Novel Load Balancing Scheduling Algorithms for Wireless Sensor Networks

Endre László, Kálmán Tornai, Gergely Treplán
Faculty of Information Technology
Pazmany Peter Catholic University
Budapest, Hungary
lasen@digitus.itk.ppke.hu, tornai.kalman@itk.ppke.hu, trege@itk.ppke.hu

János Levendovszky
Department of Telecommunications
Budapest University of Technology and Economics
Budapest, Hungary
levendov@hit.bme.hu

Abstract—In this paper, optimal scheduling mechanisms are developed for packet forwarding in wireless sensor networks, where clusterheads are gathering information. The objective is to monitor real-life processes for a given time interval and forward packets with minimum loss probabilities to the base station. In order to achieve this objective we develop an optimal scheduling algorithm, which determines the time slots in which packets must be sent by the nodes. The scheduling algorithm, on the one hand, guarantees that all the packets will be sent within a predefined time window and thus meeting delay constraints and, on the other hand, it provides uniform packet loss probabilities for all the nodes. The algorithm we propose is capable of providing optimal scheduling with given constraints and guarantees balanced load in polynomial time.

Index Terms—load balancing, scheduling, wireless sensor networks (WSN), numerical optimization

I. INTRODUCTION

Data gathering from a set of sensor nodes to a Base Station (BS) by using a cluster-based routing topology is commonly used in wireless sensor networks (WSNs) [1], [2]. In this kind of networks tiny sensor nodes communicate in short distances and collaboratively work to fulfill the application specific objectives of WSN. Many of the envisioned applications involve the collection of bursty data traffic generated by events, which are to be delivered to the BS as quickly and as reliably as possible in order to recognize emergency situation. In these applications packet delay and packet loss probability are of crucial importance [3], [4].

Resources (energy, time, bandwidth) are limited in WSN applications, hence cross-layer optimization are the key approach to minimize the utilization of resources if the Quality of Service (QoS) is prescribed.

Because of these strong limitations, scheduling becomes even more important in WSN to save on energy consumption. In this paper we develop a new scheduling algorithm for packet transfer in WSN, which guarantees reliable information transmission to the BS in terms minimizing the packet loss probabilities. Packet loss occurs when the amount of packets forwarded to the CHs exceed their capacity. Scheduling has been intensively researched in the telecommunication literature [5], [6], [7], however the main focus was on buffered architectures. The problem in our case will be transformed into a binary matrix optimization. So far, this has been solved by quadratic programming (see [8]). Now a novel solution is presented, which will provide the optimal scheduling matrix in polynomial time. Since clusterhead (CH) based routing is a commonly used solution in WSN (e.g., LEACH protocol [9] or other hierarchical solutions proposed in [10], [11]), we assume that each node can only send one packet to a selected CH at each time instant.

After the system and the application are briefly introduced, the network model and the mathematical model are detailed. The formulation and transformation of the original problem into a matrix optimization task is followed by the performance analysis and the conclusion of our work.

II. SYSTEM DESCRIPTION AND APPLICATION

The algorithm introduced in this paper is concerned with applications when the traffic generated by the sensor nodes are classified into traffic classes (e.g., motion detection, acoustic signals, video signals, ... etc.) The WSN architecture is single-tier clustered with heterogeneous sensors and performing centralized processing [12]. The CHs forward the received packets to the BS over a link with capacity \( V \). As a result \( V \) presents a limit to the packets can be collected from nodes in a given time instant. In order to collect all the packets the CH assigns timeslots to the nodes for acquiring their packets. This timeslot assignment is referred as a scheduling. In this scheme, each traffic class has two parameters:

- the amount of packets to be transmitted to the CH;
- the time duration in which this given amount of packets must arrive at the CH.

The target platform is the Berkeley Mica2 mote [13], which is one of the most widely used WSN platforms. The platform has an 8 MHz processor, 4 kB of RAM, 128 kB of flash memory and a 433 MHz wireless radio transceiver. The transfer rate is 38.4 kbps and it is powered by two AA batteries.

