

A STUDY ON DIFFERENT DATA GATHERING AND DISSEMINATION MECHANISMS IN SENSOR NETWORKS

Rasidul Karim

Dhruba Chand Halder College, Dakshin Barakat, South 24 Pgs., West Bengal, India

Abstract:

Data collection or gathering is the core tasks of sensor networks. In data gathering methods, sensors collect required information from the environment or the area where the sensor are deployed. As per query given by the desired authority or users, sensed data are aggregated by the selected leader node and disseminated to Sink nodes through base stations (BS). Then the collected information analyzed by the users or dedicated authorities for specific purposes such as environment monitoring, event tracking or making decision. In this study I, briefly overview the different data gathering and dissemination methods of sensor networks which includes how data are gathered, aggregated and forwarded to the sink nodes for analysis. In this research article I, also discuss several data gathering algorithm and their explanation based on maximization data gathering considering with the constraints such as energy, latency and channel bandwidth of the medium.

Keywords: Sensor Networks, data gathering, data dissemination, SEAD, TTDD

1. Introduction

With the advancement of technology, small and low power devices integrate sensors with actuators and transceivers that have limited on board processing capabilities. These devices Known as sensor devices. Here I concentrated with the network of such sensor devices called sensor network. I focus on the data collection mechanism and aggregation by the networks leaders that help to monitoring and control system like habitat monitoring [1], home automation [2], target tracking [3], inventory monitoring [4], environment monitoring, safety monitoring [5] etc. The most important task of sensor network gathered desired sensing information, aggregating the gather data and lastly disseminated to sink nodes for further analyses or helps to make decision instantly as per requirement by the users or authorities. This study focusses on different mechanism of effective data gathering, aggregating and forwarding to the sink for further processing. Fig 1. Shows the relationship between users and sensors.

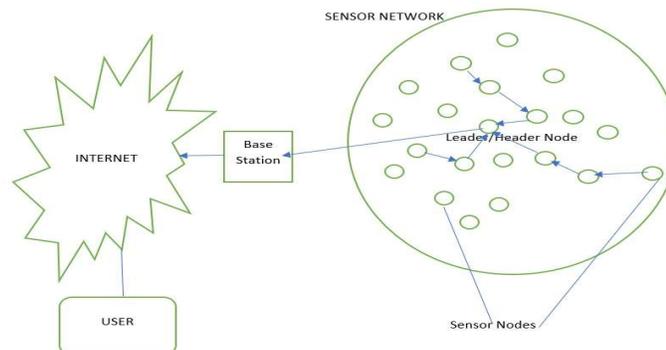


Fig 1. The relationship between users and sensors.

Data gathering methods of sensor networks not only focus on data sensed by the sensors but also send to the centralized data server through base station for further processing and help to make proper decision. Traditionally sensors deployed densely so that the network does not suffer with coverage problem i.e., ensuring that the network covers the whole target area and also not suffer with scalability problem. All sensor are self-configurable capabilities so that configure itself automatically. Sensor network also vulnerable with reliability due to environmental changes, irregular terrain and hardware failure. As sensor low powered devices suffered with energy issue, low bandwidth and short storage issue. So, data gathering methods in sensor network must considered several criteria such as target area coverage, self-configuration, robustness and energy efficiency.

Basically, one should have to follow some mechanism or technique for data sensing in sensor networks. Some of these briefly describe here.

Finding suitable localized algorithms:

Each sensor node of the sensor network that follow localized algorithm for data collection from the external environment /event that collected information locally and preprocess the data locally and then send to the base station as per necessary. Localized algorithms reduce communication overhead. So, it is helps to increase life time of the network.

Data Aggregation:

Data aggregation is another challenge in sensor network. Sensor nodes are so densely deployed so that redundant data is common as neighboring nodes may be collected same data or same data may be collected by same sensor in different times. So, it is must important to filter and aggregation the data is very much necessary in data communication. some intermediate nodes that have more residual energy (remaining energy) are may be aggregate or digest concisely so that duplicate data are removed and then forward the filter data to the network or the sink node that may eventually forwarded to the base station or local server for further processing. It saves the communication overhead, saving the power and hence the increase the lifetime. This technique called as data fusion or data aggregation method.

Self-configuring capacity:

As environment or situation changes time to time i.e., dynamic in nature, topology of the sensor network should be changes dynamically as when as it requires. Self-configuration capacity (adaption) helps to adopt the environmental changes or situation that make the system autonomous.

