International Journal of Research (IJR) Vol-1, Issue-7, August 2014 **ISSN 2348-6848**

A Survey on Data Approximation Techniques for Wireless Sensor Networks

Pavitha N

Department of Computer Engineering, Sinhgad Academy of Engineering, Pune,Maharashtra, India

Email: pavithanrai@gmail.com

Abstract:

Wireless sensor networks are very useful in a variability of locales. A main contest in these networks is to minimize energy consumption. By minimizing energy consumption we can increase the lifetime of wireless sensor networks. The recent studies on wireless sensor networks shows that by applying data approximation techniques we can significantly improve the lifetime of wireless sensor network. In this backdrop this paper formalizes various data approximation techniques that can be applied to wireless sensor networks. By applying these data approximation techniques we can significantly reduce the number of messages transmitted. Thereby communication overhead will be reduced. Reducing communication automatically increases the lifetime of wireless sensor networks.

Keywords:

wireless sensor network, energy consumption, lifetime, data approximation, communication overhead.

1. INTRODUCTION

Recent technological improvements have made the deployment of small, inexpensive, low-power, distributed devices, which are capable of local processing and wireless communication, a reality. Such nodes are called as sensor nodes. Each sensor node is capable of only a limited amount of processing. But when coordinated with the information from a large number of other nodes, they have the ability to measure a given physical environment in great detail. Thus, a sensor network can be described as a collection of sensor nodes which co-ordinate to perform some specific action. Unlike traditional networks, sensor networks depend on dense deployment and co-ordination to carry out their tasks. Sensor networks open up new opportunities to observe and interact with the physical world around us. They enable us to gather data that was until now difficult, expensive, or even impossible to collect.

Due to the wide variety of possible applications of WSNs, system requirements could change significantly. For instance, in case of environmental monitoring applications, the following requirements are typically dominant: *energy efficiency*, nodes are battery powered or have a limited power supply; *low data rate*, typically the amount of data to be sensed is limited; *one-way communication*, nodes act only
as sensors and hence the data flow is from nodes to sink(s); *wireless backbone*, usually

in environmental monitoring no wired connections are available to connect sink(s) to the fixed network.

II. NEED OF DATA APPROXIMATION

For long term wireless sensor networks with bounded bandwidth Approximate Data Collection is a good choice. In many situations with densely positioned sensor nodes, the sensor data collected by sensor nodes may have a spatial – temporal correlation characteristic. By comprehending such characteristics, the sensor nodes can be collecting data in an aggressive manner. Consequently the data accuracy will be maximized and data traffic will be minimized. This technique is adaptive to physical environmental changes and it reduces number of messages transmitted i.e. congestion is controlled. This efficient technique will improve the lifetime of wireless sensor network.

III. DATA APPROXIMATION TECHNIQUES

A. Ken Method [1]

This paper proposes a robust approximate technique called Ken that uses replicated dynamic probabilistic models to minimize communication from sensor nodes to the networks PC base station. This approach is based on a form of compression using replicated dynamic probabilistic models. The basic idea is to maintain a pair of dynamic probabilistic models over the sensor network attributes, with one copy distributed in the sensor network and the other at a PC base station.

At every time instance (i.e., with a frequency f), the base station simply computes the expected values of the sensor net attributes according to the model and uses it as the answer to the "SELECT * FREQ f" query. This requires no communication.

An attractive feature of the Ken architecture is that it naturally accommodates applications that are based on event reporting or anomaly detection; these include fire-alert-and response and vehicle tracking. In these scenarios, the sample rate is typically quite high, but the communication rate should remain quite low under most circumstances. The model reflects the expected "normal" state of the environment being monitored; anomalies result in reports being pushed to the base station for urgent handling by infrastructure logic. In addition to naturally supporting these applications, Ken enhances them with additional functionality: the ability to support interactive query results with well-bounded approximate answers. In essence, approximate data collection and event detection become isomorphic. Advantage of this method is that the readings which are visible to the user are with in fixed error bound. Disadvantage of using this method is that it is robust to communication failure that is, if the data lost during transmission it cannot be recovered.

B. Directed Diffusion [2]

This is a data centric approach, in this method all the communication is for the named data. In directed diffusion method sink broadcast its request to its neighbour. The neighbour cached the interests and the neighbour data which matching to the interest will respond back to sink. The human operator's query would be transformed into an interest that is diffused towards nodes in regions an interest that is diffused towards nodes in regions X or Y. When a node in that region receives an interest, it activates its sensors which begin collecting information about pedestrians. When the sensors report the presence of pedestrians, this information returns along the reverse path of interest propagation. Intermediate nodes might aggregate the data, e.g., more accurately pinpoint the pedestrian's location by combining reports from several sensors. An important feature

of directed diffusion is that interest and data propagation and aggregation are determined by localized interactions (message exchanges between neighbours or nodes within some vicinity).

Directed diffusion is significantly different from IP-style communication where nodes are identified by their end-points, and inter-node communication is layered on an end-to-end delivery service provided within the network. In this approach the number of transmissions will be reduced because of interested neighbour will be responded. The disadvantage of this method is that it is not suitable for long term large scale wireless sensor networks.

