**Abstract**— Recently wireless sensor networks (WSN) became an interesting topic because of its increasing usage in many fields; medical systems, environment monitoring, military applications and video surveillance. Usually sensors are placed in the desired locations to gather information frequently and then transfer it to the observers. WSN consists of a collection of application specific sensors, a wireless transceiver and a simple general purpose processor. In heterogeneous wireless sensor network, researchers found many challenging issues including the limited energy, the efficient usage of the energy, and the problem with the hierarchy of the network as imbalance network. Many studies indicated that the node clustering is a promising solution for such issues. Clustering has been shown to increase the efficiency of the energy consumption where clusters are formed dynamically with neighboring sensors and the power is assumed to be distributed equally among nodes. One of the nodes is considered as the cluster head that is responsible for transferring data among the neighboring sensors. In this work, we propose a modification based on SEP protocol. EM-SEP aims to prolong the stable period of the sensor network by maintaining balanced energy consumption. This means that we choose the advanced nodes to become cluster heads more often than the normal nodes as the case with SEP protocol. Furthermore, EM-SEP takes in account the number of nodes that are associated with each cluster head. Another important enhancement of EM-SEP protocol that if there are more than one sensor available to be a cluster head at certain round, we choose the sensor with highest energy.

**Keywords**— Wireless Sensor, Clusters, SEP, Energy Consumption, EM-SEP

I. INTRODUCTION

WSN is an emerging technology that helps to bring enormous changes in data collection, processing and dissemination for different environments and applications [1, 2]. Nodes have limited power batteries which shorten its lifespan and battery replenishment is often not practical. Increasing the network lifetime depends on management of the sensing nodes energy resource. The wireless sensors can be deployed in large different places to form wireless sensor networks for a wide variety of purposes. Such deployment can be random that can be dropped from an airplane or installing fire alarm sensors in a building [3]. The lifetime of the sensor is limited by power that every sensor has in the network and will determine the validity of the sensing task. In [4], distributed cluster-based routing protocol for sensor network called Low-Energy Adaptive Clustering Hierarchy (LEACH) is proposed. The authors developed Low-Energy Adaptive Clustering Hierarchy (LEACH) protocol which works for micro sensor networks that link the ideas of energy-efficient cluster-based routing and media access together with application-specific data aggregation to result in better performance in terms of system lifetime, latency, and application-perceived quality. Leach contains a distributed cluster which helps to organize a large numbers of nodes. Also, it includes the algorithms used for adapting clusters and rotating heads positions to distribute equally the power between all the nodes in the network. Also, Leach enables the signal distribution among nodes in order to increase the communication performance. It is shown that the Leach increases the lifetime of the system by an order of magnitude compared to other approaches.

To design an efficient protocol for wireless sensor networks, the following properties of wireless sensor network must be taken into consideration:

A) Architecture of the deployment: wireless sensor networks consist of hundreds of nodes deployed in different environments. Such environments might not be easy to reach weather they are dangerous or located in distant places. In such cases, sensors enable users to communicate and exchange data with other users even if there is no existing network infrastructure.

B) The lifetime of the system: the network should remain alive as long time as possible to have better performance since
recharging or changing the batteries might not be feasible. All nodes must be designed to prolong the lifetime of the system for better efficiency.

C) Latency: latency means how much time the data takes to be retrieved from the other nodes. In wireless sensor network, the time is sensitive and data must be received within certain time to have scalable system.

D) Quality: the user does not need redundant information because it is connected to information of neighboring nodes and user needs only the information from a higher-level description of events occurring in the monitored environment.

In general, the quality of the networks depends on the quality of the data which means the protocols must be designed guarantee a specific level of quality of the data.

The rest of the paper is organized as follows. Section 2 presents some of the related work. Section 3 gives an overview of the main characteristics of the Leach protocol architecture and how it works. Section 4 we addresses the Optimal Clustering. Section 5 presents SEP protocol and the stable election protocol based on the energy of normal and advanced sensors. Section 6 introduces performance measures that we have used. Section 7 provides the design details of EM-SEP protocol. Section 8 shows a comparison between SEP and EM-SEP. Section 9 provides the conclusions.

II. RELATED WORK

Clustering algorithms for wireless sensor networks can be categorized in two broad types : (1) probabilistic algorithms that run in synchronized rounds which are mostly used for static networks [4], and weight-based algorithms [5], which are used in dynamic networks[6]. In this section, we explain the weight-based clustering algorithms.

