International Journal of Database Theory and Application
Vol.8, No.4 (2015), pp.179-182
http://dx.doi.org/10.14257/ijdta.2015.8.4.18

Distributed Data Storage in Wireless Sensor Networks

Gao Weimin and Zhu Lingzhi

School of computer and information science, Hunan institute of technology, Heng yang 421002, China
E-mail :gwmhy@163.com

Abstract

This paper studied the techniques of distributed data storage in wireless senor networks. Firstly, the challenge and the need for such techniques were summarized; Secondly, some representative distributed data storage and retrieval schemes were introduced in detail; finally, the future research directions and open issues were pointed out.

Keywords: *Wireless sensor networks; distributed data storage; systems*

1. Introduction

Wireless sensor network (WSN) is composed of a large number of sensor nodes with small volume, low cost, and with wireless communication, sensing, data processing ability. These nodes are formed by means of network self-organizing, which can real-time senses, collects and processes the data that is interested by people in surrounding environment with the aid of the various forms of sensors embedded in the node. Then we can get more detailed and accurate information for the observers. WSN is considered to be one of the most important technology in 2lth century, which has potential practical value in military defense, industry and agriculture, urban management, biological, medical, environmental monitoring, disaster relief, anti terrorism, dangerous areas remote control and many other important areas [1, 2].

WSN is a data-centric [3], so the data storage technology is a key technology of WSN. But the traditional method of external data storage based on a single base station deployment will cause communication bottlenecks and energy consumption imbalance, but bring more base stations will increase the cost of network .Because the sensor node has a small storage capacity, low computing power, small communication range, limited energy and other characteristics of the traditional network data storage and retrieval solutions and protocol can not be applied to the WSN. Thus, we need to depend on these characteristics of sensor nodes design suitable WSN data storage scheme. In view of this, researchers have made deep researches into WSN distributed data storage technology, and made some preliminary results.

For WSN deployed base stations, allowing users access to the event for the data of interest, the traditional approach is to collect the date and put the data-aware nodes to the base station generates for users to query [4-6]. However, this method has three drawbacks: First, when there are large amounts of data generated in the network, and transfer to the base station at the same time, the surrounding base stations will occur communication bottlenecks ,and it also will cause the problems of the node energy consumption imbalance; Second, because the user may only interested in part of the collected data, but indiscriminately all sensory data will be sent to the base station, it will waste energy of the sensor nodes [7]; The third, the method of data collection rely on this base too much so that in some inconvenience deployed base stations in the harsh environment can not be applied.

However, if the WSN as a distributed database, all the sensory data generated in WSN nodes stored internally, through the rational design of distributed data storage mechanism,

ISSN: 2005-4270 IJDTA

these problems can be solved basically. First, the data stored in the network to avoid or mitigate communication bottlenecks surrounding nodes and avoid the base station load imbalance; secondly, the data storage can provide the users they interested part in the network, greatly reducing the amount of data transmission and reducing the energy consumption of nodes; Finally, the data stored in the network nodes can effectively reduce dependence on the base station, the user can initiate queries in every way, WSN nodes will make a deal with queries and return the result to the user.

Currently, the researchers in the study of WSN distributed data storage technology focused on the following aspects:

(1) Distributed data storage technology research for specific types of queries. Common types of data query temporal data query (spatio-temporal data query) [8, 28], the scope of the query (range query) [9], multi-resolution data Query (multi-resolution data query) [10], the approximate location of the query (approximate location query) [11], historical data query (historical data query) [12], the extreme range queries [13] as well as the approximate Skyline queries [14] and the like.

(2) Data distribution and found [15-18]. In some WSN applications, data consumers (users) need efficient access to real-time data-aware data producers (sensor nodes) generated. Solve the problem of data distribution and discovery usually takes push-pull technique [16]. Push technology means a data producer will perceived within the network flood event, the consumer can obtain data in a local sensory data; and Pull technical means, the data generated in the producer local storage, data consumer the fully distributed queries in the flooding, return query results when the query encountered interesting data. However, simply relying on technology or Push Pull technology to solve the problem of data distribution and discovery needs in the whole network broadcast news, obviously inefficient, data distribution and discovery issues need to be resolved is how to find success in ensuring data achieve a better balance between conditions push and pull.

