Data Communication Protocols in Wireless Sensor Networks

1. Alara Hergün
Computer Engineering
Yildiz Technical University
Istanbul, Turkey
alarackck@gmail.com

2. Aslı Şeyda Özdemir
Computer Engineering
Yildiz Technical University
Istanbul, Turkey
aseydaozdemir@gmail.com

3. Selahattin Ulaşcan Kılıç
Computer Engineering
Yildiz Technical University
Istanbul, Turkey
ulascank1@gmail.com

4. Selcan Yağmur Atak
Computer Engineering
Yildiz Technical University
Istanbul, Turkey
selcanyagmur.atak@gmail.com

***Abstract*—Wireless Sensor Networks (WSNs) consist of sensor nodes that measure the physical conditions like humidity and temperature. Individual sensor nodes are resource-constrained in terms of processing speed, storage capacity, and communication bandwidth.Therefore, there are numerous data communication protocols based on different approaches for WSNs. These protocols differ in costs, availability, physical and environmental conditions, energy consumption and effectiveness in data communication. In this paper, we first explain the present problems and design challenges for WSNs and then explore the main protocols that have been found until today. We study the advantages/disadvantages and performance issues of each routing protocol and give a brief comparison. The paper concludes with possible research areas for the future.[2]**

***Index Terms*—Wireless Sensor Networks (WSN), protocols, routing, energy-efficient.**

I. Introduction

A wireless sensor network (WSN) is a computer network made up of a large number of small independent sensor nodes or particles. These small independent sensor nodes and particles combine a vast network of computer technology; hardware, software, networking and scheduling methodologies. A wireless sensor network consists of spatially distributed autonomous devices that are used as sensors to cooperatively monitor physical or environmental conditions such as temperature, sound, vibration, pressure, motion or environmental pollution at different locations. The development of the wireless sensor network was initially mobilized for military applications such as battlefield surveillance. But today, wireless sensor networks are used in many civil applications, including healthcare applications, home automation, traffic control, environmental and natural habitat monitoring.

WSNs have different approaches on how to import data to a target because of several design challenges and routing issues such as limited hardware resources, scalability and massive and/or random node deployment; therefore there are some different general protocol types. In this paper we have divided these types into three main categories : flat, hierarchical and location based protocols. Flat-based routing protocol is a network communication protocol based on a routing system in which all routers are each other's peers and which have no structure in between. Hierarchical-based protocols are the ones that choose to cluster some nodes and prioritize some nodes in terms of several factors. Location-based protocols are based on minimizing the distance between the package and the target area. In this paper, we have researched on the most known protocols: SPIN, LEACH, PEGASIS, TEEN, APTEEN and GEAR.[2]

II. Research on Protocols

A. *SPIN (Sensor Protocols for Information via Negotiation)*

SPIN is a data centric protocol which overcomes the problems like implosion, redundant information passing, overlap in flooding protocols, thus SPIN is achieving a lot of energy efficiency. In this protocol, each node is considered as a potential base-station. Each node has the capability of tracking the changes of its' own situation. The sensors running the SPIN protocols are able to compute the energy consumption required to compute, send, and receive data over the network[2]. This will extend the protocol's operating life time. When SPIN operates nodes will name the collected data "meta data", then after meta data negotiation will occur. This will block the transmit the redundant data.

SPIN uses 3 different message types for communication by nodes : ADV(advertise new data.), REQ(receive the actual data), DATA(actual message itself).

When a node has data, instead of sending directly the data, it will send an advertise package (with broadcast) to its neighbour nodes (Using the negotiation will block the overlap and useless communications). This is the reason why the SPIN is an energy efficient protocol. if neighbour nodes are interested with data, they send a request message to main node for receiving the actual data. When main node receive the request it will send the actual data.

Besides there are some disadvantages of using SPIN, it is almost impossible to be sure of the data will delivery.

B. *LEACH (Low-energy adaptive clustering hierarchy)*

LEACH is a hierarchical protocol in which most nodes transmit to cluster heads, and the cluster heads aggregate and compress the data and forward it to the base station, also can be defined as cluster-based protocol used for random rotation of local cluster base stations. The goal of this protocol is to increase the life of the network. To achieve this, efforts are made to reduce the energy consumption required to create and maintain clusters.

In LEACH, nodes are divided into clusters and each cluster consists of Cluster members and a coordinator cluster called CH, called the Head of Cluster. Cluster heads are not selected from statically shaped nodes that lead to rapid termination of the network, but its randomized protocol is used to balance energy consumption between nodes by distributing the CH role to other nodes in the network.

