

A Survey of Data-Centric Protocols for Wireless Sensor Networks

Qasim Zia*

Department of Computer Science and Information Technology, FAST-NU University, Pakistan

Abstract

Many changes have been made in sensor fields which are different for different applications and there are many more which are under development. It is under research to develop sensor nodes which utilize low power and are of low cost. In this paper we have overviewed different data centric protocols and their up gradation. After this there is comparison in some of the latest data centric protocols on different performance metrics that affect the application or wireless sensor network.

Keywords: Wireless sensor networks; Data centric protocols; Performance metrics; Applications

Introduction

Due to large number of nodes exist in the network and their random position they are lack of global identification. Due to many such wireless sensor network applications it faces difficulty in querying a particular set of sensors [1]. Mostly it would lead to repeated transmission of data from all sensor nodes with inefficient energy acquisition. A valuable solution is the definition of routing protocols which are capable of selecting particular sets of forwarding sensor nodes and to make use of data aggregation in the transmission of data [2,3].

This routing technique is known as data centric routing. It varies from traditional address based routing in which routes are based on addressable nodes. In the data centric technique, the sink sends queries to particular regions and then looks for data from the sensors situated in the particular regions [4,5]. Attribute based naming is mandatory to describe the characteristics of data requested in the queries. SPIN and Directed Diffusion are the initial two data centric protocols proposed and they motivate other data centric techniques as well.

SPIN (Sensor protocol for information via negotiation)

It fulfills the problems of classic flooding approach by implementing two innovative aspects negotiation and resource adaptation.

To fulfill the deficiencies of implosion and overlapping, SPIN states that nodes have to negotiate each other before exchanging their data because of which only desired data will be transmitted in network. It saves our energy. To negotiate properly, nodes must be able to specify the information that they have gather. The high level description of data is called Meta data according to their idea [5,6].

Meta data actually describes and resume the data that is collected by a sensor. However, till now there is no standard format for Meta data that is why data description way is application driven. SPIN can only benefit us if the capacity of Meta data is shorter than the size of original data [7].

The family of protocols that belongs to SPIN consists of two discrete protocols which are known as SPIN-1 and SPIN-2 that uses negotiation before sending data which makes it possible that only important information will be transferred to the other node. In the network nodes do a Meta data negotiation before any data is sent to other nodes. Before exchanging the actual data, data advertisement is used to exchange the Meta data. There are three types of messages in SPIN-1 ADV, REQ and DATA. ADV is used to advertise the new data; REQ is used in order to request the desired data; DATA is the actual data itself. It is the on

demand protocol in which every time when the node gets a new data it advertises an ADV message having metadata to its neighbors by broadcasting. That is how the neighbor that does not have the data that is advertised can have it by sending a request message REQ to the node. After that, DATA is transmitted to that neighbor node. The neighbor sensor node then repeats these steps with their sensors. In the end all sensor area will receive a copy of data [8-10].

This protocol works in time-driven manner and collects the information from all nodes in the network assuming as a base station. It provides the facility that a person can query any node and get the required data immediately.

SPIN-1 can't describe any energy policy but it defines an interface for application to query the available resources. Before data is exchange nodes query resources in order to become familiar with the fact that how much energy is available. Each node has a resource manager in order to know the resource consumption and calculate the cost of computations, sending and receiving data [11,12].

Its advantage is that each node only needs to concern with its neighbors which are one hop away. But its disadvantage is that we are not sure that our data is transmitted due to advertisement of Meta data because if the data required by the node is far from it then the data cannot be transmitted. It is not feasible for applications that require reliable delivery of data e.g. an application that is used to detect intrusion [13].

A new version of SPIN-1 is SPIN-2 which uses threshold based resource awareness system in integration to negotiation. SPIN-2 works similar like SPIN-1 three stage protocol, when energy level is sufficient. Meanwhile, if the energy level is down node limits its indulgence in the protocol. Which means it takes part only if it has energy enough to complete all the three stages without going into low level threshold.