(3) Adaptive data access [19-20]. In some distributed data storage system, a node is responsible for data storage what is determined in advance. Data producer sends the data to the data storage node to store, data consumer sends the query to the data storage node for data retrieval. Selecting the data storage node is a key factor in this type of distributed data storage and retrieval solutions it affect performance advantages and disadvantages. Depending on data producer position, the data generation rate, the location of the data consumers, the frequency of issuing a query request, and the network topology and other factors adaptive selection data storage node, is of such a distributed data storage system to be solved key issues. There are a number of distributed data storage system, where is responsible for the non-deterministic, each node based on the actual situation whether it become data storage node. With this strategy WSN data storage adapted, it used for target tracking.

(4) Network data encoding and recovery [21-23]. In some harsh environments, the base station is not convenient deployment. The energy of sensor nodes is limited, as time progresses, some of the nodes may be dive due to energy depletion. In order to recover from the remaining nodes in the network all the data generated, the researchers propose a coding method based on distributed data storage. Node-based encoding of data storage mechanism needs to solve, the problem what is produce the perception of how the data transmission cost as little as possible to perceive the rest of the data diverge to each node in the network, the node receives a packet-aware what kind of coding using encoding method, and how to use the data as little as possible on the network nodes to restoring all of the data and so on.

2. A Typical Wireless Sensor Network Distributed Data Storage System Analysis

2.1 DCS and GHT System

DCS [24] the main idea is to name the data, the perception of the same name by the corresponding data stored in the same node. In order to find an event name corresponding to the data storage node, S. Ratnasamy et al proposed a method called GHT in the literature [25]. GHT using geographic hash table will be mapped to a name of a sensor network location from the location closest to the node (called home nodes) to store the name corresponding perception data. In order to avoid death or home node network topology changes in data loss, GHT surrounding update protocol using PRP (perimeter refresh protocol), the home node periodically sends data back to the hash location neighboring nodes; home node in order to avoid excessive load weight, GHT replication technique using a structure, in the network to find a plurality of image points for the home node, the data generating node will send data to the perceived mirror node closest to their storage.

In fully distributed query flooding, to avoid the data retrieval GHT but there are some problems: First problem is the poor distance sensitivity, when the position data between data producers and consumers close. You must send a request to the query itself may be relatively far away from the home node; secondly it may cause communication bottlenecks and hot spots (hot spot) problem. When a named data query multiple data nodes are frequently accessed, or when the network belonging to the named data more, it would be a burden to the named data node corresponds to home too heavy, will influence the life of network nodes. At the same time, home node also can occur communication bottlenecks occur; structure again, in order to avoid home node and its neighboring nodes deaths lead to data loss and the introduction of replication technology will bring greater communication overhead; Finally, GHT not support multiple types of data aggregation queries efficiently.

2.2 DC System

For DCS hot issue, Keng-Teck Ma et al proposed a divergence caching techniques based data-centric data storage and retrieval solutions [26], using the system of the program, also known as DC system (Diffuse Caching, referred to as DC). The main method of dispersing the burden of the program is the home node, home node selected subset of nodes from a neighbor node as a cache node, and sends the data itself is stored on the cache node; cache node can also elect part of their neighbors cache node as home node, and to the elected cache node itself forward data received from the home node; when the cache node receives a query, the query cache node can provide query results directly. Because home node sends event data to the cache node consumes energy, home or home node in the choice of the cache, you need to determine the selection of a node as its potential benefits (potential saving) when the data cache node is greater than its potential waste of energy (potential waste, refers to the home node sends data to the cache nodes need to consume energy), and only when the potential benefits outweigh the potential energy waste, it will select a node as a data cache node.

