogeneous) sensor network consists of nodes which may have two or more types of functionalities [39, 41]. For instance, the authors in [42] have proposed a sensor network solution which consists of two types of nodes, type 0 and type 1. Nodes belong to type 0 group acts as pure sensor nodes, and type 1 group nodes perform the duty of CHs. In case of a single hop routing protocol, as the CHs are fixed, CHs those are far from BS consume more energy as compared to CHs located near to the BS; this results in CHs located far from BS become dead earlier, and a portion of the network become isolated. The nonhomogeneous scheme performs better in case of multihop routing protocols. The major benefit in nonhomogeneous networks is that the hardware complexity is just limited to the CHs. As compared to homogeneous WSNs, the heterogeneous system is less rigorous to the failure of nodes [39]

(4) The Threshold (TH) level attribute shows either threshold level is used or not for the selection and rotation of CHs. If TH level is not used, then in every round, routing protocol is responsible for the selection of CHs. In every round, a lot of energy is consumed in the election of CHs. On the other hand, routing protocols which use TH level for the selection of CHs, in every round CH, remain unchanged until energy level of CHs reached up to TH level. The routing protocols which use TH level for the selection of CHs consume less energy as compared to routing protocols which do not use TH level

(5) The Path Establishment (PE) attribute of the routing protocol represents the routing strategy adopted for the communication of data from sensor nodes up to the final destination BS. There are three modes of operation for path establishment which are proactive, reactive, and hybrid, contingent on how source to destination route is established [43, 44]. In proactive routing protocol, route computation is done in advance before they are categorically desired and make sure to maintain an updated topology table on the continuous intervals. It is considered that all network paths are known to every node, and a connected graph is maintained proactively. The benefit of the proactive approach is that communication starts without any initial delay, but constant traffic overhead increases energy consumption and reduces overall network lifetime. In the reactive protocol, route formation is done on-demand. If any node tries to send data to a destination node and if that route is not previously used in the network lifetime, then the routing protocol will have to establish that specific route before the communication can be started. This technique reduces traffic overhead and increases overall network lifetime, but there is an initial delay at the start of the communication. Routing protocols those use an amalgamation of both proactive and reactive protocols is known as hybrid protocol [43]

