Communication Protocol for Wireless Sensor Networks for Energy Consumption Optimization

Design and Simulations

Marius Popescu, Lăcrimioara Grama, Corneliu Rusu
Basis of Electronics Department
TUCN, FETIT, Signal Processing Group
Cluj-Napoca, Romania

Marius Sîrbu
Sc WsComm SRL
Cluj-Napoca, Romania

***Abstract*—The goal of this paper is to develop a communication protocol for wireless sensor networks in order to optimize the energy consumption. The main idea is to eliminate the redundancy of data transmission from sensors to base station. Innovation lies in the approach used to achieve the goals. A probabilistic model based on data correlation is generated at the base station. The new values that will be measured by each node are computed using the probabilistic model; these values are sent to nodes. If the measured value is different from the received one, the difference is send to base station; otherwise the node does not send anything. The architecture of the proposed protocol is presented and also the protocol is validated through simulations. (*Abstract*)**

Keywords—wireless sensor; communication protocol; energy consumption optimization (key words)

I. INTRODUCTION

The aim of this paper is to develop a communication protocol for wireless sensor networks (WSN) to optimize the network energy consumption, by eliminating the redundancy of data transmission from sensors to base station (BS). Nowadays, the WSN has reached a degree of maturity; there are market solutions in terms of hardware and software architecture and in terms of specific communication protocols [1]. This field still provides many problems that need to be solved; one of the open issues is to minimize energy consumption in devices. A possible solution to energy consumption problem is to reduce the amount of the transferred redundant information. For this, in the literature there are two basic approaches for WSN: network coding/data aggregating [2] and distributed compressive sampling [3]. One issue raised by eliminating redundancy techniques mentioned is the optimal segmentation network to obtain the best possible compression.

Proposed protocol aims the energy optimization based on correlation using a different approach. To reduce energy consumption other techniques can be identified, such as improving software architecture [4] or use routing algorithms based on a metric dependent on power consumption [5].

The paper is organized as follows: Section II provides a brief introduction to WSN; Section III is focused on protocol design, starting from assumptions up to specifications. Based on mathematical model some algorithms are proposed to achieve the objectives. In Section IV the application and the experimental results are illustrated. Section V is dedicated to conclusions and future work.

II. WIRELESS SENSOR NETWORKS

We propose a communication protocol for WSN designed for an efficient management of energy. In many situations, the nodes of such a network are powered using chargeable cells, thus energy is a critical resource. If we add the fact that the network nodes are placed in areas difficult to reach (mountains, mining galleries, underwater environments, etc.) or are spread on large areas (network monitoring agricultural areas or those for surveillance of protected areas; in the latter case is imposed also a limited human intervention) it is easy to justify the need for such an approach.

Protocols used in other types of communication networks disregard the power consumption in the transmission and processing of data (there are other priorities: the bandwidth used, network and data security ability), the interconnected equipment being not the subject of restrictions related to energy because they are powered directly from the power distribution or their recharge do not raise any special problems. Hereinafter we will present some issues regarding the WSN, the main purpose being to establish the goal of the protocol.

A generic WSN is a set of equipment for acquisition of interest data through some sensors and carry them via radio waves to a central device where they are stored, processed and passed on. The modules directly responsible for data acquisition, primary processing and transfer to other devices are the *nodes*. Nodes have in their structure: sensors, data processing unit, memory, radio transceiver and power source[6]. The BS can be a computer able to communicate with nodes via a radio module, performed at a storage database processing, for obtaining information and data sharing through Internet; it is usually used for network administration, allowing control and configuration information transmission.

III. PROTOCOL DESIGN

The protocol is designed to be implemented with minimal changes to hardware and software architecture, typical for this kind of applications. Related to the network components the following assumptions are done: the process is a distributed one; it has a certain correlation degree in time and space; only the network nodes have energy restrictions. The main objectives are: knowledge of the state process with a certain resolution in time and space; minimizing energy consumption and maximizing the lifetime of the network; use of simple sensory nodes architecture with reduce processing capabilities.

Data structures associated to the protocol which are transmitted through the network are called data frames. Traffic data frames are transmitted periodically, during the entire network lifetime, their main purpose being to carry useful data. The BS acquaints with each node in two phases: it transmits the predicted value; if the predicted value differs from the measured one, through the BS is transmitted the difference between the predicted and correct value, otherwise nothing is transmitted.

The model used for prediction is a simple probabilistic model .For each node a prediction function depending on the previous measured values is defined . This function will take as input the values measured in the network in the last epoch and will output the predicted value for the current node and epoch.

This function will be approximated using data collected from the network. The simplest approximation scheme consist in a discretization of the definition domain and for each resulting “cell” the function will output the most frequent value in the available data set .The output domain is also discretized.

