JCSE International Journal of Computer Science and Engineering Open Access
Review Paper Volume-2, Issue-1 E-ISSN: 2347-2693

Analysis of Uniform Distribution of Storage Nodes in Wireless Sensor Network

Madhavi S. Kukade[1*], Prof. Kapil N. Hande[2]

[1*]*PG Student, Department of CSE, PBCE Nagpur, India, sakshikhamankar@gmail.com*
[2]*Department of CSE, PBCE Nagpur, India, kapilhande@gmail.com*

Received: 12/01/2014 Revised: 22/01/2014 Accepted: 28/01/2014 Published: 31/01/2014

***Abstract*—**Sensors networks are capable of collecting an enormous amount of data over space and time .Often ,the ultimate objective is to "sample, store and forward " that is to sense the data , store it locally and ultimately forward it to accent almost and analyzed. Typical sensor nodes are wireless nodes with limited storage and computational power . Furthermore they are prone to "failure" by going out of wireless range, interference running out of battery etc. The sensor and storage nodes are distributed randomly in some region and can not maintain routing tables or shared knowledge of network topology .Some nodes might disappear from the network due to failure or battery depletion overall this problem has occurred for overcoming this problem so many techniques are studied in this paper.

***Index Term*—**Distributed data collection algorithm ,Storage nodes ,Sensor nodes ,WSNs

I. Introduction

Wireless sensor networks (WSNs) often consist of small devices (nodes) with limited processing ability, bandwidth and power. They can be deployed in isolated or dangerous are to monitor objects, temperatures, etc. or to detect fires, floods, or other incidents. There has been extensive research on sensor networks to improve the inutility and efficiency. distributed data collection algorithm to accurately store and forward information obtained by wireless sensor networks is proposed. The proposed algorithm does not depend on the sensor network topology where there is arrangement of a network ,including its nodes and connecting lines, routing tables which contains the information necessary to forward a packet along the best path toward its destination .[16] Each packet contains information about its origin and destination .When a packet is received, a network device examines the packet and matches it to the routing table entry providing the best match for its destination, or a geographic location of sensor nodes which takes location information of nodes, are very valuable for sensor networks, but rather make sure of uniformly distributed storage nodes.[17] Analytical and simulation results for this algorithm shows that, with high probability, the data disseminated by the sensor nodes can be precisely collected by querying any small set of storage nodes.

In a large scale wireless sensor network with a set of sensing nodes and a set of storage nodes. The sensing nodes have limited memory and bandwidth, and they might disappear from the network at any time due to limited battery lifetime.[10][11] The storage nodes have large memory and bandwidth, but they do not sense information about the region. We assume that the data collector (base station) is far away from the nodes as shown in Fig. 1,[17] but it is connected with a set of storage nodes. The sensor nodes are able to sense data and distribute it to the storage nodes.

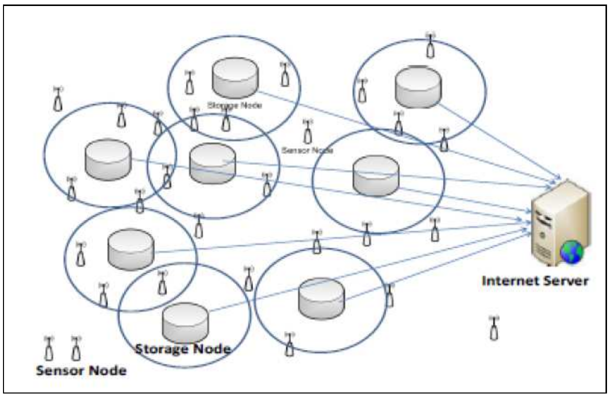

Figure1. Network model representing a Wireless Sensor with sensing and storing nodes , and base station nodes.