TinyOS operating system [14], [15] is designed to be used with networked sensors and it supports the Mica2 platform. TinyOS handles task scheduling, radio communication, clock and timers, ADC and power management.
In the case of Mica2, a fully programmable Media Access Control (MAC) layer is available. A time division multiple access (TDMA) channelization is supported in the case of such simple and energy constrained nodes. Plenty of channel access methods are proposed for WSN and there are two common solutions. One solution is the contention based scheme where nodes try to assign the channel randomly, independently from each other, such as ALOHA [16]. ALOHA is improved in many ways (such as B-MAC [17], Z-MAC [18], X-MAC [19]) to guarantee performance metrics required by different applications. Contention based protocols are very important channel access methods in the case of rare but bursty traffic, which is the typical case of event driven applications. On the other hand TDMA-based scheduled solutions could be very efficient if the users are known and fix, the data arrives regularly (e.g.: every one minutes the temperature value must be delivered to the BS) and topology is static. TDMA-based protocol can also save more energy, since each node can stay in sleep mode except for its own slot time. However TDMA-based protocols need synchronization protocols [20] affecting the guard time [21], time slot assignment [22] determining the end to end delay and scheduling. Scheduling protocols choose the order of the data transfer at the different links [23]. In this paper we propose Load Balancing Scheduling (LBS) algorithm, which minimizes the probability of cell loss.

In the following section the mathematical model will be introduced.

III. THE MODEL

The overall number of packets node $j$ wants to send to the CH is denoted by $X_j$. All of the packets must be sent within a time interval $K_j$ : $K_j > X_j$. The scheduling of packet transmissions by node $j$ can be expressed by a binary vector $e(j)$ of length $K_j$ with weight $X_j$, where component $c_j(l) \in \{0, 1\}$ indicates whether packet is transmitted or not to the CH at time instance $l$.

Composing a transmission matrix $C$ of these vectors, one may seek the optimal matrix ($C_{opt}$). To find the optimal matrix, we present a novel approach based on a constructive iteration, solving the problem in polynomial time.

The data acquisition network can be divided into two parts:

- The sensors performing measurements in the different rooms.
- Relay nodes, which are forwarding the packets containing the measurements to the BS.

In Figure 1 the different application classes correspond to different rooms in a building. Each room has different $X_j$ i.e., one requests a given amount of data from every room. The data is collected by the sensor node from its sensors and stored for transmission. In the beginning of the communication $X_j$-s are transferred to the sensors nodes, where the calculation of $K_j$-s are performed based on the battery status and returned to the CH. The optimal scheduling matrix is then constructed by the CH. In each time slot, $V$ amount of data is sent to the CH corresponding to the $l$-th column of the scheduling matrix.

This $V$ amount of data is then sent to the relay nodes of the network in order to forward them to the BS.

The nodes performing the measurements are small size units with a variety of sensors included with strong energy and processing limitations. Energy is used either in stand-by mode or in active mode when they are engaged with radio transmission (packet forwarding). The available energy at node $j$ is denoted by $G_j$, while the energy consumption in active node is $g_{tx}$ and in stand-by is $g_{sleep}$. If the node sends $X_j$ amount of packets and it switches off after $K_j$ timeslot, then the overall energy consumption can be given as

$$G_j = (K_j - X_j) \cdot g_{sleep} + X_j \cdot g_{tx}.$$  \hspace{1cm} (1)

As $G_j$ is limited (due to the battery capacity) based on this expression one can determine what is the maximal $K_j$ needed to transmit a given amount packets by the sensor node. As far as the relay nodes are concerned, they have more energy than the sensor nodes. Each room has a relay node, which acts as a CH. The CH can only receive $V$ number of packets in a time frame due to its capacity. In the case of TDMA, $V$ is the number of slots accommodated in a time frame. The task of scheduling is to assign time slots to the sensor nodes in such a way that the $X_j$ number of packets are sent in $K_j$ time and the probability of packet loss (in a given time slot more than $V$ sensor nodes are sending packets to the CH) must be uniformly distributed.