For example if we have a network with 5 nodes designed for temperature measurement in a house, in range -20 … 40 Celsius degrees the input values will be the 5 values measured in the previous epoch. Discretization can be made with a resolution of 1 degree. If the last measured temperatures are 10.1 , 12.3,12.6,12.22 and 11.99 then the cell will be [10,11]*[12,13]*[12,13]*[12,13]*[11,12] and if for this cell we have the following collected data (representing the next measured values for this node when the previous value are in this cell) : 12,12,10,10,11,12,12 ,the predicted value will be the value in the data set with the highest frequency ,therefore 12 (all measurements are in Celsius degrees).

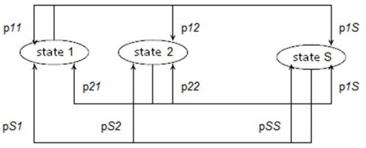
IV. SIMULATION AND EXPERIMENTAL RESULTS

A. Application

A MatLAB application was developed to prove the efficiency of the algorithm. The evaluation criteria of the protocol are: energy savings using prediction and impact on the network lifetime by optimizing the transmission scheme. Each *N*-dimension random vector (from a multivariate Gaussian distribution) represents the set of values purchased by nodes in an epoch. Covariance matrix elements are in inverse ratio to the distance between nodes; averages vector is initialized with arbitrary values; the vector obtained after each iteration is assigned to simulate temporal correlation. The application consists of two basic modules: training/adaptation module and measure and comparing the energy consumption module. For training, vectors using the same distribution are generated (the logic scheme is as in Fig. 2). Fig. 3 presents the algorithm which assesses the energetic performance of the protocol relative to another generic protocol (same architecture and application, with continues transmission at the same intervals, but which does not use the prediction mechanism).

B. Considerations on the Implementation

To easy introduce the simulation data and for rapid results visualization the MatLAB GUI was used. The user can choose between simulations with/without training phase. The first situation is closer to reality. In general no data to be used for training before commissioning of the network are available.



```
define node.slot
while (network.function)
    I={1 … S}
    for k=1 … N
        val.predict = arg max M(I, j)
        tx(val.predict)
        wait_answer
        if (error.predict)
            slot(k)=val.rx
        else
            slot(k)=val.predict
        end.if
        I={I ∈I/ node.slot(k)=slot(k)}
    end.for
    state.prevoius=state.current
end.while
```

Fig. 1. Markov network and prediction algorithm

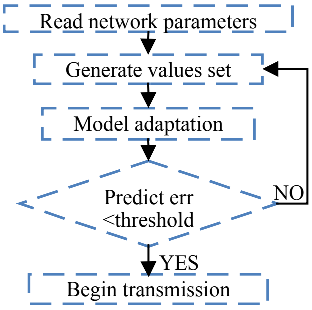


Fig. 2. Training algorithm (logic scheme)

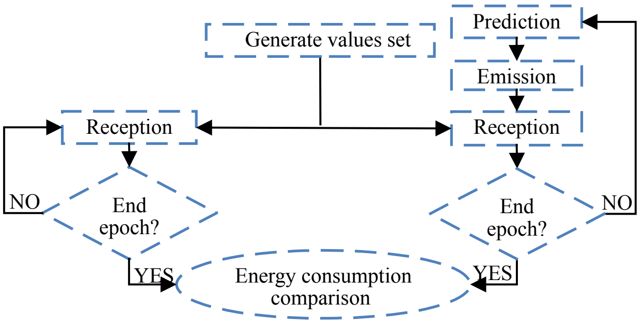


Fig. 3. Comparative energy consumption analysis algorithm (logic scheme)

In order to obtain conclusive results several steps of simulation are needed to be applied, the results of early epochs being irrelevant. The next step is to determine network characteristics: number of nodes and their location, BS coordinates, number of values that can be measured by nodes and maximum values of size of interest. Position of nodes is established in a 2D system (easy visualization of nodes spatial distributio. By marking training after setting network parameters, *Simulation* and *Training* buttons are available; otherwise there will be only *Simulation*. If one does not want to use default parameters (training/simulation epochs=100), they can be changed. Pressing *Training*, a Markov model based extract data, from a multivariate Gaussian distribution, will be generated. The steps involved for an effective simulation are:

A. If training phase is skipped a prediction model with some default parameters is used, otherwise training data is generated and with this data the training algorithm is executed.

B.For each transmission epoch the following steps are done:

1. Generate a data set from the same probability distribution used in training phase; the number of values equals the number of nodes, signifying measurements made in the current epoch;
2. Using described prediction algorithm the most likely value for each node is determined;
3. If the predicted value equals the one "measured" by the node then the energy consumption in this step is 0, otherwise the node will have to transmit the difference between measured and predicted values to BS therefore the energy consumption will be considered proportional to squared distance from node to BS ;

C.The final result is given as a percentage: for each node energy consumption is plotted in the case of described protocol, and the energy that would be consumed if it would be a simple transmission of measured values (without prediction).

