COMPUSOFT, An international journal of advanced computer technology, 4 (2), February-2015 (Volume-IV, Issue-II)

ISSN:2320-0790

COMPUSOFT
An International Journal of Advanced Computer Technology

Data-driven Energy Conservation in Wireless Sensor Networks: A Survey

Sunil Dhimal, Kalpana Sharma

Sikkim Manipal Institute of Technology
Majhitar, Rangpo, Sikkim, India

Abstract: Wireless Sensor Network (WSN) consists of large number of sensor nodes deployed at physical location for sensing & monitoring some physical phenomenon like temperature, humidity, pressure, soil moisture etc. Energy Conservation is one of the prime challenges in WSN owing to the limited battery power and unattended deployment of sensor nodes. There are various techniques motivated towards conserving energy in order to increase overall lifetime of the network. Various schemes and protocols like energy efficient media access control and Routing protocols, Data cycling techniques, Power Management, Sleep Management, Data-driven techniques etc. are used for energy conservation. This paper mainly focuses on exploitation of data-driven techniques in which the sensed data is statistically analyzed in order to find a clue which can save energy requirements. Spatial and temporal correlation between sensor node data can be exploited to propose and efficient data-acquisition technique. The inter-node data association can be measured by calculating the most commonly used metrics for measuring degree of similarity. Finally, data-driven energy conservation technique can be applied if high correlation between inter-node observations is identified.

Keywords: WSN, Similarity Metrics, Euclidean Distance, Cosine similarity, Pearson correlation coefficient.

I. Introduction

WSN has an immense scope in designing an application for remote sensing, monitoring and analyzing parameters of interest like temperature, humidity, pressure, seismic wave, soil moisture etc. Energy conservation is one of the major challenges of wireless sensor nodes. The sensor nodes are tiny devices which operate on very low battery supply and are deployed for sensing some physical parameter of interest.WSN discussed in this paper typically composed of number of sensor nodes deployed at a geographical location, a sink node or the base station [7]. Data are sensed by sensor nodes are transferred to the base station through intermediate nodes using multi-hop communication. Both the sink and sensor nodes are considered to be static (i.e. mobility of sensor node in not accounted). The architecture of WSN discussed is Fig 1. As sensor nodes are primarily operated on a limited battery power, the major concern is to reduce overall energy consumption. Typical WSN applications require dense sensor deployment in order to achieve satisfactory coverage. As a result, multiple sensors record information about a single event in the sensor field. Due to high density in the network topology, spatially proximal sensors observations are highly correlated with the degree of correlation increasing with decreasing inter node separation [1][6].

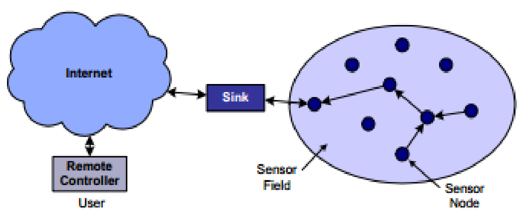

Fig 1: Sensor Node Architecture [2]

Although various media access control(MAC) and routing protocols, sleep management schemes, power management schemes has been proposed for energy conservation, no promising techniques has been proposed in the direction of energy efficient data acquisition. There are various data-driven techniques that are implemented to conserve energy – like in-network processing and data compression techniques. In-network processing mainly focuses on performing data aggregation at intermediate nodes in order to reduce communication. Data compression technique is used in order to reduce the size of data to be sent between sensor nodes[2] This paper explores data-aware energy conservation scheme which primarily reduces energy consumption by preventing redundant communication which is identified applying statistical analysis of datasets. The statistical analysis can be conducted to find out the inter-node data similarity metrics like – Euclidean Distance, Cosine similarity and Pearson correlation coefficient. These metrics helps to identify the degree of association between datasets [6]. Once high correlation between inter-node datasets is found, the redundant sensor node is switched to sleep mode in order to conserve energy efficiently [4]. Section II of this paper discusses various data conservation schemes, Section III if focused on survey of statistical tools for finding the degree of similarity between sensor nodes, Section IV discusses about Labelled Sensor Data and Section V about the conclusion of this paper.