<table>
<thead>
<tr>
<th>Timeframe, size: $V$</th>
<th>W</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 : 2 : ... : $v$</td>
<td>1 : 2 : ... : $v$</td>
</tr>
</tbody>
</table>

Fig. 2. Typical structure of TDMA timeframes

The number of timeframes available in a relay node is denoted by $W$. The number of available timeslots in each
timeframe is denoted by $V$. These slots are summarized on Figure 2. In a TDMA protocol each node has a dedicated timeslot in every timeframe wherein it may send its packet to its parent node.

A. Mathematical model

Let us assume that there are $J$ number of nodes transmitting packets to a specific cluster head. The capacity of the cluster head in a time frame is denoted by $V$. This capacity means the amount of packet can be received in one time frame, the number of time slots. The amount of packets to be sent by node $j$ is denoted by $X_j$, while the time constraint in which the transmission is to be finished is denoted by $K_j$. $X_j$ is initially defined and using the equation (1) the $K_j$ can easily be calculated.

The time is measured in discrete units thus $K_j, j = 1, \ldots, J$ are assumed to be integers. The scheduling of node $j$ is represented by a binary vector $c_j \in \{0,1\}^{K_j}$ where if $c_j(l) = 1$ then a packet is sent to the cluster head at time instant $l$. The scheduling matrix $C$ can be constructed from vectors $c_j, j = 1, \ldots, J$, which form the row vectors of $C$ and the number of columns is taken as $L = \max_j K_j$. For example in the case of $J = 6$, $X_1 = 5$, $X_2 = 5$, $X_3 = 4$, $X_4 = 3$, $X_5 = 3$, $X_6 = 7$, $K_1 = 6$, $K_2 = 8$, $K_3 = 7$, $K_4 = 6$, $K_5 = 4$, and $K_6 = 8$ one specific scheduling matrix looks as follows:

$$C_{\text{valid}} = \begin{pmatrix}
1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 \\
0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 \\
1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 0
\end{pmatrix}.$$ 

In this example the capacity of the CH was $V = 4$, and as it can be seen the number of ones is balanced. The following example for $C$ matrix is not valid and unbalanced, and this scheduling overloads the CH.

$$C_{\text{invalid}} = \begin{pmatrix}
1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 \\
0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\
0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\
1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 0
\end{pmatrix}.$$ 

The aggregated number of incoming packets in the cluster head at time instant $l$ is given as $\sum_{j=1}^{J} c_j(l)$ being the weight of the $l$th column vector in matrix $C$. The cost of reception (which needs the power of CHs) is proportional to the number of received packets at time instant $l$ expressed as $a_l := \sum_{j=1}^{J} c_j(l)$. We seek an optimal matrix $C_{\text{opt}}$ where the aggregated number of incoming packets are balanced with respect to time. More precisely, we try to achieve that nodes schedule their packet transmission in such a way that each time instant the cluster heads receive more or less the same number of packets. This objective can be formulated as follows:

$$C_{\text{opt}} : \min_{C} \sum_{l=1}^{L} \sum_{k=1}^{L} \left( \sum_{j=1}^{J} C_{j,l} - \sum_{j=1}^{J} C_{j,k} \right)^2. \tag{2}$$

The constraint can be expressed as $\sum_{l=1}^{L} C_{j,l} = X_j, j = 1, \ldots, J$ and if the last nonzero component of row $j$ is at location $M_j$, such that $M_j : \sum_{l=1}^{L} C_{j,l} = X_j, j = 1, \ldots, J$ then $M_j \leq K_j, j = 1, \ldots, J$. We seek the balanced solution by optimizing (2).

Furthermore we deal with the assumption that there is a loss in the transfer of packets to the CH because of the overload of the link. Namely, there exists no scheduling matrix for which the aggregated load does not exceed the throughput $V$ of the CH node, i.e., for every optimal scheduling matrix $C_{\text{opt}}$ there exists at least one time instance when the aggregated load exceeds the throughput:

$$\forall C_{\text{opt}} : \exists k, V < \sum_{j=1}^{J} X_j(k) \tag{3}$$

This assumption is important because it means that the available bandwidth is fully utilized.

B. The scheduling protocol

In order to optimize matrix $C$ we need parameters $X_j, K_j, j = 1, \ldots, J$. In the first stage of the protocol we define the amount of packets we want to acquire data from the different sensor nodes (e.g. we want to get information about the status of the adequate room), which will specify parameters $X_j, j = 1, \ldots, J$. In the forthcoming discussion we define a communication protocol. Therefore, initially we assume that in each room $X_j, j = 1, \ldots, J$ are given.