(6) The Scalability Level (SL) attribute in the routing protocols of WSNs has critical value because of very
<table>
<thead>
<tr>
<th>WSN routing protocol</th>
<th>AF</th>
<th>D/C</th>
<th>H/NH</th>
<th>TH level used</th>
<th>PE</th>
<th>SL</th>
<th>HC</th>
<th>EE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LEACH</td>
<td>—</td>
<td>Distributed</td>
<td>Homogeneous</td>
<td>No</td>
<td>Hybrid</td>
<td>Limited</td>
<td>Single hop</td>
<td>—</td>
</tr>
<tr>
<td>MODLEACH</td>
<td>(i) Uses threshold level for CH selection</td>
<td>Distributed</td>
<td>Homogeneous</td>
<td>Yes</td>
<td>Hybrid</td>
<td>Limited</td>
<td>Single hop</td>
<td>28%</td>
</tr>
<tr>
<td>MODLEACHHT</td>
<td>(i) Uses hard threshold</td>
<td>Distributed</td>
<td>Homogeneous</td>
<td>Yes</td>
<td>Hybrid</td>
<td>Limited</td>
<td>Single hop</td>
<td>83%</td>
</tr>
<tr>
<td>MODLEACHST</td>
<td>(i) Uses soft threshold</td>
<td>Distributed</td>
<td>Homogeneous</td>
<td>Yes</td>
<td>Hybrid</td>
<td>Limited</td>
<td>Single hop</td>
<td>122%</td>
</tr>
<tr>
<td>TLHCLP</td>
<td>(i) BS installed at the center of network area</td>
<td>Distributed</td>
<td>Homogeneous</td>
<td>No</td>
<td>Hybrid</td>
<td>Good</td>
<td>Multihop</td>
<td>17%</td>
</tr>
<tr>
<td>FL-LEACH</td>
<td>(i) Apply fuzzy logic to find exact number of CHs</td>
<td>Distributed</td>
<td>Homogeneous</td>
<td>No</td>
<td>Hybrid</td>
<td>Limited</td>
<td>Single hop</td>
<td>51%</td>
</tr>
<tr>
<td>LEACH-GPS</td>
<td>(i) Nodes equipped with GPS functionality</td>
<td>Distributed</td>
<td>Homogeneous</td>
<td>No</td>
<td>Hybrid</td>
<td>Very good</td>
<td>Multihop</td>
<td>137%</td>
</tr>
<tr>
<td>A-LEACH</td>
<td>(i) Introduces load-sharing technique</td>
<td>Distributed</td>
<td>Homogeneous</td>
<td>No</td>
<td>Hybrid</td>
<td>Good</td>
<td>Multihop</td>
<td>20%</td>
</tr>
<tr>
<td>AZR-LEACH</td>
<td>(i) It is assumed that dimensions of network and BS position are known</td>
<td>Distributed</td>
<td>Homogeneous</td>
<td>No</td>
<td>Hybrid</td>
<td>No</td>
<td>Multihop</td>
<td>40%</td>
</tr>
<tr>
<td>IBLEACH</td>
<td>(i) It transfers aggregation load to sensor nodes</td>
<td>Distributed</td>
<td>Homogeneous</td>
<td>No</td>
<td>Hybrid</td>
<td>Limited</td>
<td>Single hop</td>
<td>60%</td>
</tr>
<tr>
<td>LEACH-C</td>
<td>(i) Nodes equipped with GPS/tracking facility</td>
<td>Centralized</td>
<td>Homogeneous</td>
<td>No</td>
<td>Hybrid</td>
<td>Good</td>
<td>Single hop</td>
<td>13%</td>
</tr>
<tr>
<td>Cell-LEACH</td>
<td>(i) Clusters further divided into cells with cell head and CH</td>
<td>Distributed</td>
<td>Homogeneous</td>
<td>Yes</td>
<td>Reactive</td>
<td>Very good</td>
<td>Multihop</td>
<td></td>
</tr>
<tr>
<td>LEACH-Mobile</td>
<td>(i) Nodes/CHs mobility</td>
<td>Distributed</td>
<td>Homogeneous</td>
<td>Yes</td>
<td>Hybrid</td>
<td>Very good</td>
<td>Single hop</td>
<td>155%</td>
</tr>
<tr>
<td>EE-MRP</td>
<td>(i) Hybrid communication mechanism adopted</td>
<td>Distributed</td>
<td>Homogeneous</td>
<td>Yes</td>
<td>Hybrid</td>
<td>Very good</td>
<td>Single hop</td>
<td></td>
</tr>
</tbody>
</table>
high numbers and moderately high density of the sensor nodes. A good scalable network has the competence to survive and perform well under growing and augmented workload. Arbitrary network model and random network model are the two network models on which scalability can be analyzed [45]. In arbitrary network model, the location of sensor nodes in the network area is not restricted; within that area, each node has the capability of reception and transmission of data at any time. Spatial-temporal transmission scheduling policy or amount of transmission powers are also not restricted. In random network model, sensor nodes are evenly distributed in the network area, follows random traffic pattern, and when more sensor nodes are added in the network area than fixed power of transmission can be attuned to guarantee connectivity of the network [46].

(7) Due to its significant usage in crucial environmental monitoring and target detection in both civilian and military regions, wireless sensor networks (WSNs) are gaining interest of researchers. To optimize the routing process, load-balancing, and maximize the lifetime of the network, the large-scale implementation of energy-constrained WSNs and requirement of multihop traffic aggregation need optimal coordination of routing protocol. Therefore, the WSN network included the many routing protocol that provide the reliability to the data packet delivery from end to end. In the reliability, the overall amount of information was obtained at the sink. It presents a significant method for calculating the efficiency of the algorithm. In our study, we also thoroughly explain the reliability factor in the node deployment subheading. For the efficient packet and data transmission, the reliability matters a lot because it enhances the life of the network. The LEACH protocol works very effectively to maintain the reliability of the network and transmit the data packet at the destination end safely [32].

