Asian Journal of Information Technology 15 (13): 2238-2243, 2016
ISSN: 1682-3915

Optimization Methods for Data Storage Position in Wireless Sensor Networks: A Survey

[1]P. Nitha, [2]P.S. Periasamy and [3]Mohanosundaram
[1]Department of MCA, [2]Department of Electronics and Communication Engineering, KSR College of Engineering, Tiruchengode, Tamil Nadu, India
[2]School of Computing Science and Engineering, VIT University, Vellore, Tamil Nadu, India

Abstract: Data storage has been becomes a significant issue in Wireless Sensor Networks (WSNs) since the information collected from nodes should to be used for future information recovery. For this purpose storage nodes are introduced to collect the information regarding sensors in WSNs. But the maintenance of storage nodes to obtain less energy consumption and reduced communication cost stimulated by the network query along with heavy load of transmitting each and every one data to a central place is not an easy task. To alleviate the problem of storage node, several data storage schemas are introduced in recent work. In these schemas using storage nodes in WSNs but here how to select appropriate positions for storage nodes becomes a major important issue. To alleviate the problem of storage node positions various heuristic algorithm and Meta heuristic algorithms is proposed in recent works to find the exact positions for k storage nodes while minimization of total energy cost consumption. In this survey study we study the detailed review of various heuristic algorithms and Meta heuristic algorithms. An issue of those methods in WSNs for storage node position problem is also studied in detail. Then implemented these algorithms in a Wireless Sensor Network simulator (NS2) and some experimentation is carryout. The simulation results demonstrate that the feasibility and efficiency results of various heuristic algorithms.

Key words: Wireless Sensor Networks (WSN), data storage node placement problem, total energy cost minimization, Optimal Data Storage (ODS), data rate, eographical location

INTRODUCTION

Wireless Sensor Network (WSN) have been becomes an emerging typical ad hoc network which consists of many sensors (Tang and Gupta, 2007). Now a days WSNs applications have been used in various applications such as pervasive computing, atmosphere monitoring, health care monitoring, transportation business, military, etc (Wang *et al*., 2009). This entire application huge amount of information is collected from sensors and further used for information retrieval which is considered as one of the significant issues in WSNs (Huang *et al*., 2010).

Since, data storage in WSNs (Sarkar *et al*., 2009) occupies producers for storing sensor nodes information and consumers receives information. To perform these task data storage strategy is introduced along with the storage positions of producers and the condition of responses to consumers' queries, however it becomes difficult in WSNs (Sheng *et al*., 2010). Since, the worst data storage strategy determination increase data transmission, waste of energy which results less network lifetime for all WSNs. So, finding the position of storage nodes in optimal data storage strategy should be carefully considered as one of the major important factor for reducing the communication cost and energy consumption.

There are two major important factors that impact the results of communication cost and energy cost consumption that are data rates of producers, consumers and their corresponding path distances to the storage node. The data rate of producers is represented as the producing data rate from producers. The data rate of consumers is represented as the querying data rate from consumers. The data rate frequently does not alter in a predetermined application-specific time interval. In real WSNs, the closer the storage node of producers and consumers is decided based on the path distance (hop distance).

An efficient way to perform this task is by the positioning of the nodes in WSNs with reduced communication cost and energy cost consumption. Though these two factors have been considered as important factor and have been investigated in the recent work by addressing the data storage node problem via the

Correspomding Author: P. Anitha, Department of MCA Engineering, KSR College of Engineering, Tiruchengode, Tamil Nadu, India

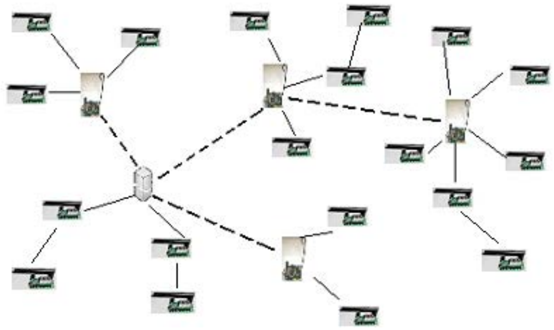

Fig. 1: A two-tier data storage strategy with four storage

representation of WSN as a tree structure, here the base station (Madden *et al*., 2005a, b) is considered storage node and receiver as consumer.