Experimental results show that, DC can effectively reduce the burden of the home node, and can reduce the transmission overhead message compared with GHT. However, the program only when retrieving data to reduce the burden of the home node, when the event data to generate a higher rate, the burden of the home node will still be relatively large, because the home node still needs to store and forward all corresponding event type of the event data; in addition, this option will have less storage space utilization efficiency, because a data block needs to occupy multiple nodes Storage space.

2.3 C-DCS System

HC Lee, who pointed out that moving the traditional data-centric data storage and retrieval solutions (e.g. GHT) is not well supported nodes and data storage, there is not necessarily a hash function to map geographic location position where is bound to focus on the hash node forwards a week, according to GPSR routing protocol for data forwarding to reach the home node, this will waste of energy [27]. To solve these problems, HC Lee et al proposed a distributed data storage and retrieval scheme what is called a C-DCS [27]. C-DCS the sensor network space into equal-sized clusters C, and all the nodes are divided into two categories: the mobile node and stable node. C-DCS provides that only stable node can participate in data storage and routing. To this end, each node needs to store a stable neighbor node list, when the data forwarding, select only stable nodes in the neighbor list as the next hop node. In the C-DCS, and between the nodes within each cluster stability through consultation elect a stable node as the cluster head. When a node has data generated, the first use of geographic data generated by the hash function is mapped to a location, and then press the geographic routing protocol data generated will be forwarded to the location, and ultimately by the geographical coverage of the cluster in storing the first event data. Similar data retrieval process and data storage process, from data retrieval query data to map the location of the node sends a query to the final query by covering the head node cluster in a data mapping the location of the query results are returned to the data retrieval node. C-DCS data storage and retrieval process is shown in Figure 1. Mobile node when the data storage and retrieval, it need to broadcast to the neighbors for help sending REQ message, after a neighbor node receives REQ message, if you wish to help move forward the data node, the mobile node returns to agree to help APV message. APV message after the mobile node receives, from which you can select the location closest to the target node as the next hop node. This process is shown in Figure 2. In order to reduce the energy consumed by data transmission, C-DCS using double-aggregation techniques (two levels of aggregation) to improve the above-described data storage and retrieval methods: event data generated by all nodes of the cluster is first sent to the cluster head aggregation, and then by the local cluster head will bring together the results ,they are sent to the cluster head covering data storage locations mapped clusters; data retrieval, cluster head receives a query for data stored once again brought together, and the results are returned to the convergence inquirer. Improved Data storage and retrieval process shown in Figure 3.

C-DCS can solve the node mobility and other issues effectively. Experimental results show that, in terms of relative to the DCS, C-DCS can reduce the number of messages transmitted over the network significantly, but it also possible to reduce the number of hot spots in the network. Disadvantages of C-DCS is in a stable state of the node and neighboring nodes need to exchange information periodically to update the neighbor node table, increasing network maintenance overhead.

