Volume 5, Issue 6, June 2015 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

An Efficient Data Collection Technique in Wireless Sensor Networks

Shruti Kaushik[1], Baninder Singh[2]
[1] Student, [2]Associate Professor
Yamuna Institute of Engineering & Technology,
Haryana, India

Abstract: ***The Wireless Sensor Network can be divided in to two parts: one that contains the colossal amount of sensor nodes and the other which includes the network formed by those sensor nodes. The sensor nodes are equipped with limited resources. Consequently, WSNs face the problems like partitioned WSNs due to which sensed data gets discarded by the base station, data inconsistency problem, and energy inefficiency. A mobile robot is needed to be used that can detect the partitioned WSNs and can effectively collect the sensed data from the sensor nodes. Our main aim is to minimize the travelling distance of the mobile robots. So that sensed data can be fetched effectively and with minimum time. We also propose and implement an approach named pattern variation discovery to detect the irregularities exist in the data collected by the sensor nodes in wireless sensor networks.***

Keywords: Wireless Sensor Networks, data mining, data inconsistency, robot routing, travelling salesman problem with neighbourhood

I. INTRODUCTION

Wireless Sensor Networks (WSNs) has proved itself as an emerging technology that can be used to solve many real-life problems. WSN's consists of a large number of sensor nodes that are smart enough to not only monitor the changes happening in the environment in real time but also intelligent enough to gather the information related to it. These sensor nodes are very cost effective if they are being compared with those traditional one's, they are small and come up with limited power, computational and processing resources. These sensor nodes also possess the capability to communicate with each other in order to transfer the collected data to the sink node which is the main central location responsible for wholesome management of all other sensor nodes in WSN. To make the communication possible, a network is formed by distributing the sensor nodes across the environment. Once the network is formed, all the sensor nodes liaise with each other to carry out computing and sensing activities. Its cost effectiveness, better output giving capability, easy implementation and troubleshooting and no mess of cables, are the reasons behind its wide popularity and adoption.
Today's wireless sensor networks are bi-directional and are intelligent enough to control the sensor activities. The main reason behind the innovation of Wireless sensor networks was the need to build and support military applications like battlefield surveillance. But nowadays, WSNs are used most effectively in monitoring and sensing many physical phenomenon happening in the environment such as change in temperature, humidity, intensity, pressure and even in the field of health care, WSN has gained popularity.

Types of Wireless Sensor Networks (WSNs)

There are mainly 2 types of WSNs:

1. Structured WSN
2. Unstructured WSN

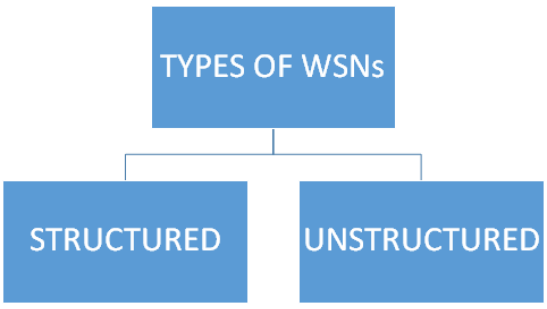

Fig1: Block diagram that depicts the types of WSNs

1. Structured WSNs:

Structured WSNs has limited number of sensor nodes that are distributed in predefined manner. Hence, maintenance is quite simpler unlike in unstructured wireless sensor networks and due to lesser number of sensor nodes involved, underlying cost is also less.

2. Unstructured WSNs:

Unlike structured WSN, unstructured WSNs possess a large number of sensors nodes which are distributed randomly in the network. Due to its ad-hoc nature and large number of sensor nodes involvement, maintenance becomes a tedious job to do. Moreover, it is more prone to connection failure as compared to Structured WSNs and the detection of connection failure is also an overhead.

So, if comparison needs to be made between structured and unstructured WSNs we can say that structured WSNs is more advantageous than unstructured WSNs as it is more cost effective and includes less maintenance and management cost [1].

II. ARCHITECTURE

WSNs have no bound when it comes to topology. These are very flexible to adopt any kind of topology. It can be deployed either by using any simple topology like star topology or by using a complex one for example multi-hop wireless mesh network.