In this article, I review the literature survey regarding data gathering mechanism used in sensor networks.

2. Data collection and broadcasting mechanism of sensor network

There are several criteria and requirements for different applications in sensor networks. According to the requirements and objective collected the information and also disseminate

or propagate the collected data throughout the network. There different data gathering mechanism and different architecture for collecting data from the targeted area or interested region. There are basically two main factors important to design a network architecture:

- a. Number of sinks and sources in the sensor networks: Sensor networks may have one or more sink node and sensor nodes depending on the application type. Basically, there are three types of applications in sensor network: multiple sources-one sink, one source-multiple sinks, multiple sources-one sink.
- b. Another important factor of dependency between network bandwidth, power/energy of the signal latency and receiving accurate information. Latency is most important criteria and needs to monitor at a level which is always beyond a threshold even it needs more energy/power. More bandwidth and low latency are always suitable for any type application of sensor networks.

3. Data gathering and forwarding mechanism in sensor networks:

There are basically three factors that are considered for categorizing data gathering and forwarding mechanism such as- (1) location of storage, (2) direction of forwarding and (3) structures of connected devices.

3.1 Storage based Data gathering in sensor networks:

Sensed data are need to be stored in proper database so that it can be process further. There are basically three types of data storage scenarios [6] such as (1) External Storage (ES), (2) local Storage and (3) Data-Centric Storage (DCS). Data are may store locally at sensor location or at external storage. Data may store at sensor node for processing that known as localization. Data may be store externally i.e., in External storage (ES) which is far from the target area.

So that we can remotely access the sensed data which far from the actual targeted area. Data also be stored based on the property of data i.e., data centric storage (DCS) store the data based on properties of data.

1. External Storage (ES):

As shown in fig 2, External storage placed externally which far from the sensor Networks and user can access the data from the external storage.

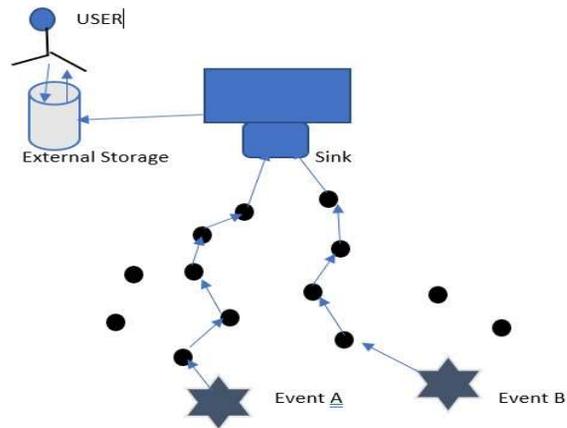


Fig 2: External Storage (ES) scenarios in Sensor networks

2. Local Storage (LS):

As shown in fig 3, data are stored at local Sensor node near to the event that reduces communication overhead. User can access the only through sink node.

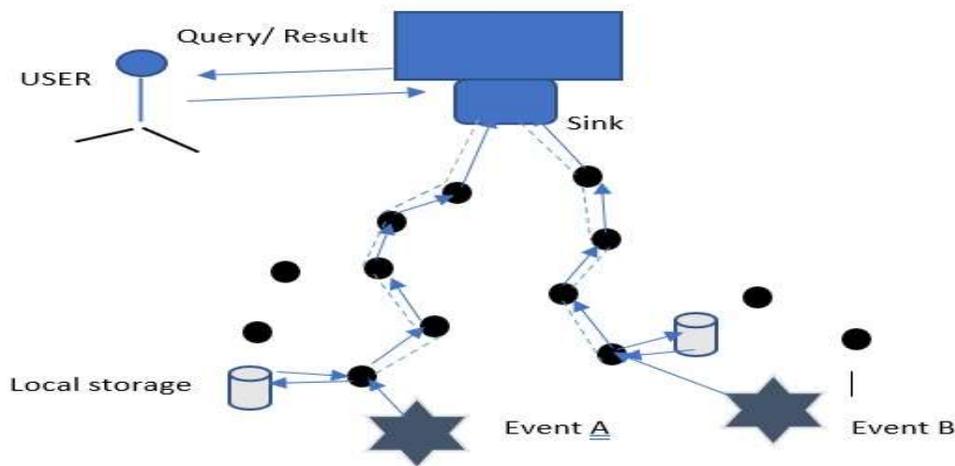


Fig 3. Local Storage (LS) scenarios in Sensor networks

3. Data Centric Storage (DCS):

As shown in fig 4, Data are store at sensor location based on type of event occur. This type of data collection known as Data Centric Storage (DCS).