1) The nodes are synchronized using the Reference Broadcast Synchronization (RBS [24]) protocol, in order to use the TDMA based package scheduling.

2) The values of $X_j$ are sent to the sensor nodes in order to specify the amount of packets to obtain a certain amount of information about the status of room containing node $j$.

3) Each sensor node determines the amount of timeframes $(K_j, j = 1, \ldots, J$, one timeslot in each timeframe), wherein the $X_j$ packets can be transmitted. In other word, by using equation (1) they calculate $K_j, j = 1, \ldots, J$. This information is sent back to the CH.

4) Having $X_j, K_j, j = 1, \ldots, J$ the CH will determine the optimal scheduling matrix $C_{\text{opt}}$.

5) If at any given timeslot an overload occurs (the weight of a given column vector exceeds $V$) then some randomly selected packets are discarded subject to uniform distribution. In this way, the fair treatment is provided for all sensor nodes.

6) Each node will receive the corresponding row vector of the optimized scheduling matrix $C_{\text{opt}}$.

7) In operation the nodes are transmitting packets to the CH only in those timeslot, which are indicated in the received row vector.
In the next section, the optimization of matrix $C$ is discussed by using a constructive polynomial algorithm named LBS.

IV. SOLUTION BY POLYNOMIAL ALGORITHM: LBS ALGORITHM

In this section a new algorithm is presented, which aims at balancing the number of "1"-s in every column vector of matrix $C$. The forthcoming algorithm is motivated by a simple consideration that if the scheduling of the load can be distributed uniformly over a given time interval, then the packet loss is reduced to the minima. Any deviation from this uniformity may cause loss to a given time instant, while in other time slots under-utilization occurs.

Let us analyze the complexity of the newly proposed LBS algorithm. We put the number of $\sum_{j=1}^{J} X_j$ "ones" into the matrix and this can be upper bounded by $JL$. There is a minimization phase in each insertion, which we evaluated in time $J$. Thus, the overall complexity is as follows

$$J \cdot \sum_{j=1}^{J} X_j \sim O(J \cdot (J \cdot L)) \sim O(J^2 \cdot L).$$

This complexity is between the cubic and quadratic time taking account the following assumption: $J < L$ therefore this algorithm runs in polynomial time.

The newly presented LBS algorithm gives optimal solution to (2) and runs in polynomial time. In the next section extensive simulations demonstrate the performance of the method.

Algorithm 1 LBS algorithm

Require: $\forall i = 1,...J : K_i, X_i$

$L \leftarrow \max_j K_j$

$C \leftarrow 0_{J \times L}$

$S \leftarrow \sum_j X_j$

while $S > 0$ do {Number of unscheduled packets}

for $l \leftarrow L$ to 1 do

Find $\arg \min_i K_i$ where $(l < K_i) \land (X_l > 0)$

$S \leftarrow S - 1$

$X_l \leftarrow X_l - 1$

$C_{i,l} \leftarrow 1$

end for

end while

The figure shows the entropy of the solutions. One may note that the horizontal axis is scaled logarithmically. (The exhaustive search is not visible as the results of the LBS algorithm fully overlaps it.)

The previous and following graphs contain the results of a Hopfield Neural Network (HNN), our previous solution to solve the specified problem. Details about the HNN may be found in [8] and [25]. The results of the HNN only slightly differs from the optimum.

V. PERFORMANCE ANALYSIS

In this section, we compare the optimal solution obtained by exhaustive search to the one achieved by the LBS algorithm and to the one achieved by a random strategy. The simulations were carried out for $J = 20$ nodes and the corresponding $X_j, j = 1, ..., J$ and $K_j, j = 1, ..., J$ constraints have been chosen randomly in the range of $X_j \in [5, 100]$ and $K_j \in [5, 100]$. The results have been evaluated after selecting several constraints randomly, then running the simulations and taking the average error of the solutions achieved by the methods and by exhaustive search. The random method and the exhaustive search are the reference solutions in order to provide a better comparison of this new LBS method.