A typical multi-hope wireless mesh network architecture is given below:

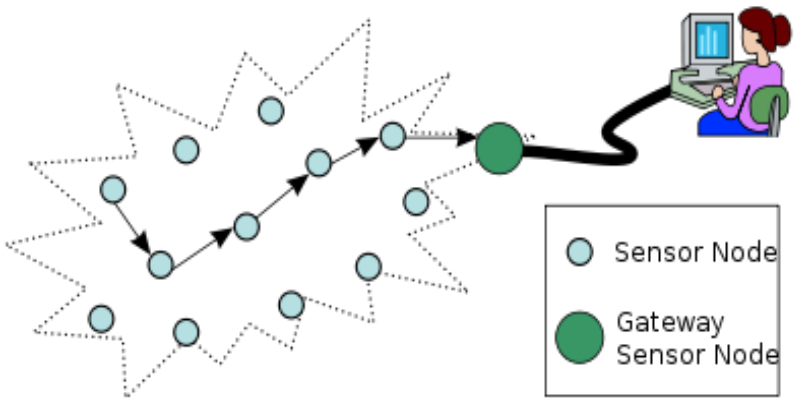

Fig. 1 A typical multi-hope wireless mesh network architecture

As the diagram itself depicts, WSN architecture is mainly divided into 3 core things:

1. Sensor Nodes
2. Gateway Sensor Node (Gateways)
3. Software installed on host computer

How the process flows:

- The sensor nodes are distributed into the real world in order to sense the data and gather the information by monitoring its surroundings
- That sensed data is needed to be sent wirelessly hop-by-hop through the gateway nodes which is in further connected to the host computer. Actually, gateway nodes relays the data received from sensor nodes to the host computer through wired transmission.
- As soon as the gathered data reached at host computer, we can analyze it, present it and can perform any other computational tasks in order to convert into meaningful information by using the software installed on the host information.

Wireless Sensor Node:

A wireless sensor network is composed of various sensor nodes which are responsible of sensing, tracking and collecting the useful data. A sensor node can be divided into following parts which help it to perform the above mentioned tasks:

1. A microcontroller with memory
2. Power Supply
3. A radio transceiver with an antenna
4. A sensor array
5. Wireless communication interface that provides interfacing of sensor array with power source

The above parts altogether responsible for the working of a sensor node. The below block diagram helps to provide the clear vision about the components of a sensor node:

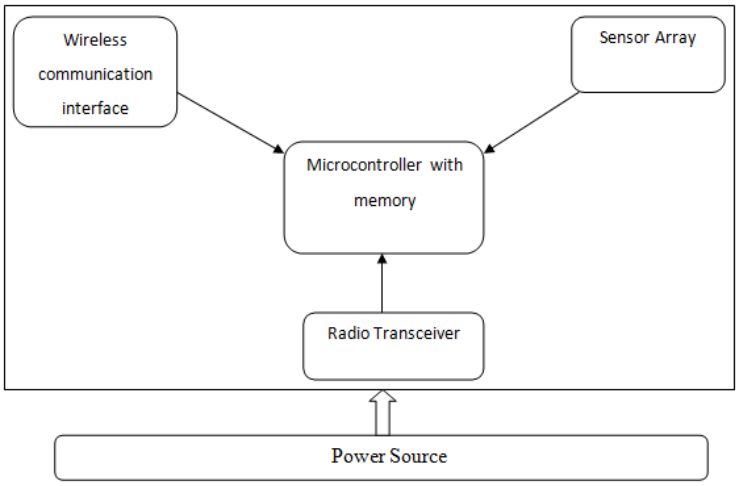

Fig. 2: Block Diagram of components of a Sensor Node

Issues Involved in WSNs

WSNs are composed of large number of sensor nodes which are provided with limited resources. Moreover, WSNs work wirelessly in the harsh environment. So, WSNs not only face the resource issues but also the network issues like connection failure, change in topology etc. Below is the brief discussion of the issues that WSNs encounter while working its tasks:

1. Limited Power Supply:

The sensor nodes are provided with the limited power resource which affects its energy efficiency. Sensor nodes require energy to perform it's all the tasks like collection of data , co-ordination with other sensor nodes, transmission of sensed data and monitoring of any change occurring in the environment. But with the limited energy, the life span of the sensor nodes gets shortened. The major problem which arises when the energy of a sensor node starts depleting is that it won't be continue its work further and won't be able to coordinate with the other sensor nodes available in the environment. Ultimately, that particular sensor node will be declared as "dead node". Due to which, the affected WSN will segregate itself from the other WSNs available in the network and all the data collected by that portioned WSN will not be transmitted to the sink node (base station) as this portioned WSN cannot be seen by the base station [2].