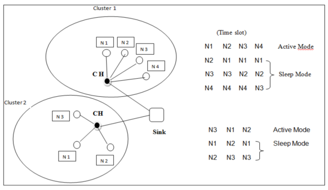

Figure 1. Leach Protocol

LEACH is designed to overcome the disadvantages of flat-architecture protocols that consume more energy. It collects and combines data from CH nodes with small size and meaningful data, and then collects the data with less energy consumption and sends it to the base station. Energy consumption depends on distance and data size. Therefore, it divides the cluster formed in each slot into two parts: the setup phase and the steady state phase.

Setup phase: During this phase the clusters will be formatted and each node decides to be CH or not. The nodes are divided into several clusters dynamically. If the probability of random numbers that a node (N) (P) chooses between 0 and 1 in the current round is less than the threshold T (n), that node is chosen as CH.

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

Once CH is selected, the new CHs introduce themselves to other sensor nodes as they are, using the CSMA protocol. Each Node joins the appropriate CHs by receiving strong signals from the cluster heads. At this stage, they must keep their receivers turned on to be able to detect the signals emitted by each non-CH CH. Heads of clusters of similar type should keep their recipients open to hear the engagement messages.

Steady state phase: At this stage, data will be detected and spread by CH's own members. CH will collect this data and send it to the base station. The network will go to the setup phase after a certain time. As a result, the steady-state stage time is longer than the setup time to reduce overhead. A node that fails to head the cluster after a few number rounds will have low energy. Thus, some CH will not be able to transmit their data to base stations. Therefore, to duplicate a number of non-CH nodes in different order, the threshold equation can be adjusted by adding a factor to multiply the node threshold. Phase continues until the end of the round.

After forming the cluster, the node within the cluster will be able to send data over shorter distances, distributing less energy. CHs consume a large amount of energy as they transmit directly to the base station. It can be said that the biggest problem of LEACH is that the CHs send data directly to the base station (especially if these CHs are far from BS). Fortunately, CHs solve this problem by using multi-hop transmission to BS via other CHs. CHs only send data to each other until they reach BS and they do not have to collect data from other CHs again.

- E-LEACH: Energy-LEACH Protocol improves the CH selection procedure.It makes residual energy of nodes as the main metric which decides whether the nodes turn into CH or not after the first round [5]. It is divided into rounds like LEACH, with the possibility that all nodes are the same to be CH in the first round.
- TL-LEACH : In LEACH, CH sends data to the BS inside a tab. However, in two-level LEACH, CH1 collects data from cluster members and transfers data through another CH2 between CH1 and BS [6].
- M-LEACH: As mentioned above, in LEACH, CH sends data to the base station in a single-hop. In the multi-tabbed LEACH protocol, data from the CH transfers to the BS using other CHs through the relay station [6]. In this protocol, the problem of large energy consumption during data transmissions in CH, which is far from the base station, has been solved.
- LEACH-C: LEACH has no information about the locations of CHs. But, the central LEACH protocol can produce better performance by distributing cluster heads across the network [9]. During the set-up phase of LEACH-C, each node sends information about its current location (possibly determined using GPS) and residual energy level to the sink. In addition to determining good clusters, the sink needs to ensure that the energy load is evenly distributed among all the nodes. To do this, sink computes the average node energy, and determines which nodes have energy below this average.
- V-LEACH: In the original LEACH, the CH is always on receiving data from cluster members, aggregating these data and then sending it to the BS that might be located far away from it. The CH will die earlier than the other nodes in the cluster because of its operation of receiving, sending and overhearing. When the CH dies, the cluster will become useless because the data gathered by cluster nodes will never reach the base station. In our V-LEACH protocol, besides having a CH in the cluster, there is a vice-CH that takes the role of the CH when the CH dies because of the reasons we mentioned above [7].

C. PEGASIS (Power-efficient gathering in sensor information systems)

PEGASIS is a protocol which is based on LEACH. It is designed to optimize LEACH in terms of power usage. In this protocol, greedy algorithm is used so there is only communication between neighbor nodes. After each iteration, the closest node among the neighbor nodes is chosen to create a pathway but as the greedy algorithm is used, the distances can be very high at some times and some loops can be formed as well. This problem causes a larger energy consumption and an increase in the dead nodes' percentage. [1, 2]

It is considered to be a hierarchical protocol since after every iteration the leader node changes with respect to other nodes' distances in the neighborhood. Different from LEACH, PEGASIS doesn't construct a clustering algorithm but tries to form a chain between the leader node and the base station. If any node dies in between the chain is reformulated to 'bypass' the node. [9]

By forming a chain, the power needed is reduced; in each iteration the distance is chosen to be a minimum; but with the greedy algorithm there may sometimes be some delay and redundancies in the chain. Nodes take turns to be a cluster head so no external energy dissipation is needed. Also, "the chain leader in each iteration of data communication is at a random position on the chain, which is important for nodes to die at random locations. The idea of nodes dying at random places is to make the sensor network robust to failures." [10, 11]

