

The Analysis and Research of Data Query Processing Technology Based on Wireless Sensor Network System

Yang lin^{1,a}

1Department of Control Engineering, Naval Aeronautical and Astronautical University, Yantai Shandong 264001, China,

^aemail: liudi5388466@163.com,

Key words: wireless sensor; data query; network system; aggregation; correlation model;

Abstract. Wireless sensor network has a broad application prospect in the national defense military, medical services and traffic control, and other fields. As the power energy of the wireless sensor network is limited, therefore the effective use of energy becomes a core problem in the sensor network query processing. The data query processing technology is analyzed and researched based on the key technology of network system in this paper. Spread the query and data collection for a process, thus reducing power consumption and time dela in data query processing.

The introduction

Sensor network has a broad application prospect in the national defense military, medical services and traffic control, and other fields. It becomes a research hotspot in recent years. Different from the traditional network, it has many notable features: sensor node computing power, storage capacity and communication ability is very limited. Node data usually contains noise, perception is uncertain: node movement, node dormancy, communication link failure and the surrounding environmental factors, such as sensor network topology changes frequently. The characteristics of sensor networks have made develop its application very difficult. Considering that the sensor network is a data-centric network, main purpose of using it is to query the data of perception or monitor events. Sensor network data management combines embedded data and distributed database technology to effectively manage the sensor data. It provides users with a simple query interface, shielding the complexity of the query processing, can greatly simplify the development of sensor network applications, has become the important supporting software for the development of sensor network applications. The energy of sensor node is supplied by batteries and cannot be replaced, energy is limited. So efficient query processing technology is the core of the sensor network data management system[1-2].

Query processing based on network gather

The most simple way is flood type for aggregation in sensor networks. The node will perceive the data sent to the neighbor node in the form of radio. After neighbor nodes receive data, and local data gathered computed locally, then the results are sent to neighbor nodes in the form of broadcast. So the gathered results can be calculated finally in the entire network with the method of the flood type. The energy cost of flood type processing method is very big. It will add unnecessary computing of some nodes. In order to overcome this shortcoming, Someone puts forward various grid gathered method on the premise of saving energy. The methods based on the tree: each node calculates its own layer in the process of establishing tree. The period that is assigned to the two layers to transmit data is called the interval. These intervals should be enough to make a pair of nodes successful completion data transmission, not to suspend the transfer by end of the period. A query delay time of the query result is generated by the length of time interval and tree layers. In order to make the tree adapt to the situation of the network, the node and its neighbors need to control each link quality. In order to save energy, the regulation behaviors frequency is less than the frequency of gathering. When a link occur problems, the node will choose a good point to the link

of father node, it can make tree have better robustness. The method based on multipath: this method based on multipath overcomes the shortcoming of low precision based on tree query method. It allows to use any gathered besides tree topology. Multipath gathered don't usually take a long time. This approach is mainly applied to the topology of the ring, because this kind of ring structure provides a good energy efficiency[3-4]. It can obviously increase the robustness, because unless each path to the root node link are all broken. Or at least it can find a link to transmit point information to the root node. Because of a node transmitting data to the upper multiple nodes, so it need to obtain the results by adopting repeat gathered profile method. Multipath structure is shown in figure 1.

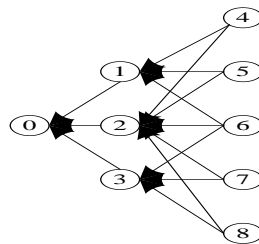


Figure 1 multipath structure figure

The use of tree combined with multipath: this method combines the advantages based on tree method and based on multipath method. As the method based on multipath has good adaptability to network, the loss of link packet will not have any impact on the result. Therefore, the method based on multipath can be used in the case of high packet loss. In addition, from the point of network topology, the method based on tree can be used in the area far away from the root node[5-7]. The method based on multipath can be used in the area near the root node. So different transmission structure can be used to improve the accuracy of the results effectively according to the location of the nodes. Tree and multipath structure is shown in figure 2.