II. DATA-DRIVEN TECHNIQUES

This section discusses about data-driven techniques used in order to conserve energy. Experimental measurements have shown that generally data transmission is very expensive in terms of energy consumption while data processing consumes significantly less [3]. The most efficient energy conserving operation is putting the radio transceiver in the low power or sleep mode whenever communication is not required. This behavior is usually referred to duty cycling and duty cycle is defined as the fraction of time nodes are active during lifetime.

The sleep wakeup frequency of sensor nodes should be coordinated in WSN. A sleep/wakeup scheduling algorithm thus accompanies any duty cycling scheme. Duty cycling schemes are generally unaware to data that are sampled by sensor node. Moreover it is observed that sampled data generally have strong spatial and temporal correlation [1].

The spatio-temporal association in inter-node sensor data is an important characteristic of WSN that can be exploited to improve the overall energy efficiency of WSN as well as the application associated with it. The spatial relation attribute causes WSN to behave in a collaborative manner which can be used for energy conservation and performance improvement [6]. This advantage can be used in development of efficient data-driven algorithm intended to improve energy efficiency even more.

The data driven energy conservation schemes is broadly classified as: In-network processing, Data Compression, Data Prediction and Energy-efficient data acquisition techniques (in Fig 2). Data Compression and In-network processing is most widely used approaches in this regard. Not much attention is given to energy-efficient data acquisition and data prediction techniques in order to conserve energy. Data compression reduces the amount of information sent by source node to sink node. In-network processing technique consists of data aggregation (sum,average,etc). Data prediction consists of describing pattern and predicting the consequent data. Energy efficient data acquisition schemes are mainly aimed at reducing the energy spent by the sensing subsystem.

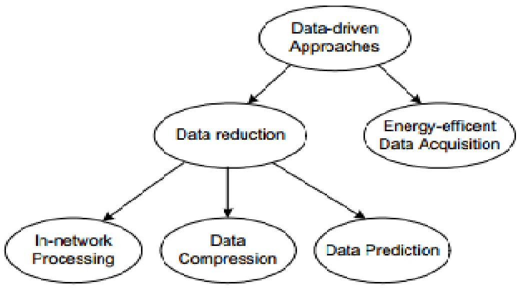

Fig 2. Data Driven Approaches [2]

III. SIMILARITY METRICS

In this section the theoretical methods to calculate inter-node data association is discussed.[**6**] Let $x[i]$ be the datasets from sensor node X and $y[i]$ be the datasets from sensor node *Y*. Let *n* be total number of samples used for analysis. The commonalities between data source due to spatial and temporal relation can be calculated using following similarity metrics:

1) *Euclidean Score:* Euclidean score is a method of calculating a score of how similar two things are. We get a value between 0 and 1, 1 meaning they are identical 0 meaning they don't have anything in common. Euclidean score is calculated as:

$$Euclidean\,Score = \frac{1}{(1 + Euclidean\,Distance)} \qquad (1)$$

Euclidean distance is the distance between two points in Euclidean space. Euclidean space was originally devised by the Greek mathematician Euclid around 300 B.C.E. to study the relationships between angles and distances. The Euclidean Distance is calculated using following formula:

$$Euclidean\,Distance = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \qquad (2)$$

2) *Cosine Similarity:* The cosine similarity between two nodes is given by:

$$Cosine\ Similarity(\cos\theta) = \frac{\sum_{i=1}^{n} x_i \times y_i}{\sqrt{\sum_{i=1}^{n} {x_i}^2} \times \sqrt{\sum_{i=1}^{n} {y_i}^2}} \quad (3)$$

The cosine similarity is in the range of $[-1,1]$, the similarity value of -1 means that data are exactly opposite, 0 meaning independent,1 meaning exactly the same, with in-between values indicating intermediate similarities or dissimilarities. Cosine similarity gives a measure of common variations observed between the samples of the nodes by evaluating the angles between the two data samples considered.

