ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 3, Issue 4, April 2014

Comparative Analysis of Different Data Dissemination Strategies in Wireless Sensor Networks

Aayushi Mahajan[1], Suhaib Ahmed[2], Swastik Gupta[3], Ashish Suri[4]

MTech Student, School of Electronics and Communication Engineering, SMVD University, Katra, India[1,2]

Assistant Professor, School of Electronics and Communication Engineering, SMVD University, Katra, India[3,4]

Abstract: Wireless sensor nodes are uniquely suited for monitoring environmental events owing to certain advantages like small size, wireless nature. However, their usage is restricted by only one engineering limitation i.e. short lifetime. There is always a need of an efficient architecture tailored towards enhancing network lifetime which strongly depends on Data Dissemination strategy used. Data Dissemination is the dissemination of the monitored information, the tasks that sensors are supposed to perform and report to the sink. So it is an important issue to design suitable Data Dissemination protocol according to different application scenarios. A variety of different energy efficient Data Dissemination protocols are proposed in recent years. However there is little comparative study of different Data Dissemination strategies available. In this paper, we present the comprehensive analysis of various Data Dissemination strategies. This paper provides useful insights for the network designer such as which Data Dissemination protocols scale well, reduce overall energy consumption or improve task completion.

Keywords: Wireless Sensor Network, Data Dissemination, Pull based strategy, Push based strategy, Push-Pull based strategy.

I. INTRODUCTION

Wireless Sensor Networks (WSNs), as a multi-hop and self-organization network, consists of sink nodes and sensor nodes with wireless communication mode [8]. Sink nodes have the responsibility to link wireless sensor network and external network. Each sensor node is a smart embedded device which includes four units: data collection unit, data process unit, wireless communication unit and battery. Figure 1 shows a schematic diagram of sensor node components. Once deployed across the monitoring region, each of them can complete data sensing and collection independently. Meanwhile, using communication unit, they can coordinate with each other to realize data delivery according to different queries. The main application of wireless sensor networks is to sense the environment and transmits the acquired information to the sink for further processing. Data Dissemination is the process of transferring desired data from active sensor nodes to data collecting nodes (i.e. sink) in the sensor network. Sometimes Routing and Data Dissemination terms are used interchangeably by many researchers. However, there is a notional difference between the two; routing refers to the process of simply transferring raw packets from source to destination without any in-network processing.

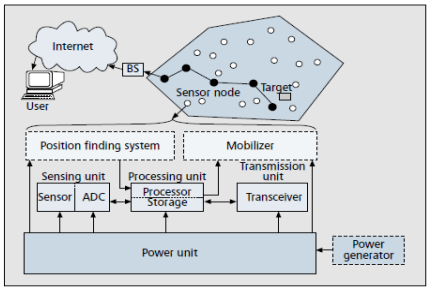

Fig. 1- Components of Sensor Nodes [7]

Data Dissemination may involve some in-network processing of sensor data like data aggregation or data fusion based on correlation among sensor data. In-network processing like data aggregation is essential to eliminate data redundancy, reduce data transmission, and improve data accuracy. Data aggregation results in better bandwidth and energy utilization which enhances the network lifetime because communications constitutes 70% of the total energy consumption of the nodes. Thus, Data Dissemination is a fundamental and very important operation for sensor networks. On the basis of function, all the nodes in WSNs can be divided into three types: source node, rendezvous node, consumer node.

Just as their names imply, Source node performs data acquisition and event detection. Rendezvous node is responsible for sensory data storage. Consumer node is the initiator of data query. Whenever we carry out the task of data storage and retrieval, only the relevant data will be disseminated among these three kinds of nodes. So it is an important issue to design suitable Data Dissemination protocol according to different application scenarios. This paper is organized as follows- In section II, the design and evaluation criterions are proposed according to the constraints of WSNs. In section III, we describe the different Data Dissemination strategies and analyze their merits and demerits. Next, in section IV, we present brief overview of various Data Dissemination protocols. Finally, section V sets out to conclude this work.