Sensor networks are especially useful in catastrophic or emergency scenarios such as floods ,fires ,terrorist attacks or earth quakes where human participation may be too dangerous .However ,such disaster scenarios pose an interesting design challenge since the sensor nodes used to collect and communicate data may themselves fail suddenly and unpredictably ,resulting in the loss of [8]valuable data .Furthermore ,because these networks are of ten expected to be deployed in response to a disaster ,or because of sudden configuration changes due to failure ,these networks are often expected to operate in a "zero-configuration "paradigm ,where data collection and

Corresponding Author: *Madhavi S. Kukade*

transmission must be initiated immediately, before the nodes have a chance to assess the current network topology.[8]

Sensors networks are capable of collecting an enormous amount of data over space and time .Often, the ultimate objective is to “sample, store and forward”, that is to sense the data ,store it locally and ultimately forward it to a central host(or “masternode”) where data from other sensor nodes is also collected and analyzed .A useful example is a traffic sensing network ,there being traffic sensors at each intersection that estimate the traffic and relay it to a central processing station. Typical sensor node are wireless nodes with limited storage and computational power .In a cooperative sensor network ,it is a good idea to have nodes’ data duplicated and spread around the networks it can be recovered from other nodes in case of failure .In particular ,every node can store some of its own data as well as data from other nodes up to its storage capacity.

II. Data Collection and Recovery Concepts

WSNs are deployed to monitor various aspects of the environment ,such as temperature and light. The WSNs are also being deployed in a wide variety of other applications . For WSN applications ,the data in the sensors are often streamed to a control centre (called sink).This process is called data collection.[1] Generally ,data gathering can be categorized as data collection which gathers all the data from a network without any data aggregation or merging and data aggregation, which obtains some aggregation values, e.g. MAX ,MIN ,SUM ,and etc .For data collection capacity ,it is defined as the average data receiving rate at the sink ,[18][19] i.e. data collection capacity reflect show fast data been collected by the sink .We use data collection capacity and network capacity interchangeably. The benefits of storing combination of data instead of original data.

Traditional error-correcting erasure codes can also be used to achieve the goal of encoding data such that if some of the encoding symbols are lost, data can still be recovered. Reed -Solomon codes are block erasure codes that have been traditionally used for error correction.[19] If part of the sensor network fails ,the data stored in the failed sensor nodes is lost .The only data which remain is the data remaining at the surviving nodes. The surviving nodes have symbols encoded from the data produced by all the sensor nodes, so there is still a chance of recovering data produced by failed nodes using the surviving encoding symbols.The failure of nodes in a region translates to a loss of memory in the global network, and any information stored in these failed regions is lost unless it is duplicated elsewhere or is already delivered to the sink. Data that is retrieved via the information collected at the sink is referred to as recovered data as shown in figure 2 also take advantage of the location information of nodes ,are very valuable for sensor networks for data collection.

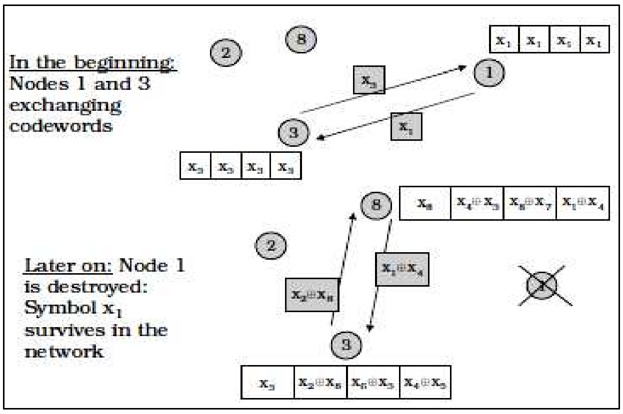

Figure 2.Sensor Networks that Need Persistence

Geocasting is the delivery of packets to nodes within a certain geographic area. Perhaps the simplest way for geocasting is global flooding. In global flooding, the sender broadcasts the packet to its neighbors, and each neighbor that has not received the packet before broadcasts it to its neighbor, and so on, until the packet is received by all reachable nodes including the geocast region nodes. It is simple method of data collection[17]as shown in figure 3.