3) *Pearson Correlation Coefficient:* Pearson Correlation Coefficient*(r)* is known as Pearson Product-Moment Correlation Coefficient (PPMC). The Pearson product-moment correlation coefficient is a measure of the linear correlation between two variables X and Y, giving a value between +1 and −1 inclusive, where 1 is total positive correlation, 0 is no correlation, and −1 is total negative correlation. As the value of *r tends to 0,* the greater is the variation. The formula for calculating PPMC is given by:

$$PPMC(r) = \frac{n\ \sum xy - \sum x\ \sum y}{[\sqrt{n\ \sum x^2 - (\sum x)^2}]\ [\sqrt{n\ \sum y^2 - (\sum y)^2}]} \quad (4)$$

IV. LABELLED SENSOR DATA

Based on the discussion in Section III, the algorithm shall be designed in order to conserve energy by identifying redundant communication and by using data prediction algorithm. In order to achieve this simulation shall be done on real time data generated by WSN. The datasets for the simulation can be taken from the Labelled Wireless Sensor Network Data Repository (LWSDR) of The University of North California (GREENBORO) [5]. The datasets from LWSDR is a Multihop labelled temperature and humidity data generated by 4 sensor nodes. The hardware platform that we used for data collection was the TinyOS-based motes, more specifically, Crossbow TelosB motes. As shown in Fig. 3, there is a base station node with ID 0 as the root of the tree and two router nodes with the IDs of, respectively, 10 and 20. Also, there are two indoor sensor nodes with respective ID numbers of 1 and 2 and two outdoor sensor nodes (3 and 4). Each sensing node has a reduced transmission range and cannot reach the base station node except through the router node. The sampling interval is 5 seconds and each packet has ten readings each of temperature and humidity. The time period for collecting the readings is 6 hours, during which anomalies were introduced to one sensor node in each scenario (nodes 1 and 3) by using a water kettle which altered the temperature and humidity appropriately.

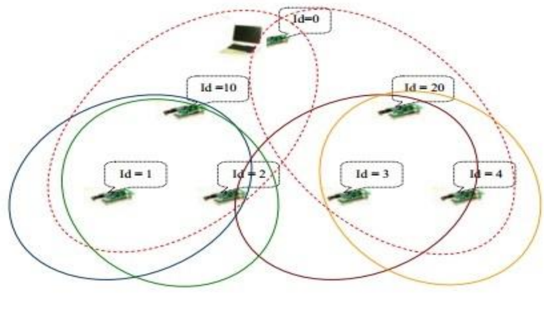

Fig.3 Multi-hop data collection distributed system [5]

V. CONCLUSION

There are various energy conservation schemes in Wireless Sensor Network. Data-driven schemes especially efficient data acquisition and data prediction schemes are not widely used. The spatio-temporal association can be exploited to propose energy efficient data-aware protocols. As sensor node deployments become dense, the sensed data becomes closely associated. As per the discussion presented in this paper it indicates that inter-node data association metrics can be used to identify degree of association thereby reducing energy consumption. The similarity metrics can be used to identify redundant nodes from the base station for the set of data discussed in Section IV. The sensor nodes with redundant samples can be switched to low energy consumption mode for energy conservation purposes.

REFERENCES

[1] Ali, Bakhtiar Qutub, Niki Pissinou, and Kia Makki. "Identification and Validation of Spatio-Temporal Associations in Wireless Sensor Networks." Sensor Technologies and Applications, 2009. SENSORCOMM'09. Third International Conference on. IEEE, 2009.J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[2] Anastasi, Giuseppe, et al. "Energy conservation in wireless sensor networks: A survey." Ad Hoc Networks 7.3 (2009): 537-568.K. Elissa, "Title of paper if known," unpublished.

[3] G. Pottie, W. Kaiser, "Wireless Integrated Network Sensors, Communication of ACM, Vol. 43, N. 5, pp. 51-58, May 2000.

[4] Raja Jurda,Antonio G. Ruzzelliz and G.M.P. O'Hare "Radio Sleep Mode Optimization in Wireless Sensor Networks".

[5] Suthaharan, Shan, et al. "Labelled data collection for anomaly detection in wireless sensor networks." Intelligent sensors, sensor networks and information processing (ISSNIP), 2010 sixth international conference on. IEEE, 2010.

[6] Vuran, Mehmet C., Özgür B. Akan, and Ian F. Akyildiz. "Spatio-temporal correlation: theory and applications for wireless sensor networks." Computer Networks 45.3 (2004): 245-259.

[7] Waltenegus Dargie and Christian Poellabauer, "Motivation for a Network of Wireless Sensor Nodes", Chapter 1 in book Fundamentals of Wireless Sensor Networks, Pp. 3-14, Wiley, 2010

[8] Ye, Wei, John Heidemann, and Deborah Estrin. "An energy-efficient MAC protocol for wireless sensor networks." In INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE, vol. 3, pp. 1567-1576. IEEE, 2002