Weight-based algorithms can be constructed as overlapping cluster where every node can belong to more than one cluster. Also, it can be constructed as disjoint cluster where a node only belongs to one cluster. In the overlapping cluster, the used algorithms choose set of nodes, termed gateway or border nodes. Each set is connected to at least two adjacent clusters. In [7], the LCA clustering algorithm is discussed where every node chooses its own cluster head from the neighboring cluster head with the smallest ID and any node that can hear two or more cluster heads is called as a gateway.

Disjoints clusters can be constructed where each node must share its information with its cluster head (such as the sensed data or sensor id). The cluster head uses this information of the node to select the cluster head nodes some protocols make their selections based on the completion of the information for certain number of hops or for the whole network. In[8], the max-min D cluster algorithm uses the d-hop information for cluster head selection. After finishing the selection process, the nodes are at most d hops away from the cluster head.

In [6], DMAC is presented that partitions the nodes of a fully mobile network (ad hoc network) into clusters, thus giving the network a hierarchical organization. Nodes are grouped by following new weight-based criteria that allows the choice of the nodes that coordinate the clustering process based on node mobility-related parameters. Like DSA, nodes choose the role based on the information from the neighboring hop. However, DMAC is not implemented in different phases. Every node responds locally to the surrounding topology. The topology can be represented by addition and deletion of links to its neighboring nodes. The cluster head is the node with highest weight between its unassigned neighbors.

In [9], the authors generalized DMAC as G-DMAC, where a new node that joins the cluster head with the highest weight in its one-hop neighborhood (like DMAC). The topology is changeable but the node is still considered as member of this cluster head v as long as there is no other neighboring cluster head u with weight w(u) > w(h) + h, given the parameter h ≥ 0. Another parameter k is introduced that defines the maximum number of cluster head neighbors that a cluster head is allowed to have.

The C4SD clustering algorithm[10] is proposed for mobile network where the node’s capability and the degree of dynamics represent the weight. In such a network, each node chooses the neighbor with the highest capability grade as parent. If such a node does not exist, this node itself is considered as a cluster head. The result of this algorithm can be represented by disjoint clusters where cluster head form an independent set.

In[8], a new algorithm called Tandem is proposed for spontaneous clustering of mobile wireless sensor nodes facing similar context (such as moving together) where every node runs a shared-context recognition algorithm, which provides a number on a scale, representing the confidence value that two nodes are together. Every node periodically makes sure it shares the same context with its neighbors. The selection of cluster heads is weight-based: the node with the highest weight among its neighbors with which it shares a common context is considered as a cluster head. A regular node joins the cluster head with the highest weight and the one that shares a common context with it.

III. LEACH PROTOCOL ARCHITECTURE

Leach protocols work in wireless micro sensor networks to monitor a remote environment. Since data is correlated among neighboring nodes in the network, the user doesn't need the redundant data but requires a high-level function of the data that describes the events occurring in the environment [4, 11]. The data signal is strongly correlated among nodes specially the one close to each other. The clustering aspect is considered as the main principle for Leach. This enables the data to be transferred and processed among nodes in the same cluster which leads to decreasing the data set that must be transferred to another user. Data aggregation techniques is an option to combine different correlated information signals to be smaller data sets that results effective and efficient data [12]. In addition, this technique reduces the amount of transmitted data between the cluster and the base station.
In Leach, the cluster is formed locally by the nodes and one of the nodes is elected to be the head of the cluster. In each round there is a new selected head for the cluster which means distributing the load equally among the nodes in the same cluster. The rest of the nodes that are not cluster head transfer their data to the head and they implement a signal procession functions on the data and transfer the data to the base station. The cluster head node consumes more power than other nodes since is responsible for the commutation in the cluster. The power load should be distributed evenly between the nodes in the same cluster. If the heads are selected in advance during the system lifetime, this causes wastage of power and that node loses its total energy in short time. If the node loses its energy, then the node is no longer operational and the non-head nodes in the same cluster will not be able to communicate with each other. For this reason, the selection of the head should not be fixed throughout the lifetime of the systems and should be randomly selected to reduce the energy consumption of the sensors in the network.

LEACH operates in rounds. Every round starts with a set-up phase when the clusters are organized, then a steady-state phase when the information is transferred between the nodes and the head and between the head and the base station. The nodes are considered to be homogeneous when all the nodes have the same initial power. However, in Leach protocol every node will be a head for the cluster every \( \frac{1}{p_{opt}} \) round which is called as epoch for the cluster wireless network\[13\].