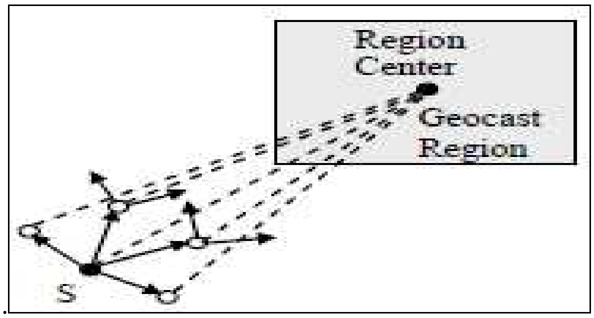

Figure 3: Progressively Closer Nodes (PCN): Closer nodes to the region than the forwarding node forward the packet further and other nodes discard it

There is a distributed storage algorithms for large-scale wireless sensor networks. Assume a wireless sensor network with n nodes that have limited power, memory, and bandwidth. Each node is capable of both sensing and storing data. Such sensor nodes might disappear from the network due to failures or battery depletion. Hence it is desired to design efficient schemes to collect data from these n nodes.

There are two distributed storage algorithms (DSA's) that utilize network flooding to solve this problem. In the first algorithm, DSA-I, we assume that the total number of sensors is known to each sensor in the network. We show that this algorithm is efficient in terms of the encoding and decoding operations. Furthermore, every node utilizes network flooding to disseminate its data throughout the network using a mixing time of approximately O(n). In the second algorithm, DSA-II, we assume that the total number of nodes is not known to every sensor; hence dissemination of the data does not depend on n. The encoding operations in this case take O(C mu^2), where mu is the mean degree of the network graph and C is a system parameter. We evaluate the performance of the proposed algorithms through analysis and simulation.[21]

III. Data Collection Techniques In WSNs

Previous work on this problem has focused on situations in which either the network topology is known or the sensor nodes are able to maintain routing tables. Previously they used a decentralized implementation of fountain codes that uses geographic routing and every node has to know its location. The motivation for using fountain codes instead of using random linear codes is that the former requires O(k log k) decoding complexity also they analyzed techniques to increase ”persistence” of sensed data in a random wireless sensor network. where they proposed two decentralized algorithms using fountain codes to guarantee the persistence and reliability of cached data on unreliable sensors & random walks to disseminate data from a sensor (source) node to a set of other storage nodes. first algorithm introduces lower overhead than naive random-walk, while the second algorithm has lower level of fault tolerance than the original centralized fountain code.

Proposing Ravine Streams to preserve data stream persistence in disruptive sensor networks. Source data is initiated and delivered in its encode form by network erasure codes On Code [20]. For data preservation, receiving nodes distributive make the acceptance decision based on local failure probability and storage space, here nodal failure probability has taken the residual energy into account and dynamically update across the process. Adaptive transmission power control in RS ensures that data acceptance probability by neighbour nodes is expected at 1 for each broadcasting under minimum transmission power. With data packet recoding, the dispensable data content redundancy is constrained for more energy efficiency.

Novel technique called growth codes to increase data persistence for data collection in wireless sensor networks, i.e. increasing the amount of information that can be recover at the sink. Growth codes are a linear technique in which information is encoded in an online distributed way with increasing degree, as shown in figure 4 where showed that growth codes can increase the amount of information that can be recovered at any storage node at any time period. These codes are also easily implemented in a distributed fashion - another important criterion for sensor networks. The code grows with time: initially codeword's are just the symbols themselves, but over time, the codeword become linear combinations of (randomly selected) growing groupings of data units. A well-designed code will grow at a rate such that the size of the codeword received by the sink is that which is most likely to be successfully decoded and deliver previously unrecovered data.[8]