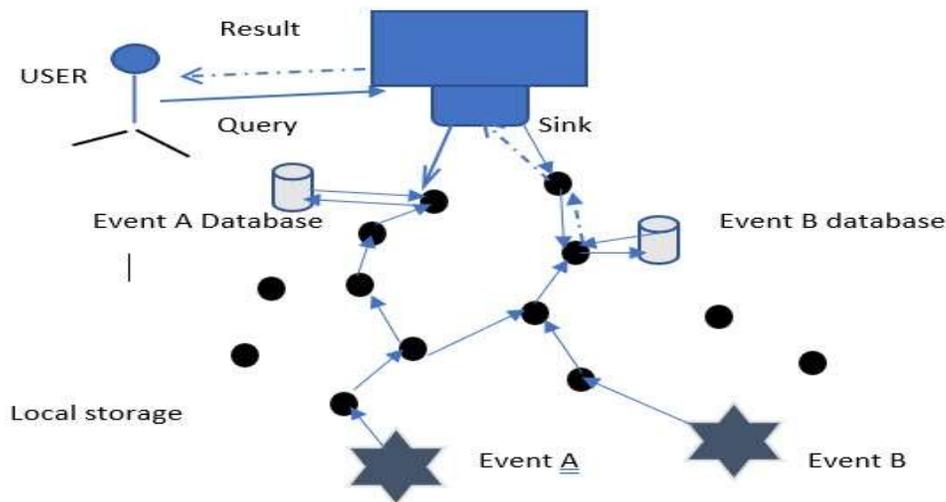


Fig 4. Data Centric Storage (DCS) scenarios in Sensor networks

3.2 Data gathering mechanism based on data forwarding/ diffusion:

In Data gathering, generally we use query and reply message in data communication. A base station or sink node sends a query to a sensor network for a specific event to be occurred and if event is occurred the sensor sends the sensed data i.e., reply to the base station for further processing. there are three types of data diffusion or forwarding technique that we follow based direction of diffusion [7]:

1. Two-Phase Pull Diffusion:

This type diffusion technique widely used where in first phase queries and replies both are broadcasted from sink or source and multiple routes are established between source and sink. in second pull phase, the sink reinforces to choose the appropriate route that have low latency and highest data rates. Two-phase pull diffusion is mainly used for the sensor network with few sink nodes and multiple sources

2. One Phase Pull Diffusion:

The two phases pull diffusion is generated high no of queries and replies for the network where exist large number of sinks and sources. Whereas, one phase pull diffusion, instead of the flooding of queries and replies.

3. Push diffusion:

A source actively broadcast the collected information when an event occurs and sinks subscribe the information with positive enforcements. This mechanism is suitable for those application where exist many sinks and few sources.it is also suitable for target tracking application where data sources are changes frequently and the routes cannot be effectively established instantly. There is a protocol named Sensor protocol for information via negotiation named as SPIN that used push diffusion.

Directed diffusion [8]:

It is a two-phase pull diffusion technique where sinks nodes are search for specific data sources on specific event and sources nodes are also try to find the sinks where the data to be transfer when a specific event occurs. This method divides into sub part- first sink forwarded request packet to all neighbors for connecting with desired data sources, then neighbors' nodes are reforwarded the data to their new neighbor if it is new one. Eventually, the data sources received request packet from the sink and finally, data transferred to the sink through the exactly reverse direction of which request packet is coming from the sink node as the request packet includes data rate and a direction to relay data (Fig 5). Initially sink send the request packet with less data rates and gradually increase it. If data sources received the request packet with high data rates, it sends the data with higher data rate and follow the updated direction for relay the data and discard previous direction of the data path.