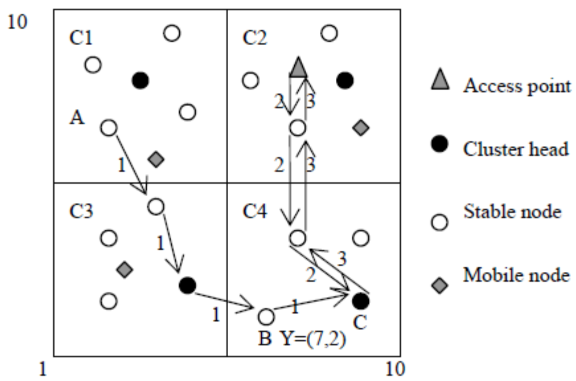

Figure 1. C-DCS Data Storage and Retrieval Process

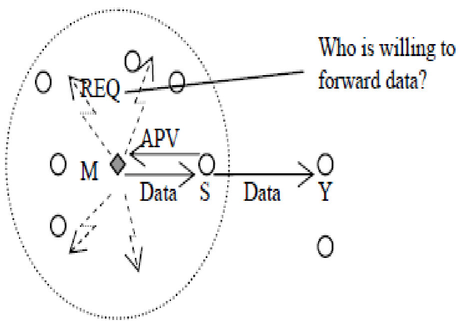

Figure 2. C-DCS Mobile Node Route

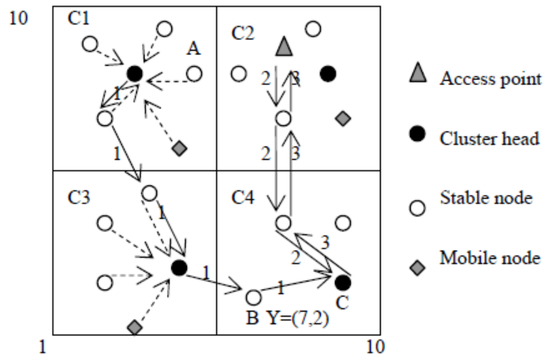

Figure 3. Improved C-DCS

2.4 ODS-NDS System

For how to determine the data rate of data generation under conditions and other factors known nodes optimal storage location problem, Wei Zhao et al proposed an adaptive data storage and retrieval scheme based on the optimal storage location (hereinafter referred to the program for ODS-NDS [19] .ODS-NDS first relational data producers and data consumers to build "one to one", "many to one", "many to many" model, and then based on these three models, the total cost of data access for a minimum goal, it give full consideration to the data generation rate data producers and data consumers send a query rate and other factors, we propose two adaptive data storage node selection algorithm: the global optimal greedy algorithm ODS (optimal data storage) and a local optimum approximation algorithm NDS (near-optimal data storage) .ODS main idea is that one by one node test, Behold a node as a data storage node enables network-wide data access can always minimum consumption; NDS data is first introduced total access cost function, and then find the optimal data storage location for the sake of making the data problem into total access cost function value of the minimum between x and y.

The advantages of ODS-NDS is that it is able to dynamically adjust the data storage location based on the rate at which the number of publishing network of producers and consumers, as well as data generated by location and query request, in order to achieve the overall cost of the optimal data access effect. However, due to ODS-NDS only select an optimal node as a data storage node, when large amounts of data in the network or network generated queries, even in the best position of the data storage node neighbors help store data, it will be a great burden on the storage node. If frequent replacement data storage node approaches to solve this problem, it will increase the cost of updating the storage node, because node position information of the base station needs frequent broadcast new storage node.

2.5 ADS System

Mentioned various adaptive data storage strategy although in different ways, but they have one thing in common, that is the entire sensor network using the same kind of data storage and retrieval strategies. However, in sensor networks, may occur more frequently in some area of the data query, and the data generation rate is low relatively , some area of the data generation rate is high relatively , but less than the number of times being queried, and therefore, the entire region using the same data storage and retrieval solutions may not be efficient. Based on this thinking, Guo-Jong Yu, an adaptive data storage conversion strategies ADS [20] .ADS's main idea is to divide the whole sensor network area into grids and each grid between the same network cells at different time periods are different data storage and retrieval strategies. ADS in each mesh data storage solutions can convert between the two data storage solutions: local storage and data-centric storage. Whenever the need to determine whether it should be stored program conversion, each grid should be selected as the number of nodes in the event of a head node (leading node), according to the head node by inside grid before time periods k generated and received event number of queries to predict the next time period will produce a number of events $Event_{num}^{t+1}$ and the number of query requests will receive $Query_{num}^{t+1}$. If $Event_{num}^{t+1}$ and $Query_{num}^{t+1}$ satisfies a formula:

$$Event_{num}^{t+1} > (Query_{num}^{t+1} + 1) * \sqrt{n} + T \quad (1)$$

Then switch to local storage mode; if the equation: $Event_{num}^{t+1} < (Query_{num}^{t+1} - 1) * \sqrt{n} - T$, switch to a data-centric data storage mode. When data is retrieved, the base station need to be dissipated to the query request target grid, the grid used if the target data-centric data storage mode, directly to the data center (data centric node) node query; if the target grid uses local storage, then you need to check all nodes within the grid.