Growth Codes:-

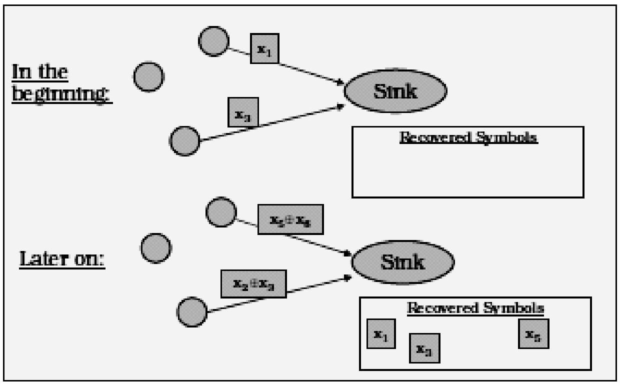

Figure4.Growth Codes in action

Also work on geographic routing where the destination is a geographic region instead of a specific node or point and recasting is a communication primitive in wireless sensor networks used, since in many applications the target is to reach nodes in a certain region. In geographic-based rendezvous mechanisms, geographical locations are used as a rendezvous place for providers and seekers of information. Geographic-based rendezvous mechanisms can be used as an efficient means for service location and resource discovery, in addition to data dissemination and access in sensor networks. A main component in geographic routing is greedy forwarding showing in figure 5, in which the packet should make a progress at each step along the path. Each node forwards the packet to a neighbor closer to the destination than itself until ultimately the packet reaches the destination. If nodes have consistent location information, greedy forwarding is guaranteed to be loop-free. [17]

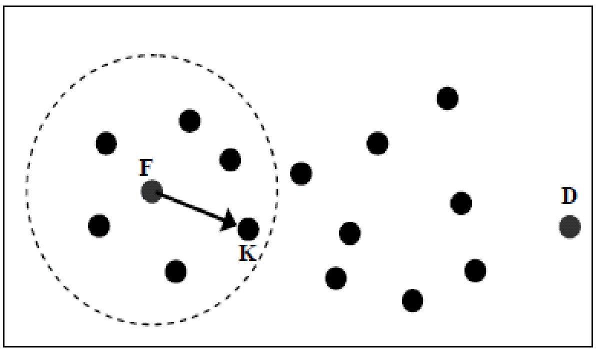

Fig 5: Greedy forwarding: Node F forwards the packet to neighbor K, which is the neighbor closest to the destination D

There is a scalable and order-optimal distributed algorithm, named Distributed Data Collection (DDC),[3]with fairness consideration and capacity analysis under the generalized physical interference model. This is the first attempt to provide detailed protocol design and rigorous capacity analysis for data collection in distributed asynchronous WSNs. DDC works in a CSMA-like manner, except for the RTS/CTS communication mode and the necessity to reply an ACK packet after receiving a data packet. In DDC, when a node has some data packets for transmission, it sets up a back off timer, and senses the wireless channel with a predefined Carrier-sensing Range (CR). If the channel is free when the back off timer expires, this node conducts a data transmission. Under this transmission manner, DDC gathers all the of data of the network to the sink (*i.e.* base station).

Working on Reed-Solomon codes and low-density parity-check (LDPC) codes which have been widely used for data persistence in distributed data storage. Usually, source data are encoded in a centralized location and then the encoded symbols are distributed to different machines for storage. However, in a wireless sensor network, data sources are distributed. It is not practical to designate an energy-constrained sensor node to perform centralized encoding, not to mention about the communication cost to transmit all sensor readings to the rendezvous point. In order to achieve data persistence in a wireless sensor network, it is desired that encoding of the erasure code can be implemented in a distributed way.[18]

Developing two distributed algorithms for solving this problem based on simple random walks and Fountain codes. Unlike all previously developed schemes, this solution is truly distributed, that is, nodes do not know n, k or connectivity in the network, except in their own neighborhoods, and they do not maintain any routing tables. In the first algorithm, all the sensors have the knowledge of n and k. In the second algorithm, each sensor estimates these parameters through the random walk dissemination. Here present analysis of the communication/transmission and encoding/decoding complexity of these two algorithms, and provide extensive simulation results as well.[22]

IV. Proposed Work

Distributed data collection algorithm

From the above things considering all issues related with WSNs proposing one algorithm to uniformly distribute the information.