An additional feature is the ability to see the correlation in time and space of generated data or simulation training phases. Correlation and covariance between values of input nodes and between values from different simulation/training epochs are available. A comparison between prediction error rate when the absolute value is chosen randomly and when the value is determined by applying the prediction algorithm is also shown.

C. Simulation Results

In order for the data to be valid, the assumptions mention in the design step must be satisfied: there must be a double dependency between data purchased by nodes, in space and time. The dependence in space refers to the correlation between measured values of nodes during an epoch. To highlight this, the significant parameters that must be set are the distances between nodes and the number of simulation/training epochs. The results will be some statistical data properties: correlation and covariance. In Fig. 4 the scenario used for nodes placement is presented and also the correlation of measured values (the number of training/simulation epochs is 100).

The dependency in time consists in the influence exercised by the values from a certain epoch on the values from other epochs. Increasing the time interval between two acquisitions, the correlation between values sets is increased. To highlight this, the significant parameter that must be set is the number of nodes. Some of the obtained results are illustrated in Fig. 5 (100 nodes). Statistical parameters obtained through simulations reveal the existence of time and space correlation between generated data, according to design assumptions. For energy consumption evaluation and comparison with a generic protocol we have used for exemplification 5 nodes, placed as in Fig. 6 (left). The energy consumption is illustrated in Fig. 6 (right): by red in the case of a generic protocol, and by blue using the described protocol. We can conclude that if the data acquired by the sensors meet the requirements, the designed protocol can achieve substantial energy savings.

V. CONCLUSIONS AND FUTURE WORK

In the theoretical part of the paper a protocol was designed according to the requirements stated.

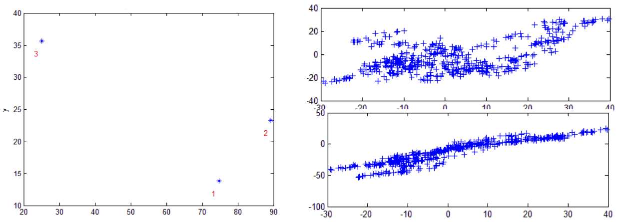


Fig. 4. Position of nodes (left) and the correlation of the measured values in nodes: 1 and 2 (upper right); 1 and 3(lower right)

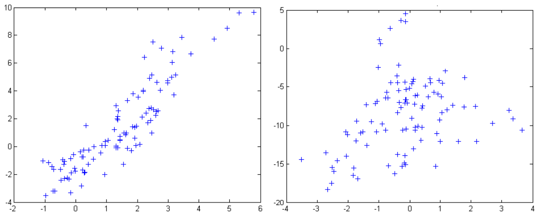


Fig. 5. Correlation of values in epochs: 1 and 2 (left), 1 and 10 (right)

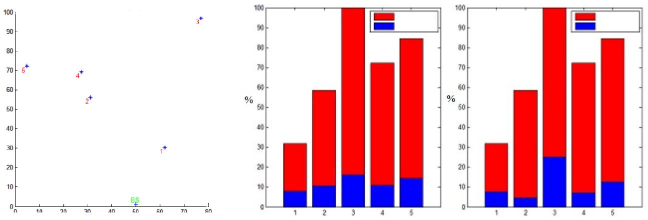


Fig. 6. Position of nodes (left) and energy consumption (right -- reference 100% is the maximum energy consumption)

Based on simulations, the protocol viability was verified. Using data generated artificially it was shown that the protocol is able to achieve considerable energy savings relative to another protocol which does not exploit time and space data correlation. As future work we can consider other predictive methods (regression or neural networks). The application can be extended to compare the performance of proposed protocol with other protocols having same objective. Also, for a full confirmation of these capabilities, testing in a real network protocol can be done.

REFERENCES

[1] ZigBee Specifications, http://www.zigbee.org/Specifications.aspx, 2012.

[2] N. S. Patil and P. R. Patil, "Data aggregation in wireless sensor network," *IEEE International Conference on Computational Intelligence and Computing Research*, 2010.

[3] Libelium Comunicaciones Distribuidas S.L., "Waspmote datasheet," 2012.

[4] D. Pfisterer, M. Wegner, H. Hellbruck, C. Werner and S. Fischer, "Energy-optimized data for heterogenous WSNs using middleware synthesis," *IEEE International Conference on Computational Intelligence and Computing Research*, 2010.

[5] S. Bilouhan and R. Gupta, "Optimization of power consumption in wireless sensor networks," *International Journal of Scientific and Engineering Research*, vol. 2, issue 5, may 2011.

[6] F. Zhao and L. Giubas, "Wireless sensor networks: an information processing approach," *The Morgan Kaufmann*, 2004.