In the tree structure, the data rates of producers and the query rate of consumer is known. From this data rates the communication cost and energy cost consumption of WSNs have been minimized by optimizing data storage placement. Another one of the schema is using mesh network topology for data storage schema. However in mesh network topology consists of several numbers of producers and consumers each and every one seeking to develop one event concurrently. But it becomes very difficult task because it is performed based on the fixed infrastructure. To meet this challenge many data storage strategies have been introduced in recent works to enhance network lifetime of WSNs. One of the schema is two-tier data storage strategy as illustrated in Fig. 1. It consists of some sensor node, storage node, with higher storage capability, high battery power and approximating the sink node to store information collected through the sensors in its closeness (Wang *et al*., 2009; Sarkar *et al*., 2009).

In this two-tier data storage strategy, how to select suitable locations for number of storage nodes to formulate energy efficient and extending the network lifetime is becomes a critical issue in WSN. Some prior works have been proposed to solve this issue however, many of the methods follows topology assumptions with fixed communication tree. This type of infrastructure consumes higher energy for different sensor nodes. In this study, we study the review of various Optimal Data Storage (ODS) methods in WSNs with the intention of permit the storage location to dynamically change in response to together the geographical locations of producers, consumers and the data rates at which data are being exchanged. In addition we also study the issues and advantages of storage node optimization methods, how it minimizes the energy consumption of the total network is also simulated using the network simulator NS2. In the simulation results ODS strategy find out the total communication cost in the one-to-one and many-to-many models.

Literature review: WSN is a set of hundreds or thousands of micro sensor nodes that have capabilities of sensing, establishing wireless communication between each other and doing computational and processing operations (Khamforoosh and Khamforoush, 2009). One of the biggest challenges in these applications is how to store and search the collected data. Since, the sensors operate on battery power it is very important to make efficient use of energy of sensors to increase the lifetime of the network. Thus the two main factors are deciding the performance and lifetime of the WSN, namely energy efficiency and data storage position. This section is divided into three sections to deal with the storage node placement problem. In the first section, the data storage problems in the conventional approaches are discussed. In the second section, the deterministic placement of storage nodes using optimal algorithms based on heuristic methods with their advantages and limitations is discussed. Finally, the application of Meta heuristic methods like GA, PSO and BAT in WSN is analyzed.

MATERIALS AND METHODS

Survey on data storage strategy in WSN: Many researchers have proposed various techniques to store data in WSN. This section describes various data storage schemes in Wireless Sensor Networks with their advantages and limitations.

Load Balanced Data-Centric Storage (LB-DCS) (Albano *et al*., 2011) is an organic approach that relies on the home perimeter for data replication and thereby overcomes the unbalanced load constraint in DCS-GHT. LB-DCS functions on top of three mechanisms:

- A density estimation protocol that is used to estimate the network density f which is included in put and get protocols
- A modified hashing function that includes f in its parameter list
- A storage protocol enforcing QoS in the selection of the number of replicas for data storage

Shen *et al*. (2011) proposed a distributed spatial-temporal similarity data storage scheme where data are stored according to their spatial and temporal similarity in order to reduce the overhead as well as the latency of a query request. Even if data-centric storage approaches are based on node cooperation, data are not fully distributed since specific nodes store all the contents generated by the others. A detailed survey on data-centric storage schemes is presented by Ahmed and Gregory (2012).

Liao and Yang (2012) tried to enhance the DCS (Data-Centric Storage) with new techniques in order to improve the energy consumption model and data-gathering, The DCS strategy enables distributed data storage within the network while allowing direct data query without message flooding. Therefore, it is more efficient than the above two alternatives in many scenarios. The DCS scheme has consequently become a popular choice for many sensor network applications and various DCS systems have been developed.