II. DESIGN RULES AND EVALUATION CRITERIONS

In order to overcome the constraints of WSNs, some following design rules must be followed:

ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 3, Issue 4, April 2014

- **Energy Efficiency**-Sensor nodes are energy constrained. For prolonging network's lifespan as much as possible, how to reduce the energy consumption is an import issue which needs to be addressed in the WSNs. In general, Energy consumption of sensor nodes comes from several aspects, e.g. communication data sensing, data computing and so on. Yet, the major; art of energy consumption is caused by communication. For this reason, the reduction of communication cost is the most important design rule.
- **Load Balance**- In the storage and query phase, if some sensor nodes take more responsibilities for data receiving and forwarding than others, they will become hotspots in the WSNs. But for enhancing network lifetime, energy consumption of all nodes should be similar. So, load balancing is another important design issue.
- **Robustness**-The topology of sensor network may be changed because of several reasons. For example: some nodes change their location, some nodes lose effectiveness or some new nodes are added in. Furthermore, the change of topology brings some new challenges to data storage and retrieval. So some measures must be considered to address these problems for system robustness.
- **Computation Simplification**- The sensor nodes, as embedded devices, have only limited computing and processing abilities. So, Apart from ensuring the correctness and validity of Data Dissemination algorithms, the algorithmic complexity should be reduced as much as possible.

In fact, these design rules also can be regarded as evaluation criterions and can be used to compare and evaluate different Data Dissemination algorithms.

III. DIFFERENT DATA DISSEMINATION STRATEGIES

Authors in [6] have classified Data Dissemination strategies into four major categories based on basis of operation [6]:

Push-based strategy, On-demand (or pull-based) strategy, hybrid strategy and data allocation over multiple broadcast channels.

- **Push Based Strategy**- All the source nodes push sensory data to sink nodes through multi-hop routing. In this way, query result can be retrieved from sink nodes without communication cost. The push-based Data Dissemination protocol is very efficient when the query rate is relatively higher compared to data acquisition frequency. Because, it can reduce the query routing cost to zero. However, this protocol has some disadvantages: (1) The method trades off communication cost in storage phase to guarantee efficient query in the later steps. (2) All sensory data must be broadcasted to sink node using multi-hop routing. Naturally, the neighbor nodes of sink nodes will undertake more data delivery task than other sensor nodes, and then the hotspots will be formed. These nodes must lose their effectiveness in advance because of energy load unbalance. Therefore, the system robustness and stability will be hard to be ensured.

There are two kinds of push based Data Dissemination strategy namely flat broadcast and broadcast disks [10].

Flat Broadcast- The simplest scheme for data scheduling is flat broadcast. With a flat broadcast program, all data items are broadcast in a round robin manner. The access time for every data item is the same, i.e., half of the broadcast cycle. This scheme is simple, but its performance is poor in terms of average access time when data access probabilities are skewed.

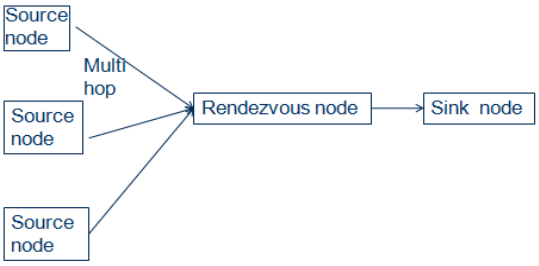

Fig.2- Push Based Strategy

Broadcast Disks- Data items are assigned to different *logical* disks so that data items in the same range of access probabilities are grouped on the same disk. Data items are then selected from the disks for broadcast according to the relative broadcast frequencies assigned to the disks.

- **Pull Based Strategy**- The pull-based Strategy adopts completely opposite idea to push-based Strategy. The source nodes will never deliver the sensor data voluntarily. Instead, they store data at home and wait for query passively. On the contrary, the consumer nodes broadcast query demands to source nodes throughout the network on their own initiatives. Obviously, communication cost takes place only when it is needed, so the method can be applied to the situation where the data production rate is higher than data query rate. The disadvantage is that even though some source nodes have no related data with query, they have to participate into data delivery.