It is assumed that every node can be a cluster head with a probability of \( p_{opt} \). In every round there is a new cluster head instead of the one in the previous round. The set \( G \) contains the rest of the nodes that are non-head nodes to keep the cluster head constant for each round. Every round \( r \), a new head is selected when every node \( s \in G \) selects a number randomly in \([0,1]\). The node is considered as a head for the cluster if that number is less than the threshold \( T(s) \).

\[
T(s)=\begin{cases} 
\frac{1}{p_{opt}} & \text{if } s \in G \\
0 & \text{otherwise}
\end{cases} \tag{1}
\]

IV. OPTIMAL CLUSTERING

Many studies have been discussed the optimal clustering with different results. In \([7]\), the authors proposed a new distributed randomized clustering algorithm for organizing the sensors in the wireless sensor network in clusters while maintaining the hierarchy of cluster heads. It is assumed that the communication environment is contention-based and without error. The results show that total energy usage is reduced with increasing the number of levels in the cluster while transmitting the data to the data processing center. Another algorithm for optimal clustering is proposed in\[14\] to organize the sensors in a wireless sensor network into clusters to minimize the consumed power in the system while exchanging the data in the network. The authors evaluate the number of cluster heads and cluster diameters that minimize the energy consumed in the network. The power used to collect information at a central node from all sensors in a clustered network is found to be significantly lower than the Max–Mind cluster algorithm. Also, the results show that energy used in the cluster with five levels can be as low as 25% of the power consumption in a non-clustered network over a large area. This clustering is optimal when energy is evenly distributed among all sensors and the total energy consumption is minimal.

Fig. 1\[7\] shows the radio energy dissipation model and to get acceptable Signal-to-Noise Ratio (SNR) in transmitting an \( L \) bit message over a distance \( d \), the energy expended by the radio is shown below:

\[
ETx(l, d) = \int L \cdot E_{elec} + L \cdot _{fs} \cdot d_2 \text{ if } d < d_0 \\
L \cdot E_{elec} + L \cdot _{fs} \cdot d_2 \text{ if } d < d_0
\tag{2}
\]

Where \( E_{elec} \) is the energy consumed per bit to transmit data between sender and receiver, \(_{fs}\) and \(_{mp}\) depend on the transmitter amplifier model we use, and \( d \) represents the distance between the sender and the receiver.

\[
d_0 = \sqrt{\frac{E_{fs}}{E_{mp}}}
\tag{3}
\]

We assume an area \( A = M \times M \) square meters, \( n \) is the number of nodes which are randomly distributed in that area and the sink is in the center of the field. Also, it is assumed that the distance between the sink and any node is \( \leq d_0 \). The energy consumed in the head node for each round can be represented in Equation (4):

\[
ECH = L \cdot E_{elec}(\frac{d_0^2}{4}) + L \cdot E_{DA} + L \cdot _{fs}d^2 \text{ toBS}
\tag{4}
\]

Where \( k \) represents the number of clusters, \( E_{DA} \) is the processing (data aggregation) cost of a bit per signal, and \( d_{toBS} \) is the distance between the cluster head and the sink.
The energy used in a non cluster head node can be computed as:

\[ E_{\text{nonCH}} = L \cdot E_{\text{elec}} + L \cdot f_s \cdot d_{\text{toCH}} \]  
(5)

Where \( d_{\text{toCH}} \) represents the distance between a cluster member and its cluster head. Assuming that the nodes are uniformly distributed, it can be shown that:

\[ E_{d_{\text{toCH}}} = \int (x^2 + y^2) \rho(x,y)\,dx\,dy \]  
(6)

where \( \rho(x,y) \) represents the node distribution.

The energy consumed in a cluster per round can be computed as in the following formula:

\[ E_{\text{cluster}} \approx E_{\text{CH}} + \sum_{k=1}^{n} E_{\text{nonCH}} \]  
(7)

The total energy in the network is:

\[ E_{\text{tot}} = L \cdot (2n \cdot E_{\text{elec}} + n \cdot E_{\text{EDA}} + f_s \cdot k \cdot d_{\text{toBS}} + \sum_{k=1}^{n} \frac{M^2}{2\pi \cdot k}) \]  
(8)

The optimal number of constructed clusters is equal to:

\[ K_{\text{opt}} = \frac{n}{\sqrt{2\pi \cdot d_{\text{toBS}}}} = \frac{n}{\sqrt{2\pi \cdot 0.765}} \]  
(9)

Because the average distance from a cluster head to the sink as in [7]:

\[ E_{d_{\text{toBS}}} = \int_A \sqrt{x^2 + y^2} \, \frac{1}{A} \, dA = 0.765 \frac{M}{2} \]  
(10)

The optimal probability of a node to become a cluster head, \( p_{\text{opt}} \), can be calculated as follows:

\[ p_{\text{opt}} = \frac{K_{\text{opt}}}{n} \]  
(11)