SPIN protocol (Figure 1) efficiently transmits data, while keeping no per-neighbor state. These protocol works well in the environment in

***Corresponding author:** Qasim Zia, Department of Computer Science and Information Technology, FAST-NU University, Pakistan, Tel: 923234400172; E-mail: qasim_233@yahoo.com

Received February 07, 2015; **Accepted** March 21, 2015; **Published** March 23, 2015

Citation: Zia Q (2015) A Survey of Data-Centric Protocols for Wireless Sensor Networks. J Comput Sci Syst Biol 8: 127-131. doi:10.4172/jcsb.1000180

Copyright: © 2015 Zia Q. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

which sensors are mobile, because they make forwarding decisions on the basis of local neighborhood information. Other protocols of SPIN family are:

SPIN-PP: This protocol works well in point to point communication.

SPIN-EC: This protocol is same as SPIN-PP, but with limited energy.

SPIN-BC: This protocol works well for broadcast channels.

SPIN-RL: This protocol is used when the channel is loss and it is used to make adjustment in the SPIN-PP protocol to make up for loss channel.

M-SPIN (Modified SPIN): This protocol is used to transmit information only to sink node instead of whole network. In this protocol total number of packet transmission is less, due to which, large amount of energy is saved.

Directed Diffusion

Directed Diffusion is described as follows a human operator uses sink node to query some specific aspects of a target region, in replay to that query the particular region collect the data that is required or specified in query, once the required data is collected result is send back to sink. There are four elements of Directed Diffusion naming, interests, gradients and reinforcement (Figure 2).

Attribute value pairs are used for data naming. A sensing task is spread in the network by the sink in the form of interests by the named data. The sink constantly broadcasts each of its neighbors with interest messages. Each node contains an interest cache with each item have different interests. Each item in the cache has different fields like timestamp, data rate and duration. Cache also keeps record of recent data items in order to prevent loops. These events set up gradient in the network that are used to draw the actions for data matching the interests and replay it back to sink. This replay uses multiple paths to access the sink. The network reinforces only one or a few set of these paths.

The nodes that are intermediate can cache or aggregate data. It uses on demand data querying system. Its advantage is that all data transfers are between neighbors with no need for addressing mechanism. Each node is capable of aggregation and caching data [14,15].

Its disadvantage is that because it is on demand query driven it is not sufficient for applications that require continuous data delivery. For example applications like environmental monitoring. Another disadvantage is that attribute based naming is application dependent because each application has its own priority.

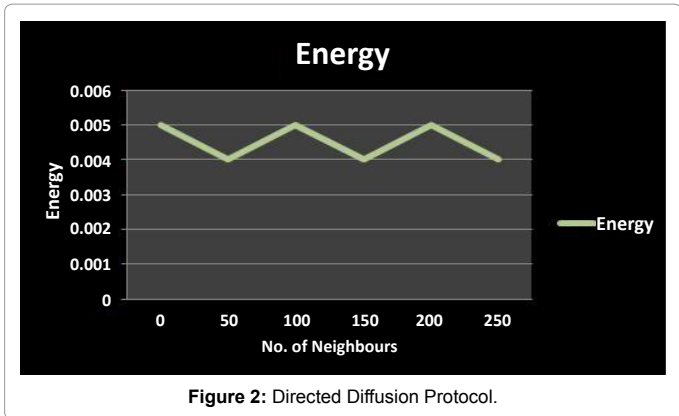


Figure 2: Directed Diffusion Protocol.

Energy Aware Routing

Shah and Rabaey [16] described the use of a group of paths which are sub optimal to increase the lifetime of the particular network.

These paths are choosing because of the probability function, which lay on the energy utilization of each specific path.

Variance paths are used with a specific probability in such a way that the lifetime of the network as a whole will increase and node's energy don't decrease.

Rumor Routing

Rumor Routing basically use agent based path creation algorithm which is basically another variation of directed diffusion. The routing took place between query flooding and event flooding. Some of the advantages of rumor routing are:-

- i. It creates only single path between destination and source.
- ii. As compared to flooding it gives energy saving.
- iii. Node failure situation is easily handled.