2. Consistency of Data

The sensor nodes not always collect the meaningful data. The reasons behind this can be interferences happen in wireless environment, harsh network conditions, and connection failures can cause the collection of unreliable data. Moreover, it is not always required to sense data which in turn consumes the energy which is not at all required. So there is a need to analyze when the data should be collected and how it should be turned into the meaningful information. Various data mining techniques can be used to perform this task [3].

3. Routing and Maintenance:

Due to the massive sensor nodes involvement and wireless nature, maintenance and management of these nodes are become necessary. So that the working of sensor nodes and the coordination between them keep on going smoothly without any hindrance. Plus, WSNs should be capable enough to adjust and revamp itself if there is any change in topology and routing have been occurred.

4. Heterogeneous nature of sensor nodes:

As discussed, WSNs are composed of a colossal amount of sensor nodes each of which has their own potential in terms of memory, processing power and sensing ability [4]. Hence, while distributing the loads to the sensor nodes, system should consider the abilities of each node and give them load according to their abilities and potential.

Remedies

- The partitioned WSNs problem, which is caused due to the energy depletion of energy of the sensors nodes, can be optimized by using the mobile robots [2]. These mobile robots can gather and send the sensed data collected by the partitioned WSNs to the sink node. But for this, the location of the partitioned WSN is necessary to be known.
- Analysts take many important decisions based on the data collected by sensor nodes in the WSNs. So, the date needs to be accurate and up-to-date. For this, various data mining techniques can be implemented to analyze and extract meaningful data from the database dump [3].
- If mobile robots are being used to collect the sensed data from the partitioned WSNs, then there is a need of using a suitable navigational strategy that can not only help the mobile robots to detect the partitioned WSNs but also guide it to how to deal with the snags comes into its path.

III. RELATED WORK

A. OVERVIEW ON DATA COLLECTION USING MOBILE ROBOTS IN WIRELESS SENSOR NETWORK:

[2] Cheng Chen, Tzung-Shi Chen and Ping-Wen Wu in 2011 have proposed a data collecting algorithm that guides the mobile robot to collect the sensed data from the partitioned WSNs which earlier used to be discarded by the base station as it could not be able to locate the partitioned WSNs. Due to the rejection of the sensed data of partitioned WSNs by base station, problems like data inconsistency comes into play. To resolve such issues, authors have proposed 2 approaches for the navigation of island WSNs: 1) Local based approach and 2) Global based approach. Furthermore, authors have also specified scheduling techniques that are required to schedule the navigational scheme of mobile robot. The scheduling techniques nare based on three things: 1) time, 2) location and 3) dynamic moving based. Ns-2 simulator has been used to simulate the scenarios and approaches.

B. NEAREST NEIGHBOUR CLASSIFICATION FOR WIRELESS SENSOR DATA:

[3] Khushboo Sharma, Manisha Rajpoot and Lokesh Kumar Sharma have used a Nearest Neighbour Classification technique to categorize the data received by sensor nodes in wireless sensor networks. This technique solves the problem of data inconsistency which occurs due to the rejection of sensed data collected by partitioned WSNs and the irrelevant data collected by the sensor nodes of WSNs. A Nearest Neighbour Trajectory is used for the categorization of data sensed by sensor nodes. A nearest neighbour classifier processes the patterns only when the request is made to classify a query vector. The technique works in two parts: 1) predicts the classes of sensor data and 2) Nearest neighbour trajectory begins the training of sensor data and construct a model.

C. HYBRID MOBILE ROBOT NAVIGATIONAL STRATEGY FOR EFFICIENT DATA COLLECTION IN WSNs

[5] Marcelo B. Soares, Mario F. M. Campos, Dimas A. Dutra, V´_ctor C. da S. Campos and Guilherme A. S. Pereira in 2007 have come up with the Hybrid Mobile Robot Navigational Strategy which is used to detect the exact location of partitioned WSNs in order to guide the mobile robot to collect the sensed data from them. This hybrid mobile robot navigational strategy works at two different layers: 1) Reactive Layer and 2) Planning Layer. The reactive layers help in constructing the model of collected data and planning layer provides the guidelines for the mobile robot. So that it can collect the data from partitioned WSNs without any hindrance.