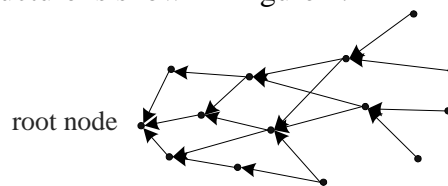


Figure 2 tree and multipath structure figure

data query down technology

Sensor network can be used to monitor the abnormal situation. In this application, the abnormal events usually can be defined as a set of query predicate, such as generate events when the temperature is beyond the normal limits. The original method is: sensors' data transmits to the server along the propagation path. Whether abnormal data is determined in the server. Due to abnormal events, this kind of query selectivity is very low. The query processing is done after a large amount of data is transmitted to the server. It will lose a lot of energy, reduce the life of network. All the query predicate are saved in query predicate list to solve the problem. The table with the query is pushed down to each node in network. Node perception data connects the predicate table. There are abnormal event occurs when the query result is not null. Then the abnormal data transmits back to the server. We can save energy and reduce the financial burden of the network. If there is no abnormal events but node is failure. or data errors because of some reason. Corresponding processing method is: a list of recording nodes to produce data is established according to the historical data. Periodically updating this list, remove those range that don't often appear. As the memory of sensor node is very limited, when data items in static table is too much, it will cause the problem of insufficient memory[8-9].

The spatial correlation model based on sensory data

The perception data of sensor nodes just shows the discrete information on some of the sample

point. Because it is approximate result, then receiving data little contribution to the results will waste a lot of time and energy, so a method based on the model can be applied. Model can be used to analyze sensor for better. For example, they can help to solve the unbalanced problem of sampling values in space. It can help to confirm the wrong information of nodes, and sampling data of necrosis node can be inferred. In addition, the model provides a framework. It can optimize the data reception of sensor. Only when the current model will not be able to deduce the accuracy of data, sensors to collect data can be allowed. Given a model, the data of node can be used to estimate the data of other node. For example, the temperature sample of node is likely to improve estimation reliability based on nodes of the model. In addition, in terms of communication costs, receiving data of node from close range saves energy more than node from far range. User query request, the query will be translated into probability operation based on the model. Error tolerance and the target confidence shows the requirement of user for approximation. If the credibility of model is lower, the model needs to obtain sampling values that can help to improve the precision from sensor network before answering check.

complex query processing

Query for wireless sensor network is to point to find k objects in n sensor nodes. K is usually from the parameters of the application specified by the user. The complex queries are done has very important practical application meaning in the wireless sensor network, such as wild biological monitoring, fire early warning and prevention control etc. Typically, wireless sensor network can produce a large number of real-time data. At the same time, as the nodes use batteries, energy is very limited, communication is in a wireless way between nodes and more energy is consumed. These determine that complex queries can be finished by using little communication cost in the wireless sensor network. So as to save the node energy maximize, the network life is prolonged. From the accuracy of the query results, complex queries are divided into precise query and approximate query in wireless sensor network. In terms of precise query, the most basic is the centralized algorithm. All data are collected, and then the results are calculated. The implement of centralized algorithm is the most simple, but no any communication cost is saved, undesirable in practical application. Naive algorithm can be thought as the improvement of the centralized algorithm, the whole sensor network is taken as tree structure, the intermediate nodes collect the value of other nodes and merge the value that they gain. Calculating k before transferred upwards, naive algorithm abandons some value that could not enter the value of final result in the process of data transmission, so saving a part of the transfer price. Both the centralized algorithm and naive algorithm, each node sends data at least once, so they are not effective algorithm of save node energy [10-12]. Its basic idea is to set a filter for each node, it is actually the node's value range. In the query request, all nodes are ordered according to interval size of filter. When a node receives a query request, first of all, it collects a new data. Whether to submit up is determined according to the filtering rules. If the new value belongs to filter interval, it does not need to submit. Otherwise, the data is abnormal, the new data need to be send back to the query port. After the query port receives all the abnormal data, the nodes need to be ordered again. The k nodes that are chosen from high to low are taken as the final query results. The useless transmission is avoided from the sensor nodes. So the energy consumption of nodes can be saved effectively.