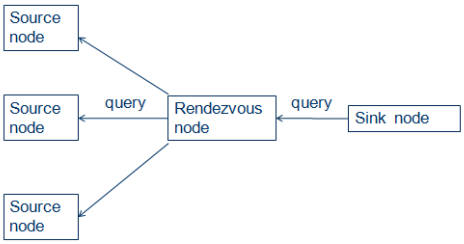

Fig. 3-Pull Based Strategy

- **Push Pull (Hybrid) Strategy**- A promising approach, called hybrid broadcast, is to combine push-based and on-demand techniques so that they can complement each other. It introduces the combination between consumer node and source node. In the first phase, source nodes get the storage location and then transfer sensory data to rendezvous nodes closest to location. Then, consumer nodes can directly transmit query to rendezvous node using same regulations. In this way, queries flooding can be avoided efficiently.

ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 3, Issue 4, April 2014

- **Data Allocation Over Multiple Broadcast Channels-** Multiple physical channels have capabilities and applications that cannot be mapped on to single channels. By having access to multiple physical channels fault tolerance is improved. For example if a server broadcasting on a certain frequency crashes, its work must be migrated to another server. Finally, being able to transmit over multiple channels has scalability benefits.

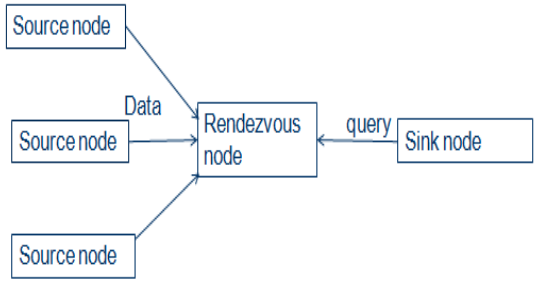

Fig.4-Push Pull (Hybrid) Strategy

IV. OVERVIEW OF DATA DISSEMINATION PROTOCOLS

This section briefly describes some of the important Data Dissemination protocols.

A. Directed Diffusion-

It [11] is a data-centric (DC) and application-aware paradigm in the sense that all data generated by sensor nodes is named by attribute-value pairs. The main idea of the DC paradigm is to combine the data coming from different sources en route (in-network aggregation) by eliminating redundancy, minimizing the number of transmissions, thus saving network energy and prolonging its lifetime.

In directed diffusion, sensors measure events and create gradients of information in their respective neighborhoods. The sink requests data by broadcasting interests. An interest describes a task required to be done by the network.

An interest diffuses through the network hop by hop, and is broadcast by each node to its neighbors. As the interest is propagated throughout the network, gradients are set up to draw data satisfying the query toward the requesting node (i.e., a sink may query for data by disseminating interests and intermediate nodes propagate these interests). Each sensor that receives the interest sets up a gradient toward the sensor nodes from which it receives the interest. This process continues until gradients are set up from the sources back to the sink. More generally, a gradient specifies an attribute value and a direction. The strength of the gradient may be different toward different neighbors, resulting in different amounts of information flow.

The possible advantages are saving of network energy thus prolonging network lifetime. Shortcoming of this protocol is that it can't be used where continuous data delivery is required as it is on demand based data model.