ADS adaptive strategies based on local data storage rate and the rate of partial query request received from the local area of event data generated, to achieve a local optimization of data storage and retrieval. However, ADS flaw is that it only focuses on the local optimization, but it is not bound to consider global optimization problems. In ADS, for which a node A the grid where no matter what kind A data storage strategy, its internal storage A grid inevitable event data generated in. A distant data of interest if the network node to node A node is generated, and send queries to these nodes rates and relatively large, the use of ADS protocol for data access is inefficient clearly.

2.6 DRIB-RDRIB System

DRIB-RDRIB system [31] is a distributed data storage system based on data transmission trajectory. DRIB first builds a virtual boundary. Virtual boundary consists of four anchor nodes and four segments constitute the axis shown in Figure 4, X, Y, Z, Z 'four anchor nodes, ZX, XZ', Z'Y, and YZ axis for the four segments. When data is stored, if the data producer to the virtual boundary within the bounded region (including the boundary nodes) the minimum number of hops of nodes is greater than 1, the data producer to the first direction closest to its borders duplicated data. When the minimum number of hops away from the event data is transferred to the virtual boundary around the area of the nodes on the node is less than or equal to 1, respectively, and copy data to the side edge YZ XZ 'direction; if the distance between data producers virtual boundary around the area the minimum number of hops within the node is less than or equal to 1, then the copy data directly to the edge side YZ and XZ 'direction. When data is retrieved, the transfer process is similar to the transmission process of data query request with the stored event data, the difference is that data retrieval, the query request to be transferred to the side YZ 'and XZ directions, respectively. From Figure 5, when the network boundary shapes rules, copy track query request transmission trajectory intersects with the inevitable event data.

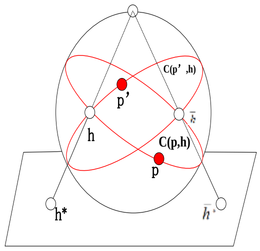

Figure 4. Double-ruling Based on the Spherical

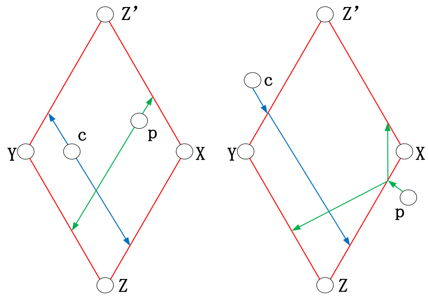

Figure 5. DRIB Intuitive Schematic Diagram

DRIB problem is that, when the network boundary irregular shape, copy or transmission of event data track locus query request is not entirely contained within the range surrounded by the virtual boundaries, this can lead to the transmission path and the query request event data replication track disjoint. To solve this problem, Chia-Hung Lin et al proposed RDRIB algorithms. RDRIB algorithm is adding some constraint rules on the basis of DRIB algorithm, these constraints can be sure to copy the track rules and event data transmission trajectory queries inevitably intersect. Experimental results show that, RDRIB only can guarantee 100% success rate inquiries, but also can reduce the cost of data transmission effectively.

2.7 EDFC and ADFC System

EDFC and ADFC [21] the use of coding techniques for distributed data storage. In a total number of nodes is N, the number of event data source node for k, sensor networks, in order to be able to use the greater probability as few nodes information recovered k data on the source data nodes to improve continuity of data, Lin et al proposed a random walk (random walk) [21] as well as distributed data distribution and storage algorithm LT codes in the literature [21] EDFC and ADFC. The basic idea EDFC algorithm, first, each node according Robust Soliton distribution of [21] to select its own encoding degree d, then each data source node b randomly selected according to random walk path of forwarding of the source packet, b Calculation the formula is:

$$b = \frac{N\sum_{d=1}^{k} x_d d\,\mu(d)}{k} \qquad (2)$$

Where $x_d d$ coding node d is the number of degrees of actual receipt of the source data packets $\mu(d)$ shows the coding node degree d of proportion.