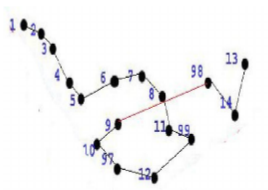

Figure 2. The chain formed in PEGASIS with Greedy Algorithm [1]

D. TEEN APTEEN (Threshold sensitive Energy Efficient sensor Network and Adaptive periodic threshold sensitive energy efficient sensor network)

We can examine network protocols in three classes in terms of working principle: Proactive Networks, Reactive Networks, Hybrid Networks

The working principle of Proactive Networks is based on sending data in time periods. When there are instant changes, data is not transmitted. Only when the specified time period is reached, the Cluster Head (CH) sends a TDMA scheme to the Nodes connected to it and the Nodes send the data they hold to CH. Such network protocols are not suitable for applications that require instant data. The period time cannot be changed after clusters are determined and the network is established. The working principle of Reactive Networks is that when a certain threshold value is reached, the relevant Node sends the data it has to CH in its cluster. In these networks, sensor nodes continuously scan the environment and transfer the data when the relevant threshold value is exceeded. This network is very suitable for time critical applications. Hybrid Networks, on the other hand, have the ability to act as both proactive and active networks. Hybrid Networks can act as proactive networks and reactive networks. It can send data based on a certain periodic time like proactive networks. Like reactive networks, by setting a threshold value, they can instantly send data when that threshold is exceeded.

1) TEEN (Threshold sensitive Energy Efficient sensor Network): TEEN protocol has a reactive network structure. In Reactive Networks, sensor nodes continuously sense the environment and transmit the value as soon as the sensed parameter exceeds a user specified threshold value [13]

In this protocol, CH sends Hard Threshold (HT) and Soft Threshold (ST) messages to the nodes connected to it in addition to attributes at each cluster change time.

Hard threshold (HT): This is a threshold value for the sensed attributes developed for reactive networks. It is the absolute value of the attributes beyond which the node sensing this value must switch on its transmitter and report to its CH. Soft threshold (ST): This is small change in the value of the sensed attributes that triggers the node to switch on its transmitter and transmit. [13]

The advantages of this protocol are as follows: If the attribute value does not reach the Threshold value, no data is sent. With Soft Threshold (ST), if the current data is the same or very close to the previous value sent, the data will not send and the sensor life is extended. The disadvantages of this protocol are as follows: The sensors do not send any data until the threshold is reached, so the user cannot receive data from the network and if the sensors expire, the user will not be able to understand this situation.

2) APTEEN (Adaptive periodic threshold sensitive energy efficient sensor network): Hybrid networks combine the best features of both proactive and active networks and reduce their limits, and APTEEN is a good example of these networks. We propose to combine the best features of proactive and reactive networks by creating a Hybrid network with that sends data periodically, as well as responds to sudden changes in attribute values[16] One of the most important differences compared to Reactive and Proactive networks is that the user can change the period duration and Threshold value himself. It uses a similar structure to the TEEN protocol in data transfer. CH broadcasts the following message: Attributes, Thresholds, Schedule and Count time.

The advantages of this protocol are as follows:

1) By sending periodic data, it gives the user a complete picture of the network. It also responds immediately to drastic changes, thus making it responsive to time critical situations. Thus, It combines both proactive and reactive policies.
2) It offers a flexibility of allowing the user to set the time interval and the threshold values for the attributes.
3) Energy consumption can be controlled by the count time and the threshold values.
4) The hybrid network can emulate a proactive network or a reactive network, by suitably setting the count time and the threshold values.[14]

E. GEAR (Geographical and energy aware routing)

GEAR is a location-based routing protocol algorithm which focuses mainly on choosing neighbors with respect to energy awareness and geographical location. While choosing a neighbor there are two phases : forwarding the packets to the target region and disseminating the data within this region using a Recursive Geographic Forwarding algorithm. While forwarding the data, two different cases are available:

1) If there is a close node available, the algorithm chooses a next-hop and proceeds to that node.
2) If all nodes are away, there is considered to be a hole in the area. GEAR picks a next-hop node considering a cost rate (which will be explained in detail in the next paragraphs). [12]

The neighbor computation is made according to a cost function $h(N, R)$. We consider that the node N is transporting the package P to the target region R. Each node is assigned to a cost and within each of the algorithm, these cost are tried to be minimized within the neighborhood of the node and re-updated as a learned cost $h(Ni, R)$. If a node hasn't got a learned cost value yet, it is assigned to a estimated cost :

$$c(Ni, R) = \alpha d(Ni, R) + (1 - \alpha)e(Ni) \qquad (2)$$

In the first case, greedy algorithm is used to decide the closest neighbor. If there are no holes in the area, the next-hop node is always a neighbor node and as the distance to the destination node minimizes, the learned cost path value will be decreased as well. In this case if all nodes have equal energy, classic greedy algorithm is adopted. If all nodes are equidistant but differing in energy GEAR tries to minimize the cost function and therefore chooses the node with the minimum energy cost. [12]