Consider a wireless sensor network N with n nodes among which S={s1,...,s_n} are sensing node and R ={r1,...,r_n−k} are storage nodes in network simulator model (NS2).The sensor and storage nodes are distributed randomly in some region R and can not maintain routing tables or shared knowledge of network topology .Some nodes might disappear from the network due to failure or battery depletion.

In this paper ,we will propose a distributed data collection algorithm for over coming the storage problem by using NS2 simulator which help to create the desired network model. The clustering storage algorithm runs in the following phases:

i) Clustering phase: In the clustering phase ,each storage node sends a flooding beacon message with its ID to all neighboring nodes in the network.

ii) Sensing phase: In the sensing phase, the sensor nodes sense data from the environment. Once the data is collected, they send their packets to the storage node, from which they have received beacon packets.

iii) Data collection and storage phase: When a sensing node senses the environment, it sends its packets to its storage nodes. The storage nodes collect the incoming packets and store them encoded in their own buffer. Based on the type of the incoming packets, the storage nodes will store these packets or update the existing data in their buffers.

iv) Querying phase: The query process can be performed by the base station or server that collects all data from the storage nodes. Total number of nodes that must be queried in order to obtain the data sensed by the sensor nodes.

V. Conclusion

In this paper a survey on various existing techniques for data collection algorithm in wireless sensor network studied, where a distributed storage problem in large-scale Wireless Sensor Networks occurred .From the techniques like growth codes, random walk ,greedy forwarding are not that much effective for data storage issues. So to overcome this problem future work involve to propose a data collection algorithm and efficient query-selection algorithm by carefully selecting a subset of queries to precisely collect sensed data and successfully store it at storage nodes and forward the data to next node uniformly.

VI. References

[1]. Xiaohua Xu, Xiang-Yang Li, Min Song “Real-time Data Collection Scheduling in Multi-hop Wireless Sensor Networks,” IEEE Global Communications Conference 2013

[2]. Mingsen Xu, Wen-Zhan Song, “Ravine Streams:Persistent Data Streams in Disruptive Sensor Networks” IEEE SECON (2013) Key: citeulike:12237835

[3]. Shouling Ji, Zhipeng Cai, “Distributed Data Collection and Its Capacity in Asynchronous Wireless Sensor Networks,” IEEE INFOCAM 2012

[4]. X. Hou, Z. Sumpter, L. Burson, X. Xue, and B. Tang, “Maximizing Data Preservation in Intermittently Connected Sensor Networks,” in IEEE International Conference on Mobile Ad hoc and Sensor Systems, Oct. 2012.

[5]. M. Valero, M. Xu, N. A. Mancuso, W.-Z. Song, and R. Beyah, “EDR2: A Sink Failure Resilient Approach for WSNs,” in IEEE International Conference on Communications, Jun. 2012.

[6]. A. G. Dimakis, V. Prabhakaran, and K. Ramchandran. "Decentralized erasure codes for distributed networked storage". IEEE/ACM Transactions on Networking, 14(SI):2809 – 2816, June 2006.

[7]. A. G. Dimakis, V. Prabhakaran, and K. Ramchandran. "Distributed fountain codes for networked storage". In Proc. IEEE International Conference on Acoustics, Speech and Signal Processing, Toulouse, France, May 14-19, 2006.