Maraiya *et al.* (2011) proposed Directed Diffusion (DD), a popular information aggregation paradigm for wireless device networks. It's a data-centric and application aware paradigm; within the sense every information generated by sensor nodes is called by attribute-value pairs. Such a scheme combines the information coming back from totally different sources en-route to the sink by eliminating redundancy and minimizing the amount of transmissions. Thus, suggested it saves the energy consumption and will increase the network lifespan of WSNs. In this theme, usually the base station broadcasts the message to the interested supply node.

Yu *et al.* (2010) proposed ODS/NDS (Optiaml Data Storage/Near Optimal Data Storage) that saves energy by storing event data at a node that will be close to the producers and consumers of the event data, minimizing communication costs. However, it does not solve the problem of querying for event data which will still require flooding of the network in order to find the storage node. A possible way to solve this problem would be to use the method for finding a storage node as the method for querying as well. This will greatly reduce the number of transmissions that it would take to find the storage node, compared to flooding the entire network.

Conventional data storage techniques discussed in this section exhibit the gaps like limitations in convergence speed, network life time, unbalanced load distribution, hot spot problem and scalability. To address these gaps, the literature survey has been extended with Heuristics algorithms in the forthcoming section.

Review of heuristic algorithms for storage node placement problem: Rebai *et al.* (2014) aimed to cover a grid fully by deploying the necessary wireless sensors while maintaining connectivity between the deployed sensors and a base station (the sink). The problem is NPComplete as it can be reduced to a 2-dimensional critical coverage problem which is an NP-Complete problem. A branch and bound (B and B) algorithm is developed to solve the problem optimally. The computational experiments are done to verify if the proposed B and B algorithm is more efficient, in terms of computation time, than the integer linear programming model developed by Rebai *et al.* (2013), for the same problem. Kabiri and Vahidi (2014) proposed an algorithm by combining Mix-Fuzzy approach and A-star algorithm. This method is capable of selecting the optimal routing path from the source node to the sink by favoring the highest remaining energy, minimum number of hops, the lowest traffic load and the lowest energy consumption rate.

Review of meta heuristic methods for storage node placement problem: Kaur and Luthra (2014) proposed an efficient method based on Genetic Algorithms (GAs) to solve a sensor network optimization problem. Long communication distances between sensors and a sink (or destination) in a sensor network can greatly drain the energy of sensors and reduce the lifetime of a network. GA can be used to minimize the total communication distance, thus prolonging the network lifetime. From the above discussions it is seen that, till now, GA is not utilized for finding optimal storage node position. With this motivation, GA based storage node placement algorithm for k storage nodes is encouraged and that can reduce the energy cost.

Li and Wen (2015) proposed a distributed localization algorithm based on PSO which can also extend to large-scale sensor networks. The main contributions of this work have the following two aspects. It reduces the initial search space by bounding box method to speed up the convergence. It proposes a distributed two-phase PSO algorithm to solve the flip ambiguity problem and localize more target nodes. Moreover, the unknown nodes which only have two references or three near-collinear references are tried to be localized in this research.

Jagtap and Shukla (2013) proposed an optimal deployment scheme called EPSO. In EPSO, the mapping between the sensor nodes and the Coordinator is optimized in order to maximize coverage, connectivity and less energy consumption. EPSO has a potential to support sensor networks with low and high density and with the coordinator. A Fitness function is presented to solve the general sensor node deployment optimization problem. To reduce the computational complexity, Node deployment is very important in view of coverage and connectivity to implement EPSO scheme. This study does not consider the storage node placement problem aiming to place limited storage nodes in the sensor network to minimize the total energy cost for collecting the raw data and replying queries at the storage nodes. So, PSO based optimal placing of storage nodes in sensor network will be proposed.

Goyal and Patterh (2013) proposed the application of BAT algorithm for distributed iterative node localization in WSNs. BAT algorithm performs quite well in terms of nodes localized and localization accuracy. The goal of localization is to assign geographical coordinates to each device with unknown position in the deployment area. In recent times, the popular strategy is to apply optimization algorithms to solve the localization problem. In this research, bat algorithm is implemented to estimate the sensor's position.