One problem with this approach is:-

- i. When numbers of events are less only then this routing performs well.

ACQUIRE (Active QUery forwarding in sensoR nEtworks)

It visualize wireless sensor network as a distributed database.

Its main property is that it divides complex queries into the sub queries. First sink sends a query. By using the cache information every node tries to answer the query and forward it to other node. If the cache info is not latest the nodes get information from its neighbor within specified hops. Once the query is solved by the neighbor it is send back to the sink through the reverse path or it is send back to the neighbor through the shortest path. Acquire allow various queries to send responses because of which it is able to deal with complex queries.

Consider the fact that Directed Diffusion cannot be used for complex queries due to energy constraints; it is because Directed Diffusion also take use flooding based query mechanism for continuous and aggregate queries. While Acquire can offer efficient querying by adjusting the value of parameter d. If d is equal to network diameter, then there is no difference between Acquire and flooding. On the other hand the query has to travel more hops if d value is too small. In order to choose the next node for forwarding the query, AC- QUIRE either choose it randomly or it is based on maximum potential query satisfaction.

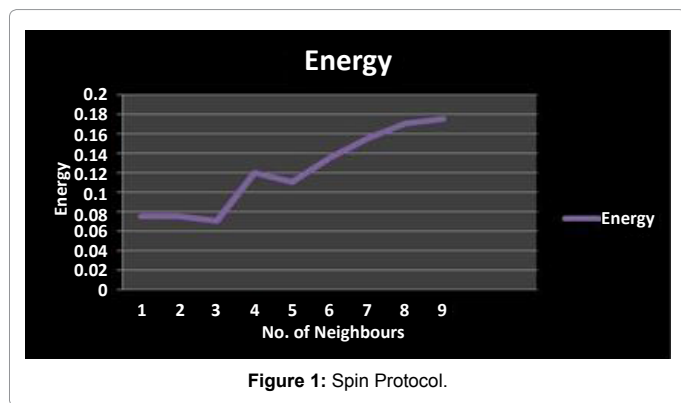


Figure 1: Spin Protocol.

LEACH (Low Energy Adaptive Clustering Hierarchy)

It is basically cluster based protocol. It is based on two phases: a setup phase and a steady phase. A setup phase is responsible for cluster creation in the network and chooses the cluster in the network. Each node decides to become a cluster heads randomly. Cluster head chooses the data to be used in its cluster. In the steady phase the node in the cluster sense and forward data to its cluster head. Cluster head gather all the data send by the node, start compress and aggregate it and send back to sink. LEACH assumes that all cluster head can directly communicate with the sink of the network. Therefore in the network having large regions it is not applicable. Nodes can sleep when there is not their turn to transmit. Cluster heads are rotated randomly. It transmits only new data to the sink.

Its advantages include that it is distributed with no global knowledge is required and we can save energy due to aggregation in the cluster head.

Its disadvantages includes that this protocol assumes that each node have enough power to transmit it to the cluster head and cluster head have enough power to transmit it to sink.

Improvements of LEACH

Because of some deficiency in LEACH. There is some latest research took place in order to improve the performance of this protocol. Some of these researches are:- E-LEACH, TL- LEACH, M-LEACH, LEACH-C, and V-LEACH.

E-LEACH: E-LEACH stands for energy LEACH in which our main concern is to improve the CH selection procedure. Similar to LEACH, it is also divided into rounds, in which the first round all nodes have the same probability and chance to become Cluster Head (CH). After the first round the energy level of every node varies with each other and node with higher energy will be taken as CH in comparison with nodes that have less energy.

TL-LEACH (Two Level): As we know that in LEACH the CH transmits the data in a single hop to the base station. On the other hand, in two levels LEACH, the CH gathers data from its cluster members and transmit it through a cluster head to the base station which is placed between the cluster head and base station.

M-LEACH: As we described above, the CH transmit the data in a single hop to the base station. In Multi hop LEACH protocol, the CH used the other CH's to transmit the data to the sink. The benefit of this protocol is that the problem that we are facing with CHs that are at a distance from the base station and consume large amount of energy during data transmission has been resolved (Figure 3).