[8]. A. Kamra, V. Misra, J. Feldman, and D. Rubenstein." Growth codes: Maximizing sensor network data persistence". In Proc. 2006 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, pp 255-266, Pisa, Italy, 2006.

[9]. Z. Kong, S. A. Aly, and E. Soljanin." Decentralized coding algorithms for distributed storage in wireless sensor networks". IEEE Journal on Selected Areas in Communications, 28(2):261– 267, Feburary 2010.

[10]. Y. Lin, B. Liang, and B. Li."Data persistence in large-scale sensor networks with decentralized fountain codes". In Proc. Twenty-Sixth Annual Joint Conference of the IEEE Computer and Communication Societies, Anchorage, AK, USA, May 6-12, 2007.

[11]. M. Goyeneche, J. Villadangos, J.J. Astrain, M. Prieto, and A. Cordoba. "A distributed data gathering algorithm for wireless sensor networks with uniform architecture". In Proc. 15th Euromicro International Conference on Parallel Distributed and Network-Based Processing (PDP’07), pages 373–380. IEEE Computer Society, Los Alamitos, CA, USA, 2007.

[12]. O. Younis and S. Fahmy." Distributed clustering in ad hoc sensor networks: A hybrid, energy-efficient approach". In Proc. Twenty-Third Annual Joint Conference of the IEEE Computer and Communication Societies, volume 1, pages 629–640, Hong Kong, China, March 7-11, 2004

[13]. Y. Lin, B. Li, and B. Liang. "Differentiated data persistence with priority random linear code". In Proc. 27th International Conference on Distributed Computing Systems (ICDCS’07), Toronto, Canada, June, 2007.

[14]. U. Niesen, P. Gupta, and D. Shah, "On Capacity Scaling in Arbitrary Wireless Networks" IEEE Transactions on Information Theory, 55(9): 3959-3982, 2009.

[15]. G. Sharma, R. Mazumdar, and N. B. Shroff, "Delay and Capacity Trade- Offs in Mobile Ad Hoc Networks: A Global Perspective", IEEE/ACM Transactions on Networking, 15(5): 981-992, 2007.

[16]. Salah A. Aly, Ahmed Ali-Eldin, H. Vincent Poor," A Distributed Data Collection Algorithm for Wireless Sensor Networks with Persistent Storage Nodes " 4th IFIP International Conference on New Technologies, Mobility and Security(NTMS) 7-10 February 2011 Paris, France

[17]. Karim Seada, Ahmed Helmy," Geographic Protocols in Sensor Networks " The 3rd ACS/IEEE International Conference on Computer Systems and Applications, 2005.

[18]. Mu Lin, Chong Luo, Feng Liu and Feng Wu.″ Compressive Data Persistence in Large-Scale Wireless Sensor Networks″ IEEE Global Telecommunications Conference (GLOBECOM 2010)

[19]. Shouling Ji,Zhipeng Cai.”Distributed Data Collection and Its Capacity in Asynchronous Wireless Sensor Networks” IEEE INFOCOM 2012

[20]. M. Xu, W.-Z. Song, and D. Heo, “OnCode: Opportunistic In-Network Coding and Delivery in Energy-Synchronized Sensor Networks,” IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops), 2013

[21]. S. A. Aly, H. Darwish, M. Youssef, and M. Zidan. "Distributed flooding based storage algorithms for large-scale wireless sensor networks.".IEEE International Conference on Communications, Dresden, Germany, June 13-17, 2009.

[22]. S. A. Aly, Z. Kong, and E. Soljanin. Fountain codes based distributed storage algorithms for wireless sensor networks. IEEE/ACM Information Processing of Sensor Networks (IPSN’08), pages 171–182, St. Louis, MO, USA, April 19-22, 2008.

AUTHORS PROFILE

Author is a student of master of engineering in computer science and engineering of Nagpur University. she had completed graduation in bachelor of engineering in computer technology from MIET,Gondia Maharashtra, INDIA in 2004