According to Zhanbei (2014) node localization is a key technology of wireless sensor network and least square algorithm has low localization precision. In order to improve the localization precision of wireless sensor network node, a novel node localization method of wireless sensor network is proposed, based on bat algorithm. Firstly, the node localization problem is transformed into a constrained optimization problem and then the problem is solved by bat algorithm. The results show that the proposed method has improved the localization precision and localization efficiency of wireless sensor network node compared with other node localization methods. From the above survey it is well known that the Bat algorithm is used for node placements and not utilized for the storage node placement problem.

RESULTS AND DISCUSSION

The simulation results is considered for the one-to-one model to decide the optimal storage node k. Here consider there is only one producer i and one consumer j to deal with information in the WSN. The optimal storage cost location must be attained with minimum $C_{total}(k)$. The network simulation parameters considered for this survey are provided in Table 1. In order to assess the performance of the proposed ODS algorithm, a WSNs is implemented in the NS2 simulator to execute some experiments for the data storage position. In the NS2 simulator, the sensor node and the storage node are randomly deployed in a 400×400 m^2 area and the sink node is in the center. The network simulator, NS2 is utilized for the purpose of simulation in the proposed research work. Here, the number of nodes taken is 400 which are randomly organized by means of uniform distribution in different parts of the deployment area with a predetermined density. The packet size is kept at 40 bytes. The Distributed Coordination Function (DCF) of IEE 802.11 is employed for wireless LANs as the MAC layer protocol. All nodes are presumed to progress simulated traffic is at Constant Bit Rate (CBR) with an average packet rate of 0.5 packets/sec. CBR is appropriate autonomously with identical average speed. The for the data that the end-systems need to predictable reaction time and a static quantity of bandwidth is constantly available for the life span of the connection.

Table 1: Simulation parameters

Parameters	Values
Number of nodes	400
Area (size)	400×400 m
Mac	802.11
Traffic source	CBR
Initial energy	100 J
Packet size	40 bytes
Transmission cost per message	0.645 mJ
Transmission cost per byte	0.0144 mJ/byte
Receiving cost per message	0.387 mJ
Receiving cost per byte	0.00864 mJ/byte
Antenna	Omni antenna
Radio propagation	Two way road
Interface queue	Queue/drop tail
Queue length	50
Channel type	Channel/wireless channel
Routing protocol	AODV
Number of bits vector saved at a node	50

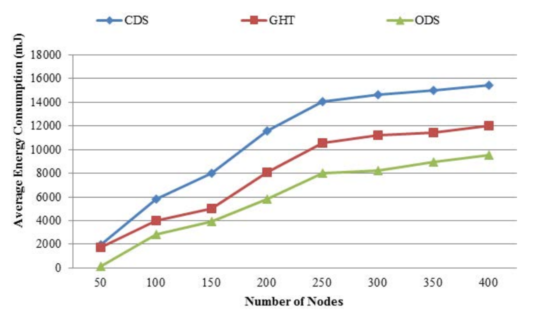

Fig. 2: Average energy cost consumption with different number of nodes

The performance of the various schemas such as Centralized Data Storage (CDS), Optimal Data Storage (ODS) and Near-optimal Data Storage (NDS) is measured. CDS is a centralized strategy (Madden *et al.*, 2015). It opts for the center node in the network area as the base station which producers deliver data to and consumers query data from. The ODS and NDS approaches are based on (Yu *et al.*, 2019). However, compared to NDS and CDS, ODS maintains higher fault stolerance.

Energy comparison: Figure 2 shows the relationship among the energy consumption and change of the number of consumers. Figure 2 shows the average energy cost consumption when the number of sensor nodes is 7 and the sensing area is 400×400 m. It is evident that the residual energy of the proposed hybrid ODS scheme is higher than that of schemes CDS and GHT after 56 runs. The data transmission cost between the nodes is