We took into consideration the properties of the platform specified in Section II in the simulations we performed.

The obtained results are depicted by the Figure 3. Analyzing the result on the figure, one may note that the solution provided by the LBS algorithm is possible as good as the results provided by the Exhaustive search. The Random strategy performs more worse than the LBS algorithm or the Exhaustive search. The figure shows the entropy of the solutions. One may note the horizontal axis is scaled logarithmically. (The exhaustive search is not visible as the results of the LBS algorithm fully overlaps it.)

The running time of the enumerated methods are compare to each other in Figure 4. The LBS algorithm, the HNN and the random strategy provide its solutions in polynomial time opposed to the exponential complexity of the exhaustive search method. (The exhaustive search due to the exponential complexity does not appear on the previous graph.) The LBS
Fig. 4. Running time of solutions achieved by different methods

algorithm is faster than the HNN recursion, the reason are detailed:

- The random strategy generates a random matrix keeping the constraints $X_j, K_j: \forall 1 \ldots J$. The complexity of this generation is $O(J \cdot L)$

- The theoretical complexity of the LBS algorithm is (as detailed previously): $O(J^2 \cdot L)$

- The time and steps of convergence for the discrete HNN using the strong Lyapunov criteria could be proven as follows: $O(J^2 \cdot L^2)$

- The exhaustive search uses huge search space and elects the optimal scheduling matrix from all possible $C$ matrix. Therefore its complexity:

$$O(2^J \cdot L)$$

Therefore we expect the results depicted by the Figure 4. (The running time is measured in seconds and the vertical axis is scaled logarithmically.) We repeated out tests several times but not enough times to provide monotone variation as one may expect. Certainly these results can be extrapolated.

In order to measure how balanced a solution is, we introduce the entropy of the weight distribution of the columns in matrix $C$ as follows

$$H(p) = \sum_{k=1}^{L} -p_k \ln p_k, \quad (5)$$

where

$$p_k = \frac{\sum_{j=1}^{J} C_{jk}}{\sum_{k=1}^{L} \sum_{j=1}^{J} C_{jk}}.$$

This entropy value provides a heuristic metric to describe the measure of equalization of the packet loads in the different solutions.

Table I demonstrates that the solution provided by LBS algorithm has the highest weight entropy (i.e., the most uniform weight distribution of the columns), hence it fulfills the constraint related to weight balancing.

<table>
<thead>
<tr>
<th>Method</th>
<th>Entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random Strategy</td>
<td>5.45</td>
</tr>
<tr>
<td>HNN</td>
<td>5.62</td>
</tr>
<tr>
<td>LBS algorithm</td>
<td>5.89</td>
</tr>
<tr>
<td>Exhaustive Search</td>
<td>5.89</td>
</tr>
</tbody>
</table>

TABLE I

ENTROPY OF WEIGHT DISTRIBUTION

Figure 5 shows the advantages of the LBS algorithm, since the number of packets hitting the CH at in time frames are equalized.

One may also see that if a capacity $V$ is given (in this example let $V = 9$) then the scheduling methods provided either by the Random Strategy or by the HNN will suffer from packet loss, as the CH cannot receive all the packets in the given time instances. This effect is very sharp if the timeframes are less than 25. In the case of LBS algorithm, or Exhaustive Search there is no packet loss even in these cases due to more uniform distribution of the load.

Fig. 5. Received packets in timeframes
The random strategy follows from a simple algorithm of constructing a scheduling matrix. In each row of the matrix the $X_j$ number of “1”-s are randomly placed within the $K_j$ time limit. This method is used for comparison, it serves as a “dummy” model.

VI. CONCLUSION

In this paper, new methods were proposed to provide optimal scheduling for packet-transmissions in WSN. The objective of optimal scheduling was defined by the Balanced Cost objective function. It is noteworthy, that the newly introduced methods yields always a valid scheduling matrix, as opposed to the HNN or random scheduling, which may yield invalid matrices. The algorithm introduced above is capable of providing optimal scheduling with given constraints and guarantees balanced load. It runs in polynomial time and yield valid solution therefore it is possible to use in real-life WSN data acquisition applications.

REFERENCES