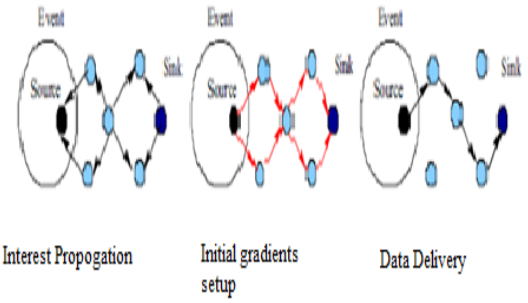

Fig.5- Process of Directed Diffusion

B. Two- Tier Data Dissemination-

TTDD [9] is based on decentralized architecture. It uses a grid structure to divide the topology into cells. Only sensors located at a cell boundary need to forward the data. The consumer actively builds the grid structure through the network and sets up forwarding points in the sensors closest to the grid boundary called dissemination nodes (DN). One tier is the cell at the consumer's current location and the other one is the DN at cells boundaries. The consumer only floods the query within its own cell. When the nearest DN that hears the query, it forwards it to its adjacent DNs. This process continues until the query reaches the producer or one of the DNs that have the corresponding data. During the query propagation period the network establishes the reverse path towards the consumer, so that it can enable the data path to be the same as that of the query propagation.

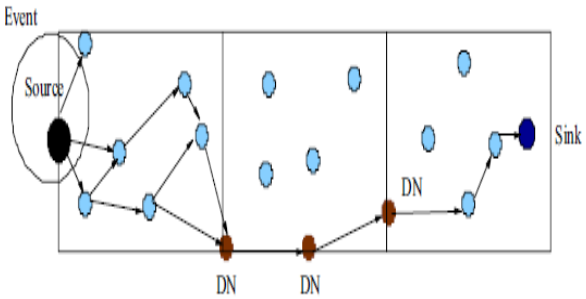

Fig.6- TTDD [9]

C. Location Oblivious Hybrid Data Dissemination-

LOHD [5] does not rely on any location information; it adaptively selects an ultra-node through a well-controlled flooding and the ultra-node maintains the gradients from sources to sinks.

It then incorporates enhanced PUSH and PULL to distribute messages along the gradients instead of flooding**.** LOHD first finds a rendezvous node called *ultra-node*, which is selected by the intersection of local flooding respectively from the sources and sinks.

The ultra-node then builds the gradients between the sources and sinks, so that the following control messages and data can be sent through the gradients instead of flooding. These operations adapt to diverse networks and sink/source distributions, and do not rely on any location information. The disadvantage of this protocol is that routing overhead increases.

ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 3, Issue 4, April 2014

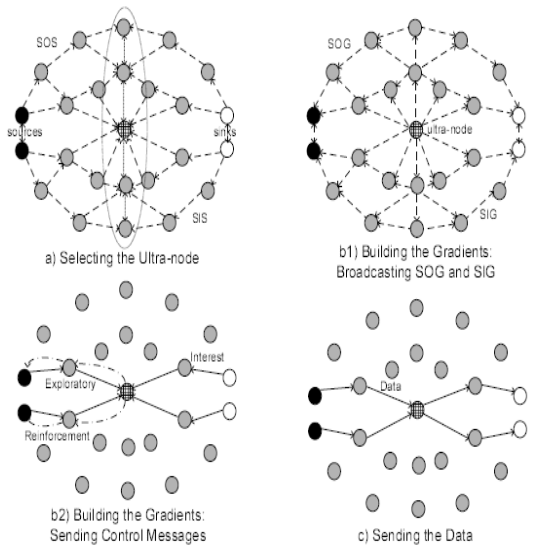

Fig.7- Work Flow of LOHD [5]

D. Balancing Push- Pull-

This model [1] combines push and pull for information dissemination and gathering. The push component features data duplication in a linear neighborhood of each node. The pull component features a dynamic formation of an on-demand routing structure resembling a comb. The aim of this protocol is to support both mobile and stationary query nodes. The scheme applies to applications where a query entry point can be anywhere in the network and occurs at any time. A major application of such a query generation mode is to support mobile information-gathering agents (mobile sinks) or hierarchical networks where higher hierarchies are more intelligent and may demand information. The push-pull query support scheme proposed resembles the action of combing for needles in a haystack or in a pool of sand, and is thus dubbed as "comb-needle." The comb needle query support model combines both push and pull in the following way: A sensor node that detects an event of potential interest pushes its data or data pointer of the event to a certain neighborhood (resembling a needle) and a query node disseminates its request to a subset of the network (resembling a comb).

