Microprocessors and Microsystems 39 (2015) 348–357

Contents lists available at ScienceDirect

Microprocessors and Microsystems

journal homepage: www.elsevier.com/locate/micpro

Energy conservation in WSN through multilevel data reduction scheme

Muruganantham Arunraja *, Veluchamy Malathi, Erulappan Sakthivel

Department of Electrical and Electronics Engg, Anna University, Regional Centre, Madurai

ARTICLE INFO

Article history:
Available online 11 June 2015

Keywords:
Wireless sensor network
Data reduction
Data aware clustering
Dual prediction
Data aggregation

ABSTRACT

Lifetime is one of the major Quality of Service factors for Wireless Sensor Networks (WSN). As sensor nodes are generally battery-powered devices, the network lifetime can be extended over a reasonable time span by lessening the energy consumption of nodes. Reducing the amount of data transmission can effectively minimize the energy consumption, the bandwidth requirement and network congestions. In a WSN, denser deployment of nodes results in a high spatial correlation between data generated by neighboring nodes. Slow varying nature of many physical phenomenon results in similar sensor observations over the period. In this proposed work, a two level data reduction technique is employed. Here the Data and Energy Aware Passive (DEAP) clustering approach is introduced to divide the sensor network into data similar clusters. A Dual Prediction (DP) based reporting is deployed between cluster members and their Cluster Head (CH). This first level data reduction is attributed to the temporal correlation of data over the time. In CHs, the data from multiple data similar nodes are aggregated to reduce the spatial data redundancy. The proposed method DEAP-DP is verified with real world datum and has achieved up to 68% data reduction at 0.5 °C error tolerance.

1. Introduction

Wireless sensor Nodes cover up significant portions of the earth's surface. Typically a WSN is a collection of spatially distributed autonomous sensor nodes deployed for monitoring physical parameters like temperature, pressure, humidity, sound, vibration and seismic events [1]. The monitored parameters are reported to the Base Station (BS) through the networked architecture. A sensor node is a miniature device with some essential components like (a) Sensing unit – for acquiring data from the surrounding environment. (b) Processor – for data processing and for a transitory storage purpose. (c) RF transceiver-for transmitting the processed data [2]. Of which the RF transceiver is the major consumer of nodal energy. WSNs are preferred over a wide variety of applications due to their small form factor and mobility compatible nature. Thus, it is impractical to have a large sized power source. This power source often consists of a battery with a limited energy budget, which results in finite lifetime of nodes. In addition, it could be impossible or inconvenient to recharge the battery, because nodes may be deployed in a hostile or unpractical environment.

In the era of IoT [3], the sensor nodes are feeding the data hungry internet servers. Thus a longer life span and optimal quality of collected data are essential. On one hand, the sensor network should have a lifetime long enough to fulfill the application requirements. On the other hand, data accuracy and latency are the major issues of WSN. Thus the data gathering process should conserve the residual energy of the node while maintaining optimal data accuracy. Hence, by exploiting the inherent properties of WSN data, energy can be conserved with minimum tolerance on the collected data accuracy. Tobler's first law of geography [4] states that "Everything is related to everything else, but near things are more related than distant things". This statistical observation implies that data correlation increases with decreasing spatial separation of sensor nodes [5]. In a WSN, the nodes are densely placed and transmit data at high sampling rates, which results in a high level of spatial and temporal correlations; hence a vast amount of data transmitted is redundant. By reducing this redundant or similar data, a significant amount of nodal energy can be conserved along with bandwidth savings.

In a WSN, both the network structure and data collection methodology decide the energy expenditure. Hierarchical structures are preferred over plain network, due to their improved reliability and energy conservation. Clustering is the prominent hierarchical architecture. By creating data similar clusters spatially redundant data can be excluded. Higher sampling rates can achieve accurate and microscopic data collection. But not all the sampled data are worth transmission. By avoiding unnecessary data transmission, huge energy savings can be realized.

* Corresponding author.
E-mail addresses: arunraja98@gmail.com (M. Arunraja), vmeee@autmdu.ac.in (V. Malathi), vlsisakthivel@gmail.com (E. Sakthivel).

http://dx.doi.org/10.1016/j.micpro.2015.05.019
0141-9331/

The major contributions of the proposed work are defining an energy efficient intra cluster data communication through adaptive nLMS based dual prediction and the construction of data similar clusters with minimal communication overhead. The proposed work constructs data similar clusters using passive clustering approach. Here the high energy nodes elect themselves as CH after an energy driven proclamation delay. The proclamation includes the features of recent data history. Based on the proclaimed information, the neighboring nodes estimate the data similarity with the proclaimed node and associate with the data similar CH. Using an adaptive nLMS based filter, the cluster members communicate their data to the CH. At the CH, spatially redundant data are aggregated and transmitted to the BS. The proposed work is intended to conserve the energy of continuous data collection systems without significant loss in data accuracy. As most of the systems are densely deployed and the phenomenon under measurement has the slow varying nature, the proposed work can achieve significant drop in nodal energy expenditure. The energy saving increases with increased temporal error tolerance.