In a particular environment, three major levels are involved to investigate the scalability with respect to previously known factors [45]. Firstly, absolute scalability is measured with respect to existing characteristics (like environment, independent parameter, and primary metrics). For instance, if the value of an independent parameter increased up to infinite level, then the efficiency of the network may not lead to evaporate. In order to evaluate whether a routing protocol satisfies the conditions of absolute scalability, all metrics should be precisely defined. Secondly, if absolute scalability did not satisfy (in most of the cases, it may not be feasible to ration absolute scalability for a routing protocol in given milieu), the optimal scalability may be achieved. A protocol is considered optimally scalable with respect to the given parameters (environment, independent parameter, and primary metrics), if none of other protocols may be more scalable with respect to the same parameters. Thirdly, relative weak scalability may be achieved. In some cases, it can be a possibility that a parameter can grow larger but cannot grow more than a certain level. In precise, it may not be possible for a parameter to cross some specific limit. If that situation occurs, then scalability level is considered as relatively weak.

(8) The Hop Count (HC) attribute of a routing protocol in WSNs has its effect on energy depletion of each sensor node and CH. The short-hop strategy is considered more energy efficient because sensor nodes consume less communication power for shorter distances. The reason behind that is signal attenuation value of which power function (Equation (7)) is proportionally equivalent [47]. Equation (7) is used to define the minimum energy requirement to establish communication between source and destination over a distance of d meters, where $P_0$ is the signal power received at destination node and $a$ represents value of path loss exponent, normally value of $a$ varies from 2 to 5 [48].

$$P(d) = P_0 \times \left(\frac{d_0}{d}\right)^a.$$  

As reception cost is also not negligible, there is the shortest range of distance from source to destination for which optimal solution for communication can be the direct transmission. The reason behind that by using multihopping strategy, it may not save enough energy because of the receiver energy cost. However, if the distance between source and destination is large enough, then a two-hop strategy is beneficial, and if the distance between transmitting and receiving devices is further increased, then the 3-hop communication can provide optimal results and so forth. Consequently, in order to calculate energy efficiency of a routing protocol for a given topology, then energy efficiency is compared on both single hop and multihop strategies on the basis of distance between source and destination BS [47].

(9) The energy efficiency with respect to LEACH (EE) attribute provides the clear picture that how much improvement in energy efficiency in the enhanced LEACH routing protocol as compared to the original LEACH routing protocol is achieved.

MODLEACH [18], MODLEACHST, and MODLEACHHT use threshold level for the selection of CHs, and in each round, CHs are not changed, as compared to LEACH [1]. CHs are selected in every round. These protocols use single hop communication and homogeneous devices and have limited scalability. IBLEACH [25] uses single hop communication and introduces load sharing by distributing the tasks of aggregation and communication. Among all discussed single hop routing protocols, MODLEACHST produces the best results, and the second best is IBLEACH, which uses load sharing technology. It can be determined that by using both improvement techniques of IBLEACH and MODLEACHST at the same time, a better routing protocol may be introduced, and more energy efficiency can be achieved.
FL-LEACH [21] optimized the CH selection mechanism by adopting the fuzzy-logic system. It enhanced energy optimization by finding the exact number of CHs required for the whole network area. FL-LEACH [21] achieves an overall better level of energy efficiency, but scalability level is limited, and its complexity level is also high. The most successful achievement of FL-LEACH is high level of scalability. As the exact location of each node is marked and CHs are selected uniformly and distributed realistically on the network field and by using multihop routing, the distance between CHs and BS is reduced. AZR-LEACH [24] divides the network area into zones and used multihoping strategy for communication and achieved an average level of energy efficiency. In EE-MRP [30], by using efficient cluster head mechanism and by use of hybrid communication mechanism, noticeable energy efficiency has been achieved.

There is need of analysis of LEACH and other routing protocols for their performance of the energy efficiency. Here, the energy efficiency means that the management and control of energy while sending data by nodes within the network. The energy efficiency makes sure that nodes should use less energy for the data communication and enhance the network life. With respect of this analysis, Figure 12 compares LEACH and other variants of LEACH routing protocol in terms of energy efficiency. It is observed that all extended variants of LEACH are comparatively more energy efficient than LEACH routing protocol. EE-MRP, MODLEACHST, and LEACH-GPS have been noticeably improved up to 155%, 122%, and 137%, respectively, as compared to LEACH routing protocol. TLHCLP, MODLEACH, and A-LEACH have got a little bit improvement up to approximately 25%. MODLEACHHT, IABLEACH, and FL-LEACH have been noticeably improved over 50%. The Y-axis is explaining the performance score of each routing protocol from min to max performance score. The LEACH protocol is the least protocol for the energy efficiency in the WSN network, and EE-MRP protocol is more efficient protocol for energy efficiency. The Y-axis of Figure 12 is helping the research community to adopt the best protocol for energy efficiency while working in WSN network as well as it is also illustrating the working directions in the protocols to make them more efficient for energy in WSN.