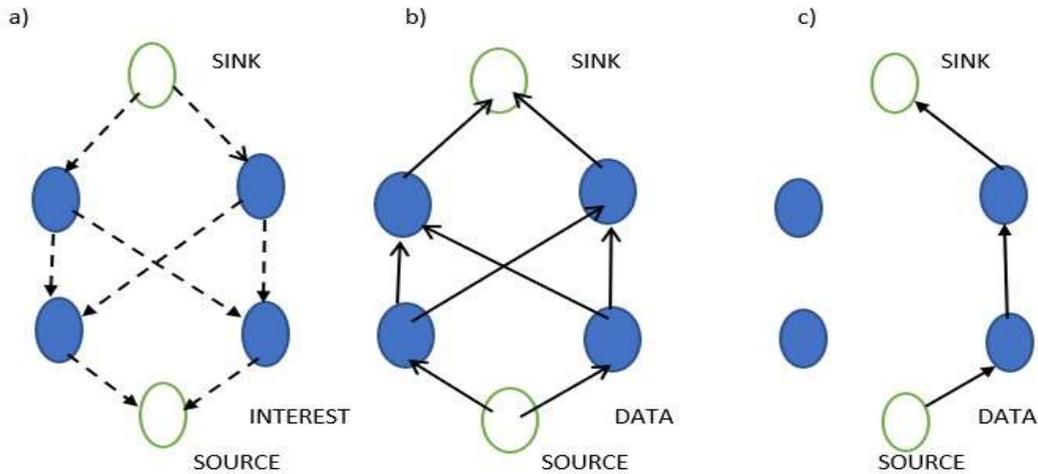


Fig 5. Three phases of directed diffusion [8] a) propagation of interest packet, b) propagation of data packet, and c) data delivery through the dedicated path

Sensor protocol for information via negotiation (SPIN):

SPIN [9,10] is a example of push-diffusion where data sources initiate the data forwarding activities. SPIN uses three stage hand shaking operations (fig 6)- such as advertisement (ADV), request for data (REQ), data message (DATA). Unlike direct flooding new data, information of new data i.e., metadata is exchanged in two advertisement and subscription phases to minimize message. When a node receives new advertisement, it replies a request data to the sender. After receiving the request data, the sender sends the real data to the receiver node from which it gets request data.

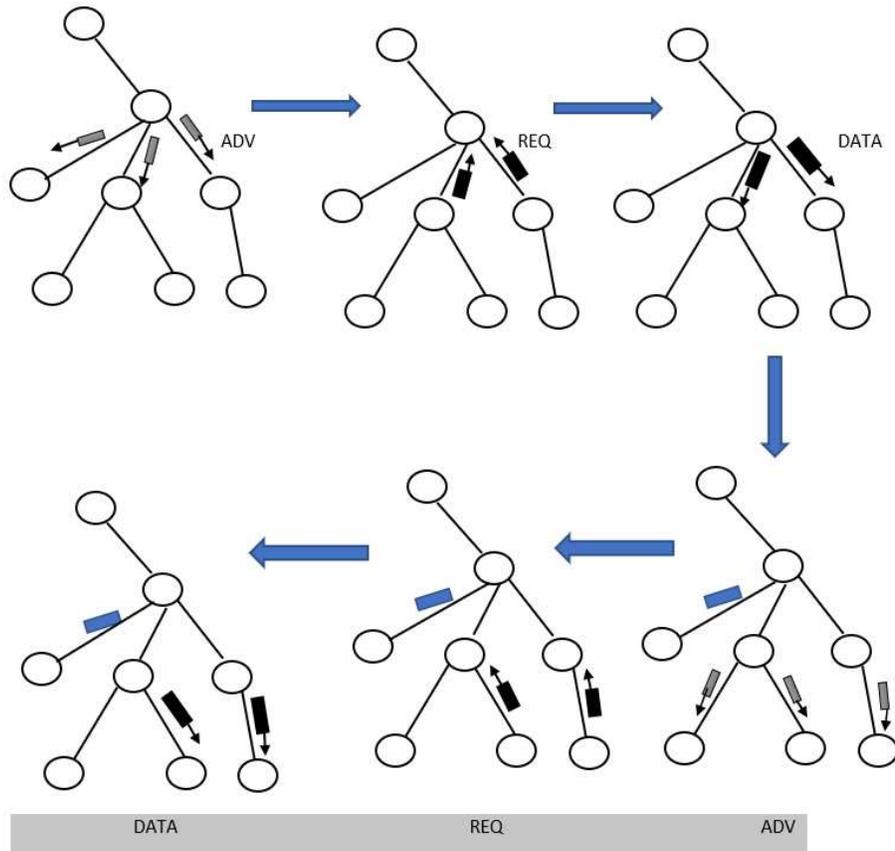


Fig 6. Three Phases of Handshaking protocols in SPIN [10]

3.3 Data gathering techniques on the basis of the structure of Dissemination

Structure of dissemination of data is play a important role in data transfer of any sensor network. Number of data sources and sink nodes depends on direction of diffusion as well as structure of dissemination in sensor networks also play important role in data transfer in sensor network from event data source to sink node of the sensor network. They are basically four types of structure dissemination structure such as Tree, Grid, Cluster, and, Chain, and their representative approaches.