After all of the source data packets are forwarded finished, each node randomly received packets from the source, select the number equal to the source encoding of the data packet itself is encoded. Data collectors can use any kind of fountain code decoding algorithm for decoding. Although EDFC algorithm to assign each node at the beginning of the coding of the accord Robust Soliton distribution, but because of the randomness of random walk brings, each node does not meet the actual coding of the Robust Soliton distribution, according to the formula to calculate the value of b may be higher than the actual value it needs to be large, thereby increasing the energy consumption of the propagation of a source packet. To solve this problem, Lin et al proposed ADFC algorithm, similar ADFC algorithm EDFC algorithm, except that, Lin et al., Designed a new distribution function in ADFC encoding algorithm, and in accordance with this new distribution function calculate the value of b. Theoretical analysis and experimental verification showed that compared with EDFC algorithm, ADFC algorithm can significantly reduce the energy consumption of the spread of the source packet, while

ADFC algorithm has a higher rate of successful decoding. Literature [22-23] also made a number of coding techniques based on distributed data storage and retrieval solutions, space limited , this article will not go into details.

3. Conclusion

This article explores the problem of WSN in distributed data storage, detailing several typical WSN distributed data storage system. By comparison and analysis of these systems, has not yet found the existence of a distributed data storage system can adapt to all applications of WSN; at the same time, the existing data storage strategies are still some drawbacks. For now, WSN distributed data storage technology need for further research in the following areas:

(1)As much as possible to reach "to minimize the overall cost of data access" and "balance node energy consumption," these two goals at the same time. At present, many research results about WSN distributed data storage technology is either total network data access cost is relatively small load balance but rather poor, either, achieving load balance, but it increases the total network data access overhead. WSN distributed data storage technology in line with the practical application is necessary to ensure the overall cost of data access as small as possible, but it also to ensure that the load node as balanced as possible.

(2)Improve the efficiency of storage space, improve the utilization of storage space imbalances .Because of the limited storage space in the WSN nodes , therefore, we need to improve the utilization efficiency of the nodes of storage space, which uses the same memory space contains as much information. For this reason, data compression technology worth to further study. Storage space utilization refers imbalances in WSN, means there may arise some nodes due to excessive amount of data stored in the data caused by overflow, while the other part node phenomenon but has more storage space. Data distribution technology is an effective way to solve this problem, we need further study.

(3)Implementation does not depend on geographic information systems, distributed data storage. Traditional WSN distributed data storage technology needs to support positioning technology, and now most of the positioning accuracy of the positioning algorithm is not very satisfactory, and if for each node is equipped with a sensor GPS, it will increase the cost. Therefore, further study is not dependent geographical information distributed data storage technology.

(4)Provide support for node mobility. Currently, WSN distributed data storage protocol basically assumes that the sensor nodes and the network is static. But in some, such as the battlefield, logistics applications, mobile nodes are likely to occur. The mobile node, in particular a mobile data storage node, easily move back and forth due to data inconsistency caused by the data storage node is lost. The new distributed data storage scheme WSN node mobility should provide better support.

(5)Improve the security of distributed data storage. So far, security problems of distributed data storage and access areas rarely cause for concern. Prior studies of people WSN security focused on the security of the network communication, such as key management, message authentication, secure time synchronization and positioning, monitoring attacks. In some applications, WSN are vulnerable to attacks such as the Byzantine, data contaminated with attack. Therefore, we need to design secure distributed data storage strategy to ensure data confidentiality, independence, integrity and effectiveness.