6. Routing Issues and Challenges

Due to the changing nature of sensor network, as compared to conventional network, it is necessary to consider different issues and challenges those affect WSN routing protocol either directly or indirectly. These factors can be categorized as node deployment, scalability, network dynamics, routing protocols, energy efficiency, congestion, coverage, security and privacy, and QoS [49, 50].

6.1. Node Deployment. In WSNs, mostly node deployment is application dependent. It can be done in two different ways, either in manual fashion or in randomized fashion [51]. In case of manual deployment, as the placement of sensor nodes is known so data transmission is done by predetermined paths, whereas in case of random deployment, the sensor nodes are randomly scattered in the network field by creating an ad hoc routing infrastructure. In random deployment, if the uniform distribution of sensor nodes is not achieved, then it is necessary to perform optimal clustering to achieve energy efficient network routing operation. The fault tolerance or reliability factor $R_i(t)$ of a sensor node with the Poisson distribution to find the nonfailure probability within $(0 : t)$ time interval is represented by Equation (8) [12, 51].

$$R_i(t) = \exp(-\Delta_i t),$$  

where $\Delta_i$ represents failure rate of “$t$” sensor nodes and “$i$” represents the time period.

6.2. Scalability. A routing protocol is considered scalable if it works well effectively, if number of sensor nodes in the network field are increased, and also effectiveness of the network increases when the hardware resources are increased [41, 52]. In WSNs, scalability can be measured in a number of
dimensions. The quantity of sensor nodes installed in the network field may be up to 100s or 1000s or much more. The routing protocol should be workable in the network field in which large numbers of sensor nodes are deployed.

6.3. Network Dynamics. In most of the network architectures, it is assumed that sensor nodes are fixed, but in some of the application mobility of both BS and sensor nodes is necessary [53]. Transmission of routing messages to a moving node or from a moving node is more difficult and challenging as compared to stationary source and destination nodes. In case of moving nodes, in addition to energy efficiency and bandwidth, it is also more challenging to maintain route stability. In most of the applications, dynamic events require periodic reporting, and accordingly, significant traffic is generated to be routed towards BS.

6.4. Routing Protocols. Currently, a lot of routing protocols for WSNs have been developed and proposed by researchers, and most of these protocols are based on the assumptions and simulations. On the basis of communication architecture and topology, these protocols are divided into different categories such as hierarchical, data-centric, and location-based QoS [54]. A lot of research in WSNs is dependent on assumptions and provide good results which are limited up to the simulations but did not provide the same results on the real devices. For instance, as the distance between communicating nodes is directly proportional to the quality of signals and less distance consumes less energy and minimized energy dissipation. This is very accurate mathematically and theoretically, but in reality, the environmental parameters like obstacles and weather conditions may directly influence the signal strength, which may badly vitiate energy efficiency and overall network lifetime.

6.5. Energy Efficiency [55]. The minimization of the energy consumption is the primary goal of energy optimization, and it is considered to be one of the difficult tasks in WSNs. Normally, WSN is scattered and arranged in an unattended location. In most of its applications, WSN is deployed only once, and it is impossible to replace nodes or even physically monitor sensor nodes, specifically in chemical industrial areas and in armed scuffle fields. In the last few years, a lot of work has been done in the field of WSNs, and the major focus of most of the research was to optimize the energy efficiency and prolong the overall lifetime of the network. The reasons behind inefficiency in energy consumption may include inefficiency of routing protocol, topology changes at runtime, mobility, nodes deployment, and difference of distance between nodes/CHs and BS. While dealing with the real-time applications, it is necessary for all nodes to remain alive in the network; if any node becomes dead due to energy dissipation, then the entire segment of the network becomes isolated, which is not affordable in real-time scenarios. In the last few years, the researchers pay maximum attention to the improvement of the critical issue (routing protocol) and proposed different energy efficient routing solutions, but still, there is a lot of space for the improvement of routing protocols.