Conclusion

Wireless sensor network is a new kind of technology and has a high application value in industrial and agricultural production, environmental protection, medical and health care, aerospace, military exercise. The data query processing technology need to be used to reduce the power consumption of nodes in the process of sensor network obtaining information. Therefore, the effective use of energy becomes a core problem in the sensor network query processing. The method of wireless sensor network system data management is proposed based on the key technology of network system in this paper. The power consumption and time delay are reduced in

the processing of data query.

Reference

- [1] Van Dam and Koen Langendoen. An adaptive energy-efficient mac protocol for wireless sensor networks. In Proceedings of the 1st international conference on Embedded networked sensor systems, SenSys'03, pages 171-180, New York, NY, USA, 2003. ACM.
- [2] Kyle Jamieson, Hari Balakrishnan, and Yocy Tay. SiR: A MAC Protocol for Event-Driven Wireless Sensor Networks, volume 3868 of Lecture Notes in Computer Science, pages 260-275.
- [3] Vern. Rajendran, Garcia Luna Aceves. Energy-efficient channel access scheduling for power-constrained Networks. In Wireless Personal Multimedia Communications, 2002. The 5th International Symposium on, volume 2, pages 509-513 v01. 2, oct. 2002.
- [4] Venkatesh Rajendran, Katia Obraczka, and Garcia Luna Aceves. Energy-efficient collision-free medium access control for wireless sensor networks. In Proceedings of the 1st international conference on Embedded networked sensor systems, SenSys'03, pages 181-192, New York, 2003. ACM.
- [5] Gang Lu, Bhaskar Krishnamahafi, and Cauligi Raghavendra. An adaptive energy-efficient and low-latency mac for tree-based data gathering in sensor networks: Research articles. *Wirel. Commun. Mob. Comput.*, 7(7):863-875, September 2007.
- [6] Linjong Rhee, Ajit Warrier, Mahesh V. S. and Jeongki Min. Zmac: a hybrid mac for wireless sensor networks. In Proceedings of the 3rd international conference on Embedded networked sensor systems. SenSys'05, pages 90-101, New York, NY, USA. 2005. ACM.
- [7] Martin Kubisch, Holger Karl, Adam Wolisz, Lizhi Charlie Zhong and Jan M. Rabaey. Distributed algorithms for transmission power control in wireless sensor networks. In *Wireless Communications and Networking, 2003. WCNC 2003. 2003 IEEE*, volume 1, pages 558-563 v01.1, march 2003.
- [8] Li Li, Joseph Y. Halpern, Paramvir Bahl, Yi-Min Wang and Roger Wattenhofer. A cone-based distributed topology-control algorithm for wireless multi-hop networks. *IEEE/ACM Transactions on Networking*, 13(1):147—159, feb. 2005.
- [9] Na Li and Jian Hou. Topology control in heterogeneous wireless networks: problems and solutions. In *INFOCOM 2004. Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies*, volume 1, pages 4 v01. march 2004.
- [10] Santosh Kumar, Ten H. Lai and József Balogh. On k-coverage in a mostly sleeping sensor network. In Proceedings of the 10th annual international conference on Mobile computing and networking, *MobiCom'04*, pages 152—158, New York, NY, USA, 2004. ACM.
- [11] Kui Wu, Yong Gao, Fulu El, and Yang Xiao. Lightweight deployment-aware scheduling for wireless sensor networks. *Mob. Netw. Appl.*, 10(6): 837-952, December 2005.
- [12] Alberto Cerpa and Deborah Estrin. Ascent: Adaptive self-configuring sensor networks topologies. *IEEE Transactions on Mobile Computing*, 3(3): 272-285, July 2004.