C. Similarity based adaptive framework [3]

This method detects data similarities among the nodes and groups them into clusters. Its main goal is to conserve energy by reducing transmission. Advantage of this method is that it reduces the amount of communication in wireless sensor networks. Disadvantage of using this method is that it only takes an advantage of temporal correlation with in the sensor data without considering the similar readings of nearby sensor nodes.

D. *Bar-B-Q Method (BBQ)[4]*

This is a query centred approach, which is used to extract data from sensor networks. It is the first method which uses Gaussian joint distribution model to extract data from sensor network. This approach will find correlations among sensor readings by using the data extracted from the sensor networks. It has numerous drawbacks because of using Gaussian joint distribution model. Firstly it is not suitable for large scale wireless sensor networks. Secondly, it needs to gather complete data set, but gathering of complete data set is too much energy consuming.

E. CONCH (Constraint Chaining) [5]

This is a continuous monitoring scheme. This method uses a spatial and temporal correlation to achieve energy saving. In a sensor network, the observed values are either spatially or temporarily correlated which implies that it is not needed to report each and every value in each time if it is not changed from the existing value. It will assign reporter and updater at the edge, the role of the reporter is to report the base station if there is a change in value and the updater job is to update the value reported by the reporter. The disadvantage of using this technique is that it goes through update message overhead when the readings changes often.

F. Approximate Data Collection (ADC) [6]

This method divides a sensor network into cluster and each cluster will have its own cluster head and it determines data association, and then performs approximation on sink node by using a parameter updated by a data model. The node which has more power will be chosen as a cluster head. The cluster head will be selected by cluster members. The doorway algorithm is implemented into the cluster head to approximate the readings of sensor nodes.

Each sensor node transmits its parameter to its cluster head, instead of intense sensor readings. If the difference between the approximated values is not greater than the original value the sensor node does not update its data to cluster head. This method is suitable if less frequent update operations are there. Disadvantage of this method is that if frequent data updates are required then transmitting parameters leads to increase in network traffic.

IV. CONCLUSION

By minimizing energy consumption we can increase the lifetime of wireless sensor networks. The recent studies on wireless sensor networks shows that by applying data approximation techniques we can

significantly improve the lifetime of wireless sensor network. In this paper a study on various data approximation techniques that can be applied to wireless sensor networks is carried out. By applying these data approximation techniques we can significantly reduce the number of messages transmitted. Thereby communication overhead will be reduced. Reducing communication automatically increases the lifetime of wireless sensor networks.

REFERENCES

[1] D. Chu, A. Deshpande, J.M. Hellerstein, and W. Hong, "Approximate Data Collection in Sensor Networks Using Probabilistic Models," *Proc. 22nd Int'l Conf. Data Eng. (ICDE '06)*, 2006

[2] C. Intanagonwiwat, R. Govindan, and D. Estrin, "Directed Diffusion: A Scalable and Robust Communication Paradigm for Sensor Networks," *Proc. MobiCom, 2000.*

[3] D. Tulone and S. Madden, "An Energy Efficient Querying Framework in Sensor Networks for Detecting Node Similarities," *Proc. Ninth ACM Int''l Symp. Modeling Analysis and Simulation of Wireless and Mobile Systems (MSWiM ''06)*, 2006.

[4] A. Deshpande, C. Guestrin, S. Madden, J. Hellerstein, and W. Hong, "Model-Driven Data Acquisition in Sensor Networks," *Proc. 13th Int''l Conf. Very Large Data Bases (VLDB ''04)*, 2004.

[5] A. Silberstein, R. Braynard, and J. Yang, "Constraint Chaining: On Energy-Effcient Continuous Monitoring in Sensor Networks,"*Proc. ACM SIGMOD Int''l Conf. Management of Data (SIGMOD ''06)*,2006.

[6] Dhivya S, S. Manoj Kumar "Cluster Based Approximate Data Collection for Wireless Sensor Networks" *International Conference on Computer Communication and Informatics (ICCCI -2013)*, Jan. 09 – 11, 2013, Coimbatore, INDIA.

[7] Chao Wang, Huadong Ma, Yuan He and Shuguagn Xiong "Adaptive Approximate Data Collection for Wireless Sensor Networks" *ieee transactions on parallel and distributed systems, vol. 23, no. 6*, June 2012.

[8] M. Li, Y. Liu, and L. Chen, "Non-Threshold Based Event Detection for 3D Environment Monitoring in Sensor Networks," *IEEE Trans. Knowledge and Data Eng., vol. 20, no. 12, pp. 1699-1711*, Dec. 2008.

[9] R. Rajagopalan and P.K. Varshney, "Data-aggregation techniques in sensor networks: a survey" *IEEE Comm. Surveys & Tutorials, vol. 8, no. 4, pp. 48–63*, 2006.

[10] You-Chiun Wang, "*Data Compression Techniques in Wireless Sensor Networks*".

[11] Yujie Zhu, Ramanuja Vedantham, Seung-Jong Park and Raghupathy Sivakumar "A scalable correlation aware aggregation strategy for wireless sensor networks" *Information Fusion Volume 9, Issue 3*, July 2008.

[12] Teerawat Issariyakul , Ekram Hossain, "*Introduction to Network Simulator NS2*".

[13] Ashu Gupta, R.Annie Uthra, "Cluster Based Approximate Data Collection in Wireless Sensor Network" *International Journal of Electronics and Computer Science Engineering.*

[14] The RED queing discipline, *http://opalsoft.net/qos/DS-26.htm.*