Acknowledgment

This work was financially supported by National natural science foundation of china (61103245) Hunan province science and technology plan projects (2014FJ6026); Hunan province office of education planning issues (No.XJK014CXX004) ;Science and technology project in hengyang(No.2014KG61);Hunan institute of technology reform project (JY201425) ;philosophy and social science foundation in Hunan province(14YBA129).

References

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, “A survey on sensor networks”, IEEE Communications Magazine, vol.8, no.40, **(2002)**.
[2] F. Y. Ren, H. N. Huang and C. Lin, “Wireless sensor networks”, Journal of Software”, vol.7, no.14, **(2003)**.
[3] J. Z. Li and H. Gao, “Survey on Sensor Network Research”, Journal of Computer Research and Development, vol.1, no.45, **(2008)**.
[4] Q. Zhang, Z. P. Xie, B. Ling, W. W. Sun and B. L. Shi, “A Maximum Lifetime Data Gathering Algorithm for Wireless Sensor Networks”, Journal of Software”, vol.11, no.16, **(2005)**.
[5] M. Liu, J. N. Cao, G. H. Chen, L. J. Chen, X. M. Wang and H. G. Gong, “EADEEG: An Energy-Aware Data Gathering Protocol for Wireless Sensor Networks”, Journal of Software, vol.5, no.18, **(2007)**.
[6] J. B. Liang, J. X. Wang, T. Sh. Li and J. E. Chen, Maximum Lifetime Algorithm for Precise Data Gathering Based on Tree in Wireless Sensor Networks Journal of Software, vol.9, no.21, **(2010)**.
[7] L. Yan, W. L. Tao, K. Liu and Y. Xu, “The Research of Data Storage and Retrieval Scheme for Wireless Sensor Networks”, International Journal of Intelligent Engineering and Systems, vol.2, no.2, **(2009)**.
[8] M. Aly, A. Gopalan, J. Zhao, A. M. Youssef, “STDCS: A Spatio-Temporal Data-Centric Storage Scheme For Real-Time Sensor net Applications”, In: Proceedings of the Fifth Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks. IEEE Computer Society, **(2008)**; California, America.
[9] X. Li, Y. J. Kim, R. Govindan and W. Hong, “Multi-dimensional Range Queries in Sensor Networks”, In: Proceedings of the 1st international conference on Embedded networked sensor systems(Sensys’03), **(2003)**; New York, America.
[10] L. Xie, L. J. Chen, D. X. Chen and X. Li, “Ring-Based Multi-Resolution Data Storage for Sensor Networks”, Journal of Software, vol.12, no.20, **(2009)**.
[11] J. L. Xu, X. Y. Tang, and W. C. Lee, “A New Storage Scheme for Approximate Location Queries in Object Tracking Sensor Networks”, IEEE Transactions on Parallel and Distributed Systems, vol.2, no.19, **(2008)**.
[12] C. Y. Ai, R. Y. Du, M. H. Zhang and Y. S. Li, “In-network Historical Data Storage and Query Processing Based on Distributed Indexing Techniques in Wireless Sensor Networks”, In: Proceedings of the 4th International Conference on Wireless Algorithms, Systems, and Applications, **(2009)**; Berlin, Germany.
[13] S. G. Xiong, J. Z. Li, L. Chen and X. B. Wang, “Querying Peak Regions in Wireless Sensor Networks”, Journal of Software, vol.8, no.21, **(2010)**.
[14] L. Q. Pan, J. Z. Li, J. Z. Luo, “Approximate Skyline Query Processing Algorithm in Wireless Sensor Networks”, Journal of Software, vol.5, no.21, **(2010)**.
[15] D. Braginsky and D. Estrin, “Rumor routing algorithm for sensor networks”, In: Proceedings of the ACM International Workshop on Wireless Sensor Networks and Applications, **(2002)**; Atlanta, America.
[16] X. Liu, Q. F. Huang and Y. Zhang, “Balancing push and pull for efficient information discovery in large-scale sensor networks”, IEEE Transactions on Mobile Computing, vol.3, no.6, **(2007)**.
[17] C. H. Lin, J. J. Kuo and M. J. Tsai, “Reliable GPS-Free Double-Ruling-Based Information Brokerage in Wireless Sensor Networks”, In: Proceedings of the 29th conference on Information communications, **(2010)**; San Diego,CA|USA.
[18] G. T. Shi, J. Zheng, J. F. Yang and Z H. Zhao, “Double Cross: A Double-Blind Data Discovery Scheme for Large-Scale Wireless Sensor Networks”, In: Proceedings of the 2011 IEEE International Conference on Communications, **(2011)**.
[19] Z. C. Yu, B. Xiao and S. G. Zhou, “Achieving Optimal Data Storage Position in Wireless Sensor Networks”, Computer Communications, vol.1, no.33, **(2010)**.
[20] G. J. Yu, “Adaptive Storage Policy Switching for Wireless Sensor Networks”, Wireless Personal Communications, vol.3, no.48, **(2009)**.
[21] Lin YF, Liang B, Li BC. Data Persistence in Large-scale Sensor Networks with Decentralized Fountain Codes. In: Proceedings of the 26th IEEE International Conference on Computer Communications (INFOCOM 2007). (**2007**)May6-12: Anchorage, Alaska, USA
[22] J. B. Liang, J. X. Wang and J. E. Chen, “An Overhearing-based Scheme for Improving Data Persistence in Wireless Sensor Networks”, In: Proceedings of the 2010 IEEE International Conference on Communications (ICC 2010), **(2010)**; Cape Town, South Africa.
[23] N. Cao, Q. Wang, K. Ren and W. J. Lou, “Distributed Storage Coding for Flexible and Efficient Data Dissemination and Retrieval in Wireless Sensor Networks”, In: Proceedings of the 2011 IEEE International Conference on Communications, **(2010)**; Princeton, New Jersey.