E. Solution Mapping on Broadcast and On-demand channels-

Algorithm SOM [2] is a composite and generic algorithm which is composed of a search strategy and a broadcast program generation algorithm. Algorithm SOM consists of two major phases: the search space pruning phase and the solution searching phase. Fig. 8 shows the architecture of algorithm SOM. In the search space pruning phase, some infeasible configurations are removed from the search space. Then, in the solution searching phase, a search strategy is used to guide the search of the optimal solutions with the aid of the employed broadcast program generation algorithm and the analytical model of the on-demand channels. . In essence, scheme BIS-Incremental is guided to explore the search space with higher likelihood to be the optimal first, thereby leading to an efficient and effective search.

F. Core Based Reliable Data Dissemination-

In this, the object dissemination is divided into two distinct phases. Before the Data Dissemination commences, a subset of nodes in the network that have reliable links and that form an approximate minimum dominating set are selected as *core nodes*. After this core construction step, in the first phase of the Data Dissemination protocol, the object is reliably propagated from the sink to the core nodes. After the entire object has been propagated to the core nodes, the second phase commences in which the core nodes disseminate the object to their neighboring non-core nodes in parallel. The core-based two-phase approach used by CORD [4] is motivated by the goal of reducing the energy consumption for disseminating the object within the network.

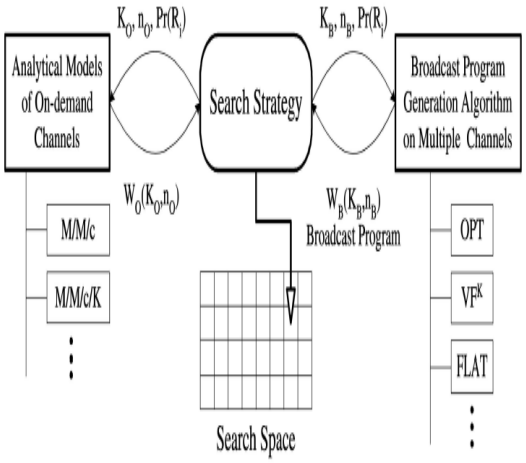

Fig.8- Architecture of SOM [2]

By constructing a core for Data Dissemination, the protocol implicitly selects the set of nodes that are responsible for disseminating the object to their neighbors. This reduces the number of control messages that need to be exchanged between neighboring nodes. Second, the two-phase core-based approach is also suitable for heterogeneous networks, where a subset of nodes in a network is more powerful than the others. A distinctive feature of CORD is that in addition to adopting a two-phase approach, it aggressively uses sleep scheduling in order to further reduce energy consumption for large object dissemination.

G. Real Time Data Dissemination-

In the proposed protocol [3], the Data Dissemination procedure consists of three steps: normal routing, projection routing and overhearing of the mobile sink. All the procedures are based on the spatiotemporal approach. A data packet is forwarded from a source node to the exit point of movable area in normal routing. Some relay nodes of normal routing within a movable area are selected as branch nodes for projection routing. The reason of two routing modes is to define the transmission distance between a source and a mobile sink. The transmission distance is the summation of the distance for the normal routing to one of the branch nodes and the distance for projection routing from the branch node to a mobile sink. This protocol considers the virtual region to be expected for the mobile sink to locate in, and calculates the desired

ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 3, Issue 4, April 2014

delivery speed based on the region. The value of TSDR (time deadline success ratio) which is defined as the ratio of successfully received data packets on the time deadline to all generated data packets from the source node is high in case of RTDD. Figure 9 depicts the Data Dissemination in RTDD.