1. **Tree:** Tree structure is most common structure that is used in sensor networks where the sink or base station placed on the root of the tree and the branches holds the sources where from data are to be gathered. TAG [11] and TinyDB [12] are two example that follow the tree structure for data dissemination. On the other hand, tree constructed a scenario of single source and multiple sinks where data source placed at root and sinks are hold the branches where the data to be send. Mirkovic et al. [13] proposed self-organizing multicast forwarding tree is used to disseminate the source data to multiple sinks. Here sinks are broadcast their interest packet for specific events. When the interest packet receives by a sensor node it calculates a minimum distance from some sink which will use to implement the shortest path tree. The tree starting from the root

and use reverse path to locate the sink. A sensor node with new stimulus joins the tree with the closest on-tree sensor. Thus, new branches of the tree are created.

Scalable Energy-efficient Asynchronous Dissemination Protocol (SEAD) [14]:

SEAD is a well-structured data dissemination tree which forward data from a source which is root of the tree and many sinks that are leaves of the tree. The tree is created with geographical routing protocol. If a mobile sink wants to receive the data from a data source, it connects through one of the nearest sensor nodes called Access node in the dissemination tree. Similarly, Home agent in mobile IP (Internet Protocol), the access node relay data to the sink as an anchor. When sink is out of threshold range of access node it informs the original access node by sending PathSetup Message and the mobile sink joins with new access node. A source node copies its data at different replicas that reduces number of messages transmitted over the tree.

- Grid:** Like the SEAD, two tier data dissemination (TTDD) [15] is developed for a single source and multiple mobile sinks. TTDD used grid structure for data dissemination. Fig 7 shows how grid structure generated from data source. In top tier, a source detects an event early by constructing grid structure where sensors nearer to the grid point elected as dissemination node. Whereas, in lower tier a mobile sink sends a interested packet data or receives data from nearest grid point on local grid. In TTDD, if a mobile sink moves to other grid, the sink quickly construct grid structure and so information access with less delay. Target tracking in the battle field is as well-suited example of TTDD.
- Cluster:** When data is generated by sensors, it processed first locally and then it eventually transfers to the sink. Cluster structure is better for this purpose. Several hierarchical cluster methods [16,17,18] are used popularly. One of the most suitable cluster approaches is low energy adoption clustering hierarchy (LEACH) [16]. It is a two-level clustering approach where sensors are portioned into clusters. Each sensor willing to become a Cluster head with some probability at each round. Once a sensor elected as a Cluster head (CH), it sends a message to all the neighbor node so that remaining sensor are uses minimum transmission power to join with their closest CHs. Cluster Heads (CHs) are uses Time Division Multiple Access (TDMA) protocol to allocate slot for cluster member to relay their sensed data to their CH. Then CHs compress their receive data and sends the aggregated data to the sink. LEACH considers CHs are apply high radio power to transfer the data base station or nearest CHs. In [17], authors design a hierarchical clustering approach similar to LEACH with considering load balancing among sensor and energy efficiency. The author of [19], design a dynamic and decentralized clustering mechanism for target tracking. Cluster form dynamically and operate in decentralized fashion. CHs are activated with some predetermine threshold values. Active cluster head broadcast a message for asking nearer sensors to join the cluster and forwarded their sensed data to the cluster head. They uses Voronoi diagram to show 1) how CHs interact with each other to select one Cluster Head (nearest to the target tracking) is active, 2) how the active cluster head request to get the sensor information from those that are closest to the target location and 3) both the request packet from cluster head and sensed packet from sensor to cluster head do not significant collided to each other.