[24] S. Scott, R. Sylvia, K. Brad, G. Ramesh and E. Deborah, "Data-Centric storage in sensor nets", ACM SIGCOMM Computer Communication Review, vol.1, no.33, **(2003)**.
[25] S. Ratnasamy, B. Karp, S. Shenker, D. Estrin, R. Govindan, L. Yin and F. Yu, "Data-Centric storage in sensornets with GHT, a geographic hash table. Mobile Network Application, &, vol.8, no.4, (**2003)**, pp.427−442. http://dx.doi.org/10.1023/A:1024591915518.
[26] K. T. Ma, K. Y. Cheng, K. S. Liu and V. Tam, "Improving data centric storage with diffuse caching in wireless sensor networks", Wireless Communications and Mobile Computing, vol.9, **(2009)**.
[27] H. C. Le, H. Guyennet, N. Zerhouni, "Mobile Effect Reduction in Data-Centric Storage for Wireless Sensor Networks", In: Proceedings of the 3rd IET International Conference on Intelligent Environments, **(2007)**; Berlin, Germany.
[28] H. Y. Shen, L. Y. Zhao and Z. Li, "A Distributed Spatial-Temporal Similarity Data Storage Scheme in Wireless Sensor Networks", IEEE Transactions on Mobile Computing, vol.7, no.10, **(2011)**.
[29] H. Y. Shen, T. Li and T. Schweiger, "An Efficient Similarity Searching Scheme Based on Locality Sensitive Hashing", In: Proceedings of Int'l Conf. Database Theory, **(2008**）; Bucharest, Romania.

Author

Weimin Gao (1975---), Male, Hunan Province Qidong country, Associate Professor, School of Computer Science Master, Hunan Institute of Technology, mainly engaged in the research of Computer Network and Information Security. He has received Master degree from Hunan University, and his major is electronic engineering. His research interests include wireless sensor network, Computer Network and Information Security