V. SEP PROTOCOL

SEP protocol proposed a new solution which is called the stable election protocol. SEP is based on the energy of normal and advanced sensors. SEP approach proposes two different (Popt) weighted optimal election probability; one for the normal nodes and the other for the advanced nodes. SEP defines the (Pnrm) which is the weighted election probability for normal nodes, and the (Padv) the weighted election probability for the advanced nodes. Therefore, the weighted probabilities of normal and sensors are formed as following respectively.

\[ P_{\text{nrm}} = \frac{P_{\text{opt}}}{(1 + a + m)} \]  
(12)

\[ P_{\text{adv}} = \frac{P_{\text{opt}}}{1 + a + m} \]  
(13)

SEP define two different thresholds as well. \( T_{\text{nrm}} \) is the normal nodes threshold. On the other hand, \( T_{\text{adv}} \) the threshold for advanced nodes. Hence, the normal and advanced nodes equation threshold consequently is

\[ T_{\text{nrm}}(r) = \left\{ \begin{array}{ll}
\frac{P_{\text{nrm}}(r)}{(1 - P_{\text{nrm}}(r) \mod \frac{1}{\text{nrm}})}, & \text{if } S_{\text{nrm}} \in G' \\
0, & \text{otherwise}
\end{array} \right. \]  
(14)

\[ T_{\text{adv}}(r) = \left\{ \begin{array}{ll}
\frac{P_{\text{adv}}(r)}{(1 - P_{\text{adv}}(r) \mod \frac{1}{\text{adv}})}, & \text{if } S_{\text{adv}} \in G' \\
0, & \text{otherwise}
\end{array} \right. \]  
(15)

where \( r \) is the current round, \( G' \) is the set of the normal nodes that have not become cluster heads within the last \((1/P_{\text{nrm}})\) round.

VI. PERFORMANCE MEASURES

We have identified the following parameters to measure the performance of the clustering protocols:

1. **Stability period**: known as a stable region and it refers to the duration of time when the network is operational.
2. **Instability Period**: known as unstable region and it refers to the time between the death of the first and last node.
3. **The lifetime of the network**: the time between the beginning of the network and the death of the last node.
4. **Number of cluster head per round**: refers to the number of nodes that can communicate with sink from their cluster members.
5. **Alive node every round**: refers to the number of nodes that still have not used their total energy.
6. **Throughput**: refers to the rate of the data sent in the network, the data exchanged between the nodes and cluster head and between the cluster head and the sink.

There is a relationship between the lifetime of system and reliability. If there is at least one node alive, we can receive an update regarding the sensor’s field but these updates may not be reliable because we cannot guarantee that the existence of one cluster head per round in the last round.

VII. EM-SEP: AN EFFICIENT MODIFIED STABLE ELECTION PROTOCOL

In this section we describe EM-SEP protocol, which improves the stable period of clustering hierarchy in the sensor networks. The modification is based on the SEP algorithm as described in the previous section. This modification aims to prolong the stable period of sensor network by maintaining well balanced energy consumption. This means that we choose the advanced nodes to become cluster heads more often than the normal nodes as the SEP protocol does. Furthermore, we take into account the number of nodes that are associated with each cluster head. Our modification tries to evenly distribute the nodes between the selected cluster heads. Another important enhancement in our work that if there are more than one sensor available to be a cluster head at certain round, our
technique chooses the sensor with the highest energy. Those two factors prolong the stable period of the sensor network as shown in the result section.

VIII. ANALYSIS OF EM-SEP PROTOCOL

EM-SEP have been simulated over 100 * 100 m² field with 100 nodes sensor network randomly located and for other experiments we increases the number of nodes to 150. For the sink location, we simulated the environment where the sink located in center of the area and for other scenarios it was placed in the corner of the area. Hence, the location of each sensor is randomly selected and uniformly distributed over the field dimensions. All sensors sense and transmit 4000 bits packet every round. The nodes aggregate the data they receive from other nodes with their own data, and produce only one packet regardless of the number of received packets. The base station (sink) is considered as a cluster head. Therefore each sensor sends the sensed data to the closest cluster head. All sensors start with 0.5 J. The results are compared with SEP algorithm of clustering heterogeneous sensor networks using the following factors:

- Stability period (the time where the first sensor dies).
- Number of nodes alive per round.
- Number of dead nodes per round.
- Number of packets sent to base station.
- Number of packets sent to cluster head.