D. OPTIMAL ROBOT ROUTING PROBLEM IN WIRELESS SENSOR NETWORKS

[6] Bo Yuan, Member, IEEE, Maria Orlowska, and Shazia Sadiq in 2007 have discussed how the mobile robots can be used to visit given sensor nodes distributed in Euclidean plane in order to collect data from them.The authors have proposed an approach which can be used to solve NP-hard problem which is a special case of Travelling Salesman Problem and Neighbourhood (TSPN) which in turn helps in minimizing the total travelling distance covered by the mobile robot while collecting the sensed data.

IV. PROPOSED WORK

Research Goals:

Given a set of sparsely distributed sensors in the plane, a mobile robot is required to visit all sensors to download the data and return back. The effective range of each sensor is specified and the robot must at least reach the boundary to start communication. The primary goal of optimization is to minimize the travelling distance by the robot. The problem can be regarded as a special case of the Travelling Salesman Problem with Neighbourhoods (TSPN), tiny sensor nodes, equipped with sensing, communication capabilities and computation can be deployed in large numbers in geographical areas to monitor, detect and report events. Wireless networks consisting of such sensors create exciting opportunities for large-scale and surveillance applications. In many of the applications, it is essential to mine the sensor readings for patterns in real time in order to make intelligent decisions.

Detection of sensor data irregularities is useful for practical applications because the patterns found can be used for both decision making in applications and system performance tuning. For example, irregularities in sensory data are of interest of monitoring applications. For this kind of applications, the communication cost can be reduced if only normal sensory values, as opposed to all values, are transmitted.

The problem of irregularities detection is to find those sensory values that deviate significantly from the norm. This problem is important in the sensor network setting because it can be used to identify abnormal or interesting events or faulty sensors.

Research Methodology:

A new approach named pattern variation discovery is used to solve this problem. In this, first we need to give the definition of normal patterns. This definition can be models of multiple sensory attributes or constraints among multiple attributes. The next step includes the discovery of irregularity. Whenever a normal pattern is broken at some point, irregularity appears. That is, the pattern variation happens. Detection of sensor data irregularities

For example, we want to discover the irregular distribution pattern among multiple sensory attributes along time. Then, for each time point, we can put the values of a group of sensory attributes at a series of sensor nodes into a matrix, which represents a distribution status. The problem then becomes to discover the irregular matrix among a set of matrices. An irregular matrix represents that, at the corresponding time point, the distribution pattern of all the sensory attributes on all the nodes are irregular.

Detection of irregularities is tightly interrelated to modelling of sensor data. Therefore, we propose to detect irregular single-attribute sensor data with respect to time or space by building models.

V. RESULTS

The proposed data mining technique is implemented in Tcl and results are simulated under the environment of Network Silulator-2 (NS-2). Firstly, analysis has been made to check the number of packets with and without anomalies populate in wireless sensor networks as we keep on increasing simulation time. The following table and line graph predict the analysis that has been made:

TABLE 1 NUMBER OF NORMAL AND ABNORMAL PACKETS AT DIFFERENT SIMULATION TIME

Simulation Time	Abnormal Packets	Normal Packets
10	1493	195
15	1802	436
20	2043	635
25	2332	878

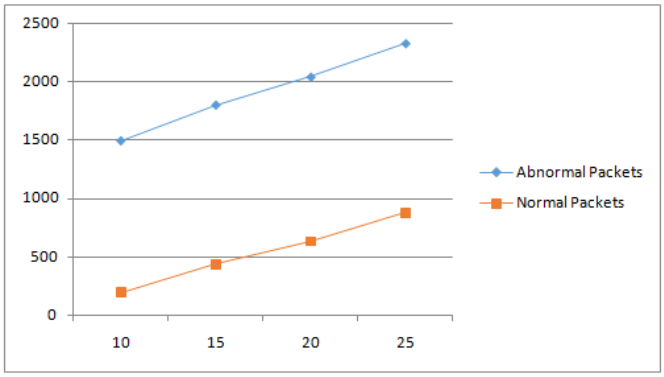

Fig. 3: A line graph between Number Of Packets (normal and abnormal) and simulation time

After the analysis, we applied our data mining technique to obtain the results that show that this technique is quite helpful to save the energy of wireless sensor networks. There are total 30 nodes in the network and out of which only 10 are using the energy of complete WSN. Below are the screenshots of the results that are obtained using ns-2 simulator:

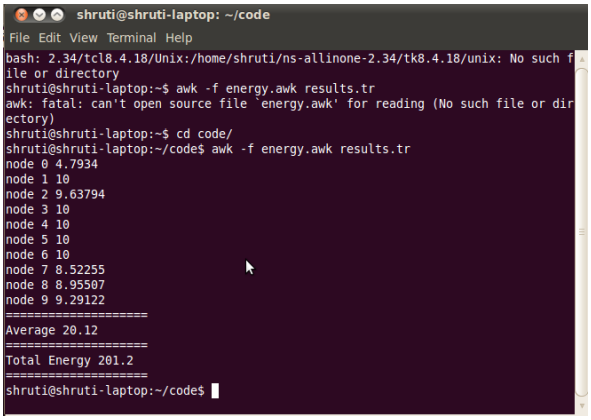

Fig. 4: Total energy and average energy consumed by the 10 nodes in WSN

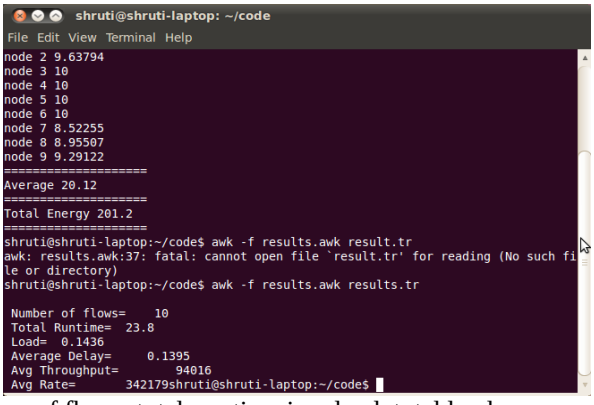

Fig. 5 Results on the basis of no. of flows, total run time involved, total load, average delay faced, average rate and average throughput achieved

VI. CONCLUSION

A pattern variation discovery is used to improve the energy efficiency of WSNs. The packets with irregularities are discarded to save the energy and improve the throughput of the WSN. The screenshots make it much clear that if a pattern variation discovery technique is used, the total load can also be reduced. In spite of 30 nodes, only 10 nodes are consuming the energy that carries relevant data. Thus, the introduction of this data mining technique can improve the sensed data-collecting and energy efficiency of WSNs.

REFERENCES

[1] Jennifer Yick, Biswanath Mukherjee, DipakGhosal, Department of Computer Science, University of California, Davis, CA 95616, United States

[2] Tzung-Cheng Chen, Tzung-Shi Chen, Member, IEEE, and Ping-Wen Wu.

[3] Khushboo Sharma, Manisha Rajpoot, Lokesh Kumar Sharma ,Department of Computer Science and Engineering, Rungta College of Engineering and Technology, Kohka Road, Kurud, Bhilai, India.

[4] AdiMallikarjuna Reddy V AVU Phani Kumar, D Janakiram, and G Ashok Kumar, Distributed and Object Systems Lab, Department of Computer Science and Engineering, Indian Institute of Technology Madras, Chennai, India – 600036,fadi,phani, d.janakiram,ashokg@cs.iitm.ernet.in, May 3, 2007

[5] Marcelo B. Soares, Mario F. M. Campos, Dimas A. Dutra,V´_ctor C. da S. Campos and Guilherme A. S. Pereira

[6] Bo Yuan, Member, IEEE, Maria Orlowska, and Shazia Sadiq, School of Information Technology and Electrical Engineering, The University of Queensland, QLD 4072, Australia, E-mail: {boyuan, maria, shazia}@itee.uq.edu.au.

[7] L. K. Sharma, O. P. Vyas, S. Schieder, and A. K. Akasapu,"A NearestNeighbour Classification for Trajectory Data", Springer CCIS, Volume 101, Part 1, pp. 180-185, 2010.

[8] H. Park, W. Liao, K. H. Tam, M. B. Srivastava, and L. He. A uni_ed network and node level simulation framework for wireless sensor networks. Technical report, September 7 2003.

[9] E.M. Arkin and R. Hassin, "Approximation Algorithms for the Geometric Covering Salesman Problem," Discrete Applied Math., vol. 55, no. 3, pp. 197-218, 1994.

[10] Wang Zhangqi ,Zhu Xiaoguang ,Han Qingyao, School of Mechanical Engineering,North China Electric Power University.071003 Baoding, China