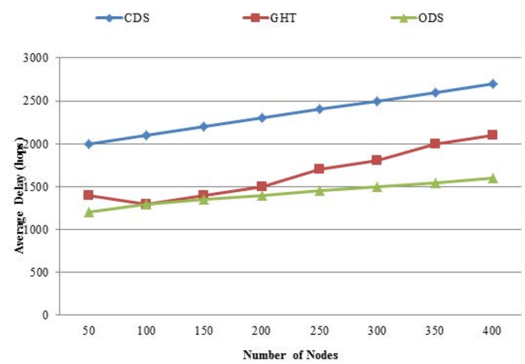

Fig. 3: Average delay consumption with different number of storage nodes

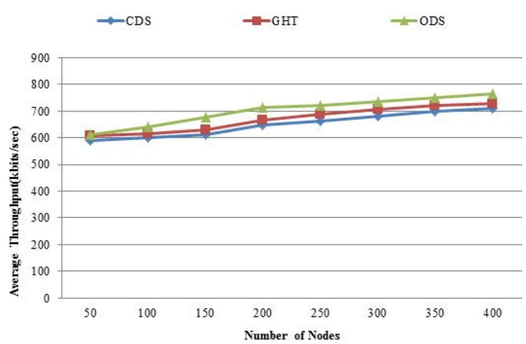

Fig. 4: Average throughput comparison with different number of storage nodes

minimized. Thus, the energy consumption is saved. When the number of consumer is increasing, then the network will increase the energy consumption but in the proposed hybrid ODS data storage energy consumption increased less than CDS and GHT. The difference between the proposed algorithm and the optimal begins to increase when the storage node number becomes larger.

Average delay comparison: Figure 3 shows the comparison of the average delay caused by the approaches in one-to-one mode. It is observed from the Fig. 3 that the proposed ODS approach provides significant results with lesser average delay when compared with the CDS and GHT approaches. When the number of nodes increases, the average delay also increases. However, the proposed ODS attains less average delay than the existing CDS and GHT approaches.

Average throughput comparison: Figure 4 shows the comparison of the throughput obtained by the approaches. When the number of nodes increases the throughput also increased. The proposed ODS attains high throughput when the number of nodes increases. It is observed from the figure that the throughput is higher for the proposed ODS approach when compared with CDS and GHT approaches. This is mainly due to the communication overhead.

CONCLUSION

The most important issue in designing Wireless Sensor Networks (WSNs) is the support of the data latency as functional requirement and support of non functional requirements such as information integrity, communication cost, energy consumption, etc. While designing WSNs positions of sensor nodes have been becomes very difficult task for achieving the desired design goals. This survey addresses the data storage position problem in WSN. In these data storage strategies, storage node have been introduced to decrease the data transmission cost of data and to increase the network lifetime of all WSN. This survey also addresses the issues of all previous methods and it is considered as the scope of future work. Also implement the data storage strategies and conduct some experiments which show that heuristic algorithm is practical and well-organized.

It can be concluded that the Meta heuristic method has not been utilized so far to find the optimal storage node position with less energy cost. So in the future work we formulate the storage node position problem and propose a Meta heuristic algorithm for it. Future work, also regard as the energy balance between sensor nodes, predominantly for the sensor nodes with the intention of forward raw data and in addition it also consider how to solve the data storage position problem in terms of other performance metrics such as packet delivery ratio, higher throughput and less power consumption, etc.

REFERENCES

Ahmed, K. and M.A. Gregory, 2012. Techniques and challenges of data centric storage scheme in wireless sensor network. J. Sensor Actuator Networks, 1: 59-85.

Albano, M., S. Chessa, F. Nidito and S. Pelagatti, 2011. Dealing with nonuniformity in data centric storage for wireless sensor networks. IEEE. Trans. Paral. Distrib. Syst., 22: 1398-1406.

Goyal, S. and M.S. Patterh, 2013. Wireless sensor network localization based on bat algorithm. Int. J. Emerging Technol. Comput. Appl. Sci., 13: 507-512.

Huang, Y., J. Cao, B. Jin, X. Tao and J. Lu *et al*., 2010. Flexible cache consistency maintenance over wireless ad hoc networks. IEEE Trans. Paral. Distrib. Syst., 21: 1150-1161.

Jagtap, A.M. and M.A. Shukla, 2013. Optimized node deployment using enhanced particle swarm optimization for WSN. Int. J. Emerging Techonol. Adv. Eng., 3: 170-174.