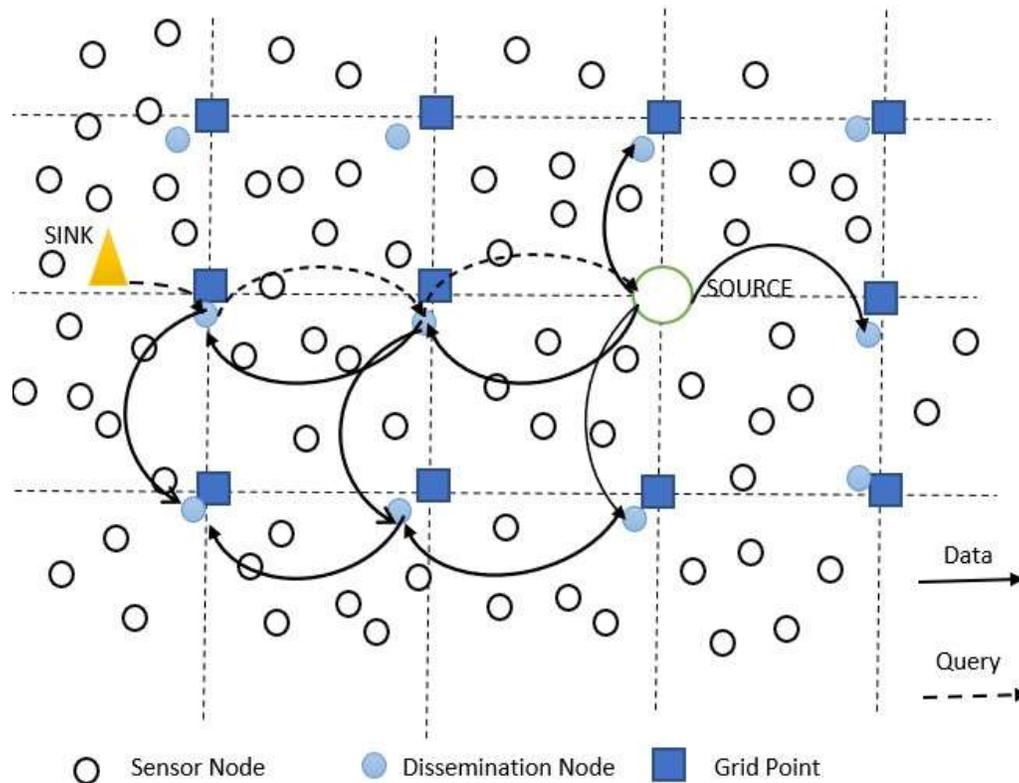


Fig 7. Two Tier Data Dissemination (TTDD) grid structure [15].

4. **Chain:** When latency is more important rather than energy efficiency or bandwidth, Chain structure is more efficient to aggregate the data along its path and send to a sink. The Power-efficient gathering in sensor information system (PEGASIS) [20] is used to aggregated the data collected from all sensors of the network. Here only one leadership is elected each time and the leadership is rotate among all the sensor. Here all the leader directly connected with all the sensor such that it follows mesh topology which complete graph. From the starting sensor node at the one end of the chain, data are propagated and aggregated along the chain and other end of the chain continue data dissemination and aggregation process. After those aggregations from two end comes to the leader and the leader directly send whole aggregated data to the sink.

Conclusion:

In this article, I briefly explain different strategies to gathering data by the sensor network. I also explain data gathering mechanism based on storage location whether data storage are local storage (LS), External storage (ES) or Data Centric Storage (DCS). Then different data diffusion technique such as two phases pull diffusion, one phase pulls or push diffusion mechanism discussed briefly. Directed diffusion and SPIN protocol explained how sink and sources searches to each other for diffuse their request or data packets. After that I discuss

different data gathering strategies based on the structure of the dissemination or connection of physical nodes such as tree, grid, cluster or chain structure of the sensors or nodes. I also explain different protocol such as SEAD, TTDD, LEACH and PEGASIS for each of the structure of dissemination of data. This article helps the researcher to know about the different data gathering strategies and data dissemination techniques in sensor networks.

Conflict of Interest: There is no conflict of interest in this article.