The rest of the paper is organized as follows. Section 2 discusses the related work. Section 3 explains the proposed methodology. Experimental setup is explained in Section 4. Simulation results are discussed in Section 5. Finally, Section 6 concludes our work.

2. Related works

The problem of WSN lifetime maximization, in general, has been addressed in several literatures. Anastasi et al. [2] listed multiple approaches for saving energy. Since radio communication is the governing power consuming activity in WSN, by minimizing the amount of data transmitted, a noteworthy level of energy conservation can be achieved. We divide the previous literature into the two main categories. In the first category, we discussed various energy aware clustering methods and their issues. In the second category, we discussed different energy efficient data collection methods and highlighted the importance of the dual prediction.

Clustering is the most significant and widely used hierarchical routing algorithms. Based on the parameters used for CH election and cluster formation criteria, the clustering algorithms can be classified as Probabilistic and Non Probabilistic Clustering Algorithms. To decide the initial CH, each sensor node is provided with a node ID or some probability value in the probabilistic clustering algorithms. Better network lifetime and energy consumption can be achieved by determining the appropriate probability value for clustering the sensor nodes. LEACH [6] is a prominent probabilistic clustering algorithm, where CH is elected based on a random probability and the neighboring nodes join with the nearest CH. HEED [7] is also a probabilistic Clustering Algorithm that uses distributed scheme for cluster formation. Here CHs are selected based on node‘s residual energy, node degree or node’s proximity to its neighbors. EECS [8] is a distributed clustering method similar to LEACH with certain improvement in CH section and cluster formation. Unlike HEED, in EECS single-hop communication pattern is retained between the CHs and BS.

In case of non-probabilistic Clustering Algorithms, the criteria for CH election and cluster formation are based on the node‘s nearness parameters like degree, distance, and connectivity. Here the cluster formation requires extensive communication between the nodes and their neighbors. Highest Connectivity cluster [9] is a non-probabilistic method that proceeds in two levels namely tree discovery and cluster formation. The number of neighbor nodes is broadcasted by the sensor node to consider the connectivity. The sensor with highest connectivity that is the maximum number of one-hop neighbors is elected as CH, whereas the clustering process can be started by any node in the WSN. Biologically Inspired Clustering Algorithm [10], uses swarm agents in a stochastic algorithm to obtain near optimal solutions to complex, nonlinear optimization problems. Separation and alignment swarms are used to achieve a stable and near uniform distribution of the CH nodes. Weight-Based Clustering Algorithm [11] is a non-probabilistic clustering algorithm that elects the CH non-periodically and utilizes distributed scheme for cluster formation. In order to save power, new CH is elected when the sensor loses the connection with its CH. The parameter used for election procedure is the ideal node degree, transmission power, mobility and the remaining energy of the nodes.

The probabilistic clustering methods have the advantages of low overhead clustering and faster convergence time, but suffer from inefficient cluster formation and reduced cluster reliability. The non-probabilistic clustering methods achieve highly reliable and well balanced cluster, but the control message overhead is high and long convergence time. The proposed approach is a non-probabilistic clustering method with minimal control overhead, as it uses a passive clustering technique. The convergence time is also reduced with the help of energy aware back off delay. Hence the proposed method is reliable and highly energy balancing.

There is a new genre of clustering algorithms that considers the data similarity between the neighboring nodes. CAG [5] utilizes the correlation to form clusters in a tree-based structure sensor network. Since the clusters are data similar, only CHs data represent the entire cluster. Since the adjacent nodes are not aware of each other’s status, there is a high probability of redundant cluster formation. In CAG during the cluster adjustment phase, the existing clusters can only be bifurcated and cannot be merged. This issue leads to the formation of clusters with one member, so after a while all sensor nodes will become CHs. To cope with this, DACA [12] proposes a data and energy aware algorithm, where the clusters can be merged. Still, in both the methods, the error tolerance between the CH and its members cannot be guaranteed. The DACA includes the node degree factor for CH election and hence increases the control overhead. The proposed DEAP uses the passive clustering through which the clustering overheads are reduced significantly. Being a deterministic approach the method distributes the clusters more efficiently. The data aware nature of proposed clustering helps in achieving highly efficient data aggregation at the CH.

Among the three major energy conservation techniques, both the mobility and duty cycling methods are unaware of the inherent data redundancy, thus data driven methods are more preferable for data correlated networks. Based on problems and applications, Data-driven methods [13] are classified as Data compression, in-network processing and data prediction. The data-driven approaches aim at reducing the energy spent by the communication subsystem and the amount of data to be delivered to the sink node. Since the data generated by a WSN has a high level of spatial and temporal similarity, the method that exploits these similarities can serve better towards effective energy conservation without affecting the accuracy of the collected data.

Xiong et al. [14] use distributed source coding to encode the correlated data independently at the source nodes and decode it jointly at the destination. But it requires prior knowledge on precise correlation of data. Compressed sensing [15,16] methods can compress the data using a small number of non-adaptive, randomized linear projection samples and recover the same without any loss, provided the data has sufficient sparsity. The assumption of constant sparsity in the sensed data in conventional CS is unrealistic for most natural signals, which may cause reconstruction quality degradation.

To reduce data while traversing the network towards the sink, In-network processing [17] method performs data aggregation