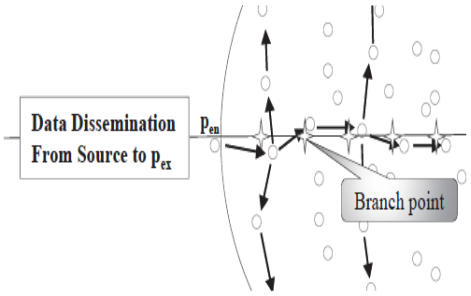

Fig.9- Routing Branching in RTDD [3]

V. CONCLUSION

By analyzing the above Data Dissemination protocols, some conclusions can be drawn. Typically, when these protocols are studied in isolation, the emphasis is on studying only the scaling behavior of the protocol (for example, the impact of network density on scaling behavior). Such an approach can mask the design weaknesses of a particular protocol. Being a relative comparison, this study is to provide useful insights to what kind of design choices are the most desirable in order to improve the performance of proposed protocols. Each of the protocols studied performed well in some cases, but displayed certain drawbacks in others. It is very hard to find a perfect Data Dissemination strategy suitable for all application requirements. Their application effectiveness also needs further evaluation on the hardware platform of sensor nodes.

Table: Comparison of Data Dissemination Strategies

Protocol	Advantage	Disadvantage
Directed Diffusion	Saves network energy and increase in network lifetime	Not suitable for continuous data delivery
TTDD	Suitable for multiple mobile links	Unexpected dissemination node failure
LOHD	Sink node location not required	Increased overhead due to flooding
Balancing Push-Pull	Suitable for continuous and query based data delivery	No search algorithm used
SOM	Bandwidth efficient	End-to-End delay not considered
CORD	Highly energy efficient	No bandwidth saving
RTDD	Effective for Real Time Data Dissemination	Fails for fast varying mobile sinks

REFERENCES

[1] Xin Liu, Qingfeng Huang and Ying Zhang, "Balancing Push and Pull for Efficient Information Discovery in Large-Scale Sensor Networks", IEEE Transactions on Mobile Computing, vol. 6, No. 3, March 2007.

[2] Jiun-Long Huang, Wen-Chih Peng and Ming-Syan Chen, "SOM: Dynamic Push-Pull Channel Allocation Framework for Mobile Data Broadcasting", IEEE Transactions on Mobile Computing, vol. 5, No. 8, August 2006

[3] Seungmin Oh, Yongbin Yim, Jeongcheol Lee, Hosung Park, and Sang-Ha Kim, " Real-Time Data Dissemination for Slowly-varying Mobile Sinks in Wireless Sensor Networks", IEEE 24th International Symposium on personal, indoor and mobile radio communication: Mobile and Wireless networks, 2013.

[4] Leijun Huang and Sanjeev Setia, "CORD: Energy-efficient Reliable Bulk Data Dissemination in Sensor Networks", 26th IEEE international conference on advanced information networking and applications, 2012.

[5] Xu Cheng Feng Wang Jiangchuan Liu, "Hybrid PUSH-PULL for Data Diffusion in Sensor Networks without Location Information", IEEE communication society , ICC proceedings , 2008

[6] .Fang Hongping, Fang Kangling, "overview of Data Dissemination strategies in wireless sensor networks", in International conference on E- Health Networking, Digital Ecosystem & Technologies,2010.

[7] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," IEEE Communications Magazine, vol. 40, pp. 102 - 114, Aug 2002.

[8] Cui-Li,Ju Hailing,Miao yong. Overview of Wireless Networks[J].Journal of computer Research and Development, 2005 ,42(1):163-174

[9] F. Ye et al., "A Two-Tier Data Dissemination Model for Large-Scale Wireless Sensor Networks," Proc. ACM/IEEE MOBICOM, 2002.

[10] Yasir Faheem, Saadi Boudjit, Ken Chen, "Data Dissemination strategies in mobile sink wireless sensor network: A Survey," IEEE communication letters, Vol. 17, No. 1, March, 2009.

[11] Chalermek Intanagonwiwat, Ramesh Govindan, and Deborah Estrin. Directed diffusion: A scalable and robust communication paradigm for sensor networks.In Proceedings of the Sixth Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom 2000), pages 56–67, Boston, MA, USA, August 2000. ACM Press.