6.6. Congestion. As WSNs generally set up by wireless ad hoc networks, in which routing protocols are running on CHs/nodes and in maximum cases, these are working in multihopping scenarios. Nodes in these networks have very limited resources and are deployed in unattended environments. In recent years, researchers have provided ample amount of weight to the congestion issue, but there are still a lot of space to optimize that issue. Congestion in the network may lead to many problems which include unselective drop of data packets. It is possible that some packets with low importance may be delivered, and packets of higher importance may plunge. Network congestion also increases energy depletion because saturation of links increased with the increase in congestion.

6.7. Coverage. In WSNs, the area of deployed sensor nodes in which all nodes can communicate is considered as coverage, and it provides the surety to accurately receive and transmit data. WSNs are deployed in denser mode because it is an energy constraint network with limited battery life and also have limited computational resources. This form of arrangement of sensor nodes is essential for the reason that nodes/CHs cannot rely on few routes; on the other hand, if network is not densely deployed, then there are chances of disconnection of route from BS if any node becomes dead or its battery power reaches to the end. In such situations, there can be a part of network or some clusters may be isolated from the other network and did not have any route to the BS. The best case for the coverage is to achieve the goal to discover a route with the highest observance, and so any recognizable activity can be captured by the nodes [56]. The study reported that an algorithm for the connected p-percent coverage problem in wireless sensor networks to enhance network life. Before the suggested solution, the authors thoroughly examined the coverage issues in the WSN. In this study, for cover p-percent coverage, the proposed method employs a learning automaton to pick the best dominator and dominate sensor nodes. The work reveals that the pDCDS algorithm performed much better than the existing algorithms [57, 58].

6.8. Quality of Service (QoS). Circumscribed latency in the delivery of data is also a compromised situation for the applications which are time-constrained. Nevertheless, in most of the WSN applications, energy conservation, on which overall network lifetime is dependent, is given more weightage as compared to quality of service for the transmission of data [30]. With the passage of time, the energy of sensor nodes is exhausted, and the quality of data transmission is degraded to achieve maximum lifetime. Energy efficient routing protocols should be improved to achieve maximum quality of service with minimum energy consumption. A study has been conducted on the QoS; in this work, the author revealed that one of the most difficult difficulties in wireless sensor networks (WSNs), particularly regarding surveillance systems, is QoS routing. Because of the higher packets drop and energy usage, the multihop data transfer in WSNs necessarily requires secure links for end-to-end transmitting data. To reduce the delay and enhance the energy usage with respect
of QoS in the WSN network, the author proposed the model. They proposed the distributed learning automaton algorithms for it. The results show that the RRDLA algorithm strikes a strong balance between a variety of QoS constraints, including end-to-end delay and energy usage [32].

7. Conclusion

WSNs have limited resources; however, energy resource is most important resource among all resources, because network lifetime is dependent on the life of sensor nodes and sensor nodes stay alive on the basis of energy. It is observed by detailed literature review that the most of the power energy has been utilized during communication process. An efficient routing protocol plays the most important role in optimization of energy consumption in the communication process. In this survey paper, a well-known hierarchical routing protocol LEACH has been described along its extended versions as well. Apart from the considerable improvement in power consumption by LEACH as compared to other nonhierarchical routing protocols, there are a lot of opportunities to improve energy efficiency by improving the routing protocol. It is observed that by improving the technology of sensor nodes, energy efficiency can also be improved. For instance, in LEACH-GPS, sensor nodes are equipped with GPS feature, and energy efficiency has been improved up to 137% as compared to LEACH routing protocol. Most of the routing protocols have improved routing efficiency by adding multihoping (decreasing communication distance) and setting threshold levels for the selection of CHs. EE-MRP has significantly improved energy efficiency up to 155% by adopting efficient techniques. Most of the discussed LEACH variants have noticeable improved energy efficiency and overall network lifetime, but still, there is a lot of space for improvement in routing protocol in terms of energy efficiency and performance. In our study, we have highlighted shortcomings in every routing protocol, which can guide to design an optimized and energy efficient routing protocol.

Data Availability

The data are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was partially supported by the Natural Science Foundation of China (NSFC) under grant no. 62072217.

References