V-LEACH (Vice): In this new Version LEACH protocol, we submit a vice-CH in addition to CH in the cluster. Its duty is to take the position of CH when the CH dies. As we know that when a CH dies, there is no benefit of cluster because of the fact that the information gather by the node members will not reach the sink.

LEACH-C: As we see that LEACH has no information about the CH's places. On the other hand, centralized LEACH protocol can give better performance by making distribute the cluster heads in the network. In the set up phase, each node transmits the energy and location which remains to the sink. In return the sink then runs a centralized cluster formation algorithm in order to find out the clusters for that particular round. But as this protocol needs location information for every sensor node which exists in the network and which is usually given by GPS, it

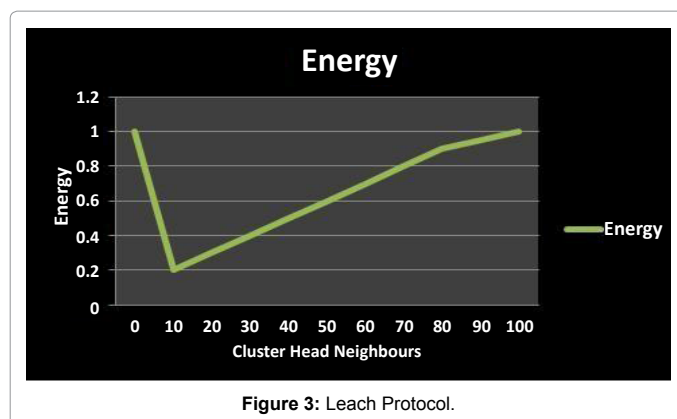


Figure 3: Leach Protocol.

is not popular, powerful and robust.

TEEN (Threshold sensitive Energy Efficient sensor Network)

It is a hierarchical protocol which is beneficial for time critical applications where network work in a reactive manner. The nodes that are closer to each other make clusters and elect a cluster head. The responsibility of each cluster head is to send the data to the sink. After the clusters came into existence the cluster head broadcasts two thresholds to the nodes which are hard and soft thresholds. The minimum the cluster node is called hard threshold.

In hard threshold the node can only transmit when they are in the range of interest. It reduces the number of transmission. If the node senses a value which is at or beyond the hard threshold it can only send when its value changes equal or greater than soft threshold. So, the soft threshold will further refine the number of transmission. One can adjust the values of hard or soft threshold in order to control the transmission. In the applications where periodic reports are needed TEEN is not a very good choice because user cannot be able to acquire any data if the values can't reach the threshold.

APTEEN (Adaptive Threshold sensitive Energy Efficient sensor Network)

There is basically another version and extension of TEEN protocol, which is named as APTEEN. It performs both periodic sensing and it is reactive to time critical events. It is different from TEEN in that it must send a data if it has not send it for a count time equal to cluster head. It consumes less energy as compared to LEACH.

They tried to overcome the TEEN's problem by adding parameters to sensor nodes in every cluster. It eliminates the ambiguity between packet loss and unimportant sensed data which shows no certain change. Through this energy conservation and network lifetime is improved.

Its main disadvantage includes the main overhead and complexity of forming clusters in multiple levels.

GAF (Geographic Adaptive Fidelity)

Geographic Adaptive Fidelity is a location based energy aware routing algorithm used for mobile ad hoc networks but it can be used in sensor networks as well.

The network is divided into fixed pieces to make virtual grid. It makes equal and fixed zones. Its size is dependent into transmitting power which is required. It uses the GPS information to associate itself in the zone.

| Data Centric Protocols | Classification | Power Usage | Data Aggregation | Scalability | Query Based | Overhead | Data Delivery Model | Q o S |
|------------------------|----------------------------|-------------|------------------|-------------|-------------|----------|---------------------|-------|
| SPIN | Data Centric | Ltd | Yes | Ltd | yes | low | Event driven | No |
| DD | Data Centric | Ltd | Yes | Ltd | yes | low | Demand driven | No |
| ACQUIRE | Data Centric | Low | Yes | Ltd | yes | low | Complex query | No |
| LEACH | Hierarchical\ Data Centric | High | Yes | Good | No | High | Cluster-head | No |
| TEEN & APTEEN | Hierarchical\ Data Centric | High | Yes | Good | No | High | Active threshold | No |
| GAF | Hierarchical\ Data Centric | Ltd | No | Good | No | Mod | Virtual grid | No |

Table 1: Classification and Comparison of data centric protocols in WSN [15].