References:

- [1] A. Mainwaring, J. Polastre, R. Szewczyk, D. Culler, and J. Anderson. Wireless sensor networks for habitat monitoring. In Proceedings of the 1st ACM International Workshop on Wireless Sensor Networks and Applications (WSNA '02), Atlanta, Georgia, September 2002.
- [2] F. Zhao, J. Shin, and J. Reich, Information-driven dynamic sensor collaboration for tracking applications, *IEEE Signal Processing Magazine*, 19:68–77, March 2002.
- [3] S. S. Intille. Designing a home of the future. *IEEE Pervasive Computing*, 1(2):80–81, April–June 2002.
- [4] I. A. Essa. Ubiquitous sensing for smart and aware environments. *IEEE Personal Communications*, 7(5):47–49, October 2000.
- [5] P. Bonnet, J. Gehrke, and P. Seshadri. Querying the physical world. *IEEE Personal Communications*, 7:10–15, October 2000.
- [6] S. Ratnasamy, B. Karp, S. Shenker, D. Estrin, R. Govindan, L. Yin, and F. Yu. Datacentric storage in sensor networks with GHT, a geographic hash table. *Mobile Networks and Applications*, 8(4):427–442, August 2003.
- [7] J. Heidemann, F. Silva, and D. Estrin. Matching data dissemination algorithms to application requirements. In Proceedings of the 1st International Conference on Embedded Networked Sensor Systems (Sensys '03), Los Angeles, California, November 2003.
- [8] C. Intanagonwiwat, R. Govindan, and D. Estrin. Directed diffusion: A scalable and robust communication paradigm for sensor networks. In Proceedings of the 6th Annual IEEE/ACM International Conference on Mobile Computing and Networking (MobiCom '00), pages 56–57, Boston, Massachusetts, August 2000.
- [9] J. Kulik, W. Heinzelman, and H. Balakrishnan. Negotiation-based protocols for disseminating information in wireless sensor networks. *Wireless Networks*, 8(2/3):169–185, March–May 2002.
- [10] W. Heinzelman, J. Kulik, and H. Balakrishnan. Adaptive protocols for information dissemination in wireless sensor networks. In Proceedings of the 5th ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom '99), Seattle, Washington, August 1999.
- [11] S. Madden, M. J. Franklin, J. M. Hellerstein, and W. Hong. TAG: A Tiny Aggregation service for ad-hoc sensor networks. In Proceedings of the 5th Symposium on Operating System Design and Implementation (OSDI 2002), Boston, Massachusetts, December 2002.
- [12] S. Madden, M. J. Franklin, and J. M. Hellerstein, and W. Hong. The design of an acquisitional query processor for sensor networks. In Proceedings of the 2003 ACM SIGMOD International Conference on Management of Data, pages 491–502, San Diego, California, June 2003.

-
- [13] J. Mirkovic, G. P. Venkataramani, S. Lu, and L. Zhang. A self-organizing approach to data forwarding in large-scale sensor networks. In Proceedings of the IEEE International Conference on Communications (ICC 2001), Volume 5, pages 1357–1361, Helsinki, Finland, June 2001.
- [14] H. S. Kim, T. F. Abdelzaher, and W. H. Kwon. Minimum-energy asynchronous dissemination to mobile sinks in wireless sensor networks. In Proceedings of the 1st International Conference on Embedded Networked Sensor Systems (Sensys '03), Los Angeles, California, November 2003.
- [15] F. Ye, H. Luo, J. Cheng, S. Lu, and L. Zhang. A two-tier data dissemination model for large-scale wireless sensor networks. In Proceedings of the 8th Annual International Conference on Mobile Computing and Networking (MobiCom 2002), pages 148–159, Atlanta, Georgia, September 2002.
- [16] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan. Energy-efficient communication protocols for wireless microsensor networks. In Proceedings of the 33rd Hawaii International Conference on Systems Science (HICSS), Volume 8, page 8020, Maui, Hawaii, January 2000.
- [17] S. Bandyopadhyay and E. J. Coyle. An energy efficient hierarchical clustering algorithm for wireless sensor networks. In Proceedings of INFOCOM 2003, San Francisco, April 2003.
- [18] C.-C. Shen, C. Srisathapornphat, and C. Jaikeo. Sensor information networking architecture and applications, IEEE Personal Communications, 8(4):52–59, August 2001.
- [19] W.-P. Chen, J. C. Hou, and L. Sha. Dynamic clustering for acoustic target tracking in wireless sensor networks. IEEE Transactions on Mobile Computing, 3(3):258–271, July–September 2004.
- [20] S. Lindsey, C. Raghavendra, and K. M. Sivalingam. Data gathering algorithms in sensor networks using energy metrics. IEEE Transactions on Parallel and Distributed Systems, 13(9):924–935, September 2002.