A. Stable Period

From the analysis of our work results, it is found that EM-SEP prolongs the time interval before the death of the first node (stability period) more than the SEP algorithm. EM-SEP extended the stable period compared to SEP by 5%. Fig. 2 shows results for EM-SEP and SEP where m = 0.1 and a = 1, and the base station is located in the center of 100 m² area.

Fig. 2 demonstrates that EM-SEP prolongs the life time of the network by 5%. To verify the results, we repeated the previous experiment by changing the location of the base station to be located in the corner of the area (90x90). Fig. 3 shows that EM-SEP algorithm result is better than the SEP algorithm. Fig. 4 shows the comparison between EM-SEP and SEP while increasing the number of nodes.

A. Throughput

Fig. 5 and Fig. 6 show that the throughput of EM-SEP is close enough to the throughput of SEP in most of the rounds. We have measured the total number of packets sent to both the base station and the cluster heads. It is found that EM-SEP throughput is little bit less than SEP throughput, but EM-SEP is much better to lengthen the sensors time. The number of alive nodes per round and the number of dead nodes per round are shown in Fig. 7 and Fig. 8 respectively. EM-SEP has better number of alive nodes per round, and the number of dead nodes is less than the SEP algorithm.

It is important to note that for more than 50 tests of EM-SEP, the results showed that no sensor is dead before round 950. Meanwhile, SEP results showed that every 10 tests there are roughly 2 tests where the first dead sensor takes place in 800s.
IX. CONCLUSIONS

In this paper, we propose a modification to the SEP protocol to reduce the communications overhead by saving power and extending the network life time. EM-SEP protocol deals with the network as a number of clusters while introducing an efficient mechanism for communications among nodes. EM-SEP protocol increases the stable period of the sensor network by evenly distributing the power consumption among the nodes in the cluster. In addition, we propose a solution in case there are a number of candidate nodes to be the cluster head at any given point of time. EM-SEP protocol suggests that the node with the highest energy should be the cluster head.

REFERENCES

Arafat Abu Mallouh

Arafat Abu Mallouh is originally from Jordan. He is pursuing his Doctorate in Computer Science and Engineering at the University of Bridgeport in Bridgeport, Connecticut, USA. He received his Bachelor’s degree in Computer Science from The Hashemite University, Zarqa, Jordan. Mr. Abu Mallouh received his Master’s degree in Computer Science from Amman Arab University for Graduate Studies, Amman, Jordan. His research interests include artificial intelligence, image processing, Machine Learning, and Data Mining. Currently Mr Abu Mallouh works on new techniques for voice processing.

Khaled M. Elleithy

Dr. Elleithy is the Associate Dean for Graduate Studies in the School of Engineering at the University of Bridgeport. He has research interests are in the areas of network security, mobile communications, and formal approaches for design and verification. He has published more than two hundred and fifty research papers in international journals and conferences in his areas of expertise.

Dr. Elleithy is the co-chair of the International Joint Conferences on Computer, Information, and Systems Sciences, and Engineering (CISSE). CISSE is the first Engineering/Computing and Systems Research E-Conference in the world to be completely conducted online in real-time via the internet and was successfully running for six years. Dr. Elleithy is the editor or co-editor of 12 books published by Springer for advances on Innovations and Advanced Techniques in Systems, Computing Sciences and Software.

Zakariya Qawaqneh

Zakariya Qawaqneh, originally from Ajloun, Jordan, is pursuing his Doctorate in Computer Science and Engineering at the University of Bridgeport in Bridgeport, Connecticut, USA. He received his Bachelor’s degree in Computer Science and Application from The Hashemite University, Zarqa, Jordan. Mr. Qawaqneh received his Master’s degree in Computer Science from Jordan University of Science and Technology, Irbid, Jordan. His research interests include artificial intelligence, image processing, computer languages, and network security.

Ramadhan J. Mstafa

Ramadhan Mstafa is originally from Dohuk, Kurdistan Region, Iraq. He is pursuing his Doctorate in Computer Science and Engineering at University of Bridgeport, Bridgeport, Connecticut, USA. He received his Bachelor’s degree in Computer Science from University of Salahaddin, Erbil, Iraq. Mr. Mstafa received his Master’s degree in Computer Science from University of Duhok, Duhok, Iraq. His research interests include image processing, mobile communication, security and steganography.

Adwan Alanazi

Adwan Alanazi is originally from Saudi Arabia He is pursuing his Doctorate in Computer Science and Engineering at the University of Bridgeport in Bridgeport, Connecticut, USA. He received his Bachelor’s degree in Computer Science from University of Hail, Hail, Saudi Arabia. Mr. Alanazi received his Master’s degree in Computer Science from University of Missouri Kansas City. His research interests include Wireless Sensor Networks and Network Security.