GAF make all nodes equal inside the same zone but at the same time it makes sure that for a specific time period at least one node in the zone is active and keeps the other nodes in the sleeping position. It has no effect on network connectivity and routing fidelity and saves its energy. The active node is responsible for the data collection and forwarding data to other nodes.

GAF can be stated as hierarchical and location based protocol because zones of the grid can be classified as clusters. In order to make balance of the energy limitations in a grid a node can change its position from sleeping to active state.

Simulation Setup

Number of nodes represents the total number of nodes which are present in the network. The number represents and provides the size of the routing table at each node.

Number of sources represents the number of nodes that can send or transmit data.

Dimension of network represents the area in which the node can move.

Application layer is CBR which represents an agent that transmits data at constant bit rate (CBR).

Simulation Time represents the total sum of time that is required for a specific simulation.

Node Speed represents the interval of the speed in which a node can travel.

Pause Time represents the time between two travelling events of the node.

Transport layer Protocol UDP because it do not affect flow of data packets.

We will perform simulations on NS2 and make graph in Microsoft Excel 2010 in order to view energy consumption (Table 1). We have two files tcl and nam file. tcl file is used for coding while nam file is used to view the topology. There are two commands to run tcl and nam file:- ns file name (for tcl file) and nam file name (for nam file). The following Figure 4 shows the screenshot of directed diffusion nam file.

Simulation Results

We took three data centric protocols (Table 2) and compared their energy consumption with each other. In order to compare the performance of SPIN, Directed Diffusion and LEACH in terms of energy consumption we had compared the nodes in topology with the amount of energy consumption. We had performed different experiments. Graphs of which are shown above in Figure 1, Figure 2 and Figure 3. The figures had shown the graphs which were created using MS Excel in this term paper.

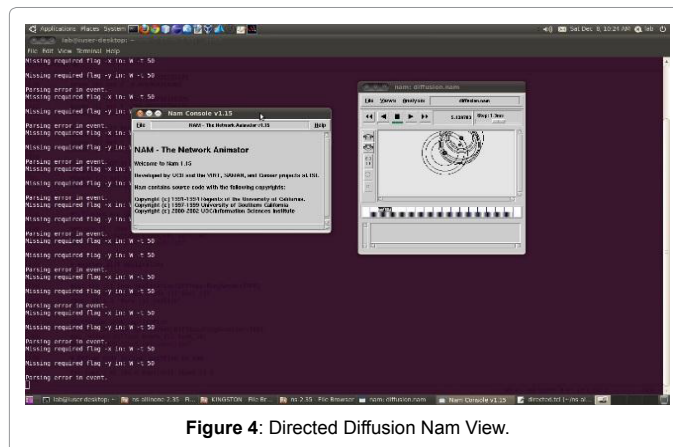


Figure 4: Directed Diffusion Nam View.

| Number of nodes | No. of nodes varies |
|--|---|
| Dimensions of network | 1500×300 m |
| Application Layer | CBR |
| Packet Size | 512 bytes |
| Packet rate | 4 packets/sec |
| Routing Layer | SPIN/DD/ACQUIRE/LEACH/TEEN & APTEEN/GAF |
| MAC | IEEE 802.11(DCF) |
| Simulation time (startup + data transmission + shutdown) | 200 + 900 + 200 sec |
| Node speed | 0-20 m/s |
| pause time | 0-900 sec |
| IFQ length | 50 packets |
| Node transmission range | 250 m |
| Transport layer protocol | UDP |
| Periodic route update interval | 15s |
| Periodic updates missed before link declared broken | 3 |
| Initial triggered update weighted settling time | 6s |
| Weighted settling time weighting factor | 7/8 |
| Route advertisement aggregation time | 1s |
| Maximum packets buffered per node per destination | 5 |

Table 2: Characteristics of Test Network [2].