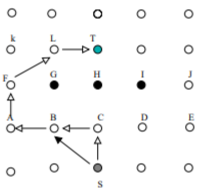

Figure 3. Learning routes for a case with holes [12]

In the second case; if there is a hole and the algorithm cannot find a closer node which is closer to the target node T, there is a recursion to the node before and the second closest node is chosen to proceed. In this case, as there is a recursion the learned cost for the node which needed a recursion is updated. [12]

Once the package is arrived at the target region R, Recursive Geographic Forwarding is used. The target region is divided into several sub-regions and the package is multiplicated for every sub-region until one reaches the target node T. When the current node is the only one in the sub-region, recursive splitting comes to an end. [12]

III. Conclusion

In this paper, we tried to summarize most popular routing protocols in Wireless Sensor Networks and have classified this techniques by three common categories: Flat, Hierarchical and Location Based Routing Protocols. We also mentioned the advantages and disadvantages of routing techniques. Even though this protocols are very famous, they still have some problems that need to be solved. We tried to cover some of these problems. Data communication in wireless sensor networks has limited solutions for now, but we believe that these protocols will improve in the near future. [2]

In the table below, the differences and outcomes of the protocols that we have covered in this paper can be observed:

Protocols	Classification	Mobility	Position Awareness	Power Usage
SPIN	Flat	Possible	No	Limited
LEACH	Hierarchical	Fixed BS	No	Maximum
PEGASIS	Hierarchical	Fixed BS	No	Maximum
TEEN and APTEEN	Hierarchical	Fixed BS	No	Maximum
GEAR	Location	Limited	No	Limited

References

[1] Hasan Al-Hasan, Mohammad Qatawneh, Azzam Sleit, Wesam Al-mobaideen. "EAPHRN: Energy-Aware PEGASIS-Based Hierarchcal Routing Protocol for Wireless Sensor Networks." Journal of American Science 2011;7(8):753-758]. (ISSN: 1545-1003)

[2] Shamsad Parvin, Muhammad Sajjadur Rahim. "Routing Protocols for Wireless Sensor Networks: A Comparative Study." International Conference on Electronics, Computer and Communication (ICECC 2008)

[3] Ibrihich OUAFAA, Krit SALAH-DDİNE, Laassiri JALAL, El Hajji SAİD. "Recent Advances of Hierarchical Routing Protocols for AdHoc and Wireless Sensor Networks: A Literature Survey.", Bilişim Teknolojileri Dergisi, Cilt:9 ,Sayı:2, 2016.

[4] Kulik, J., Heinzelman, W. Balakrishnan, H. "Negotiation-Based Protocols for Disseminating Information in Wireless Sensor Networks," in IEEE Communication Magazine, pp. 169–185, 2002.

[5] Improvement on LEACH Protocol of Wireless Sensor Network (VLEACH) M. Bani Yassein, A. Al-zou'bi, Y. Khamayseh, W. Mardini, January, 2009

[6] V. LOSCRÌ, G. MORABITO, S. MARANO, 2005, A Two-Levels Hierarchy for LowEnergy Adaptive Clustering Hierarchy (TL-LEACH), IEEE, 0-7803-9152-7.

[7] DISSERTATION, HANGZHOU, ZHE JIANG AND MO XIAOYAN, 2006, Study and Design on Cluster Routing Protocols of Wireless Sensor Networks.

[8] W. B. Heinzelman et al., "An Application-Specific Protocol Architecture for Wireless Microsensor Networks,", October, 2002

[9] Payal Jain, Dr. Anu Chaudhary. "The Comparison between Leach Protocol and Pegasis Protocol based on Lifetime of Wireless Sensor Networks.", IJCSMC, Vol. 6, Issue. 12, December 2017, pg.15 – 19

[10] Shamsad Parvin, Muhammad Sajjadur Rahim "Performance Evaluation of LEACH and PEGASIS: Two Prominent Routing Protocols for Wireless Sensor Network"

[11] Sujata, Brijbhushan. "Energy Efficient PEGASIS Routing Protocol in Wireless Sensor Network", International Research Journal of Engineering and Technology (IRJET), Volume: 04, Issue: 07, July-2017

[12] Yan Yu, Ramesh Govindan, Deborah Estrin. "Geographical and Energy Aware Routing: a recursive data dissemination protocol for wireless sensor networks", 08/14/2001.

[13] Jigisha Parmar,Ashishgoud Pirishothm. "Study of Wireless Sensor Networks Using Leach,Teen and Apteen Routing Protocols", 2015

[14] Arati Manjeshwar, Dharma Agrawal. "APTEEN: A Hybrid Protocol for Efficient Routing and Comprehensive Information Retrieval in Wireless Sensor Networks", January 2002