Kabiri, M. and J. Vahidi, 2014. Analysis of performance improvement in wireless sensor networks based on heuristic algorithms along with soft computing approach. J. Math. Comput. Sci., 3: 47-67.

Kaur, G.V. P. Luthra, 2014. A Survey on applications of genetic algorithm in wireless sensor networks. Int. J. Adv. Res. Comput. Sci. Software Eng., 4: 664-668.

Khamforoosh, K. and H. Khamforoush, 2009. A new routing algorithm for energy reduction in wireless sensor networks. Proceedings of the 2nd IEEE International Conference on Computer Science and Information Technology, August 8-11, 2009, IEEE, Beijing, China, ISBN: 978-1-4244-4519-6, pp: 505-509.

Li, D. and X.B. Wen, 2015. An improved PSO algorithm for distributed localization in wireless sensor networks. Int. J. Distrib. Sens. Networks, 2015: 1-8.

Liao, W.H. and H.C. Yang, 2012. An energy-efficient data storage scheme in wireless sensor networks. Proceedings of the 2012 IEEE Symposium on Network Operations and Management, April 16-20, 2012, IEEE, Maui, Hawaii, ISBN: 978-1-4673-0267-8, pp: 478-481.

Madden, S.R., M.J. Franklin, J.M. Hellerstein and W. Hong, 2005a. TinyDB: An acquisitional query processing system for sensor networks. ACM Trans. Database Syst., 30: 122-173.

Madden, S.R., M.J. Franklin, J.M. Hellerstein and W. Hong, 2005b. TinyDB: An acquisitional query processing system for sensor networks. ACM. Trans. Database Syst. TODS., 30: 122-173.

Maraiya, K., K. Kant and N. Gupta, 2011. Architectural based data aggregation techniques in wireless sensor network: A comparative study. Int. J. Comput. Sci. Eng., 3: 1131-1138.

Rebai, M, H. Snoussi, L. Khoukhi and F. Hnaien, 2013. Linear models for the total coverage problem in wireless sensor networks. Proceedings of the IEEE International Conference on Modeling Simulation and Applied Optimization Hammamet Tunisia, April 28-30, 2013, IEEE, Hammamet, Tunisia, ISBN: 978-1-4673-5812-5, pp: 1-4.

Rebai, M.L., M. Berre, F. Hnaien, H. Snoussi and L. Khoukhi, 2014. A branch and bound algorithm for the critical grid coverage problem in wireless sensor networks. Int. J. Distrib. Sens. Networks, 2014: 1-9.

Sarkar, R., X. Zhu and J. Gao, 2009. Double rulings for information brokerage in sensor networks. IEEE./ACM. Trans. Networking, 17: 1902-1915.

Shen, H., L. Zhao and Z. Li, 2011. A distributed spatial-temporal similarity data storage scheme in wireless sensor networks. IEEE. Trans. Mob. Comput., 10: 982-996.

Sheng, B., Q. Li and W. Mao, 2010. Optimize storage placement in sensor networks. IEEE. Trans. Mobile Comput., 9: 1437-1450.

Tang, B. and H. Gupta, 2007. Cache placement in sensor networks under an update cost constraint. J. Discrete Algorithms, 5: 422-435.

Wang, Y., X. Fu and L. Huang, 2009. A storage node placement algorithm in wireless sensor networks. Proceedings of the 2009 4th International Conference on Frontier of Computer Science and Technology, December 17-19, 2009, IEEE, Shanghai, China, ISBN: 978-0-7695-3932-4, pp: 267-271.

Yu, Z., B. Xiao and S. Zhou, 2009. Achieving optimal data storage position in wireless sensor networks. Comput. Commun., 33: 92-102.

Yu, Z., B. Xiao and S. Zhou, 2010. Achieving optimal data storage position in wireless sensor networks. Comput. Commun., 33: 92-102.

Zhanbei, W.A.N.G., 2014. Node localization of wireless sensor networks based on bat algorithm. Comput. Eng. Appl., 11: 019-94.