Conclusion

After surveying the existing data centric protocols, we compare the characteristics of protocols with each other.

After that we analyze that which protocol perform better in which environment, then we make energy consumption comparison of Spin, DD and LEACH in which we come to know although Leach clusters head takes more energy as compared to DD and Spin but the nodes which are under cluster head takes less energy as compared to DD and Spin which saves energy. After leach DD is the better option than Spin

but obviously Spin is better than gossiping and flooding.

Acknowledgment

I would like to thank DR. Saima Zafar for giving her precious time to guide me in this domain.

References

1. Norouzi A, Zaim AZ (2012) An Integrative Comparison of Energy Efficient Routing Protocols in Wireless Sensor Network. *Scientific Research* 4: 65-75.
2. Sahoo BPS, Rath S, Puthal D (2012) Energy Efficient Protocols for Wireless Sensor Networks: A Survey and Approach, *International Journal of Computer Applications* 44: 43-48.
3. Shi L, Zhang B, Mouffah HT, Ma J (2012) DDRP: An Efficient Data-Driven Routing Protocol for Wireless Sensor Networks with Mobile Sinks. *International Journal of Communication Systems* 26: 1341-1355.
4. Prasan UD, Murugappan S (2012) Energy Efficient and QOS aware Ant Colony Optimization (EQ-ACO) Routing Protocol for Wireless Sensor Networks. *International Journal of Distributed and Parallel Systems* 3: 257-268.
5. Eslaminejad MR, Sookhak M, Razak SA, Haghparast M (2011) A Review of Routing Mechanisms in Wireless Sensor Networks. *International Journal of Computer Science and Telecommunications* 2: 1-9.
6. Kannammal KE, Purusothaman T (2011) Comparison of Data Centric Routing Protocols with Random Way Point Mobility Model in Mobile Sensor Networks. *European Journal of Scientific Research* 65: 452-546.
7. Ahmedy I, Ngadi MA, Omar SN, Chaudhry J (2011) A Review on Wireless Sensor Networks Routing Protocol: Challenge in Energy Perspective, *Scientific Research and Essays* 6: 5628-5649.
8. Rehena Z, Roy S, Mukherjee N (2011) A Modified SPIN for Wireless Sensor Networks. *IEEE* 1-4.
9. Joshi A, Priya LM (2011) A Survey of Hierarchical Routing Protocols in Wireless Sensor Network, *MES Journal of Technology and Management* 67-71.
10. Roseline RA, Sumathi P (2011) Energy Efficient Routing Protocols and Algorithms for Wireless Sensor Networks_A Survey. *Global Journal of Computer Science and Technology* 11: 179-183.
11. Jose Cecilio, Joao Costa, Pedro Furtado (2010) Survey on Data Routing in Wireless Sensor Networks. *European Community's Seventh Framework Programme* 1: 3-46.
12. Tian K, Zhang B, Huang K, Ma J (2010) Data Gathering Protocols for Wireless Sensor Networks with Mobile Sinks, *IEEE Communications Society* 1-6.
13. Azni AH, Saudi MM, Azman A, Johari AS (2009) Performance Analysis of Routing Protocol for WSN Using Data Centric Approach, *World Academy of Science, Engineering and Technology* 3: 187-190.
14. Pavai K, Sivagami, Sridharan D (2009) Study of Routing Protocols in Wireless Sensor Networks, *International Conference on Advances in Computing, Control, and Telecommunication Technologies* 522-525.
15. Biradar RV, Patil VC, Sawant SR, Mudholkar RR (2009) Classification and Comparison of Routing Protocols in Wireless Sensor Networks. *Special Issue on Ubiquitous Computing Security Systems* 4: 704-711.
16. Shah R, Rabaey J (2002) Energy aware routing for low energy ad hoc sensor networks in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)* 1: 350-355.