International Journal of Advances in Computer Networks and its Security – IJCNS
Volume 4 : Issue 1 [ISSN 2250 – 3757]
Publication Date: 09 January 2014

Energy Efficient Data Aggregation using Bitmask techniques for Wireless Sensor Networks

Santhosh kamath

Abstract—**The rapid advancement of hardware technology has enabled the development of small, powerful, and inexpensive sensor nodes, which are capable of sensing, performing computation and wireless communication. This modernizes the deployment of wireless sensor network for monitoring some vicinity and collecting required information. However, sensor nodes have limited energy constraint which is a major challenge for such vision to become reality. We consider energy constrained wireless sensor network deployed over a region. The main job of such a network is to gather information from node and transmit it to sink node for further processing. So the aim of any data forwarding protocol is to conserve energy to maximize the network lifespan. Sensor nodes are capable of performing in-network aggregation of data coming from more than one source. In this paper we focused on energy consumption issue and aim to develop an energy efficient data aggregation protocol. To deliver energy efficiency we have considered a cluster-based wireless sensor network. This approach executes on each cluster independently and provides an energy efficient data aggregation in a cluster to maximize network lifetime for whole network.**

Keywords- **Data aggregation, Wireless sensor networks. Energy efficiency, Bit mask.**

I. Introduction

A wireless sensor network is a wireless network consisting of tiny devices which monitor physical or environmental conditions such as temperature, pressure, motion or pollutants etc. at different regions. The tiny device, known as sensor node, consists of a radio transceiver, microcontroller, power supply, and the actual sensor. Initially sensor networks were used for military applications but now they are widely used for civilian application area including environment and habitat monitoring, healthcare application and so on.

Santhosh Kamath
Department of Information and Communication
Manipal Institute of Technology
Manipal University
India
santosh.kamath@manipal.edu

Normally sensor nodes are spatially distributed throughout the region which has to be monitored. They self-organize into a network through wireless communication, and collaborate with each other to accomplish the common task. With the going time, sensor nodes are becoming smaller, cheaper, and more powerful which enable us to deploy a large-scale sensor network. Basic features of sensor networks are self-organizing capabilities, dynamic network topology, node failures, limited power, mobility of nodes, short-range broadcast communication, multi-hop routing, and large scale of deployment [1]. The strength of wireless sensor network lies in their flexibility and scalability. University of California at Berkele has developed Mica mote which is a special purpose sensor node. Other special purpose sensor nodes available are Spec, Rene, Mica 2, Telos etc. Some high bandwidth sensor nodes available are BTNode, Imote 1.0, Stargate, Inryonc Cerfeube etc. [2].

The most important constraint imposed on sensor network is the limited battery power of sensor nodes. The effective lifetime of a sensor node is directly determined by its power supply. Hence lifetime of a sensor network is also determined by the power supply. So, the energy consumption is main design issue of a protocol. Limited computational power and memory size is another constraint that affects the amount of data that can be stored in individual sensor nodes. So, the protocol should be light-weighted and simple. Communication delay in sensor network can be high due to limited communication channel shared by all nodes within each other's transmission range. It is widely accepted that the energy consumed in one bit of data transfer can be used to perform a large number of arithmetic operations in the sensor processor [3]. Many techniques have been suggested to aggregate the data in the Sensor Network. Data aggregation techniques explore how the data is to be routed in the network as well as the processing method that are applied on the packets received by a node. They have a great impact on the energy consumption of nodes and thus on network efficiency by reducing number of transmission or length of packet.

There are two approaches for in-network aggregation in terms of size of the data: that is with data size reduction and without data size reduction. The aggregation in networks with data size reduction refers to the process of combining and compressing the data packets received by a node from its neighbors in order to reduce the packet length to be transmitted or forwarded towards the sink. In this paper we propose a data aggregation technique which helps in keeping the precision of the data while reducing the packet size to be transmitted to the sink from cluster head (CH). This enhances the battery life by reducing the transmission energy.

For Mica Motes, TinyOS predefined a maximum packet size of 36 bytes. As shown in Figure 1, out of the 36-byte of the packet, 29-bytes are allocated to sensor data (payload) and remaining bytes to destination address, Active Message (AM) type, length, group and Cyclic Redundancy Check (CRC).The payload may consists of data, an Encryption key/s for security reason and source ID. Since the size of the payload is limited to 29-bytes there must be an optimal method in order to adjust IDs of a large number of sensor nodes in a single packet for huge WSNs.

Dest (2)	AM (1)	Len (1)	Grp (1)	Data (0 - 29)	CRC (2)

Figure 1. Tiny OS packet format

The paper is organized as follows. Section II presents related work addressing the data aggregation. Section III describes data compression technique. Section IV Experimental analysis and result and the section V conclusion and the future work.

II. RELATED WORKS

Most of the work done till now on in-network aggregation mainly deals with problem of forwarding packets from source to sink, to facilitate aggregation therein. Actually the main idea behind were to enhance existing routing protocols such that they can efficiently aggregate data. Survey says that most of the data aggregation techniques fall under three categories. They are tree-based approaches, multi-path approaches, and cluster-based approaches. There are some hybrid approaches also that combine any of the three techniques above. So, some of the approaches are described in coming sections with giving details of some of the main techniques by different authors.

i. Tree Based Approach

The simplest way to aggregate data is to organize the nodes in a hierarchical manner and then select some nodes as the aggregation point or aggregators. The tree-based approach perform aggregation by constructing an aggregation tree, which will be a minimum spanning tree, with the sink as route and source nodes are considered as leaves. Every node has a parent node to forward its data. Flow of data starts from leaves nodes and ends at the sink and therein the aggregation done by parent nodes. In [7] author proposed a data-centric protocol which is based on aggregation tress, known as Tiny Aggregation (TAG) approach. In [3] author proposed a reactive data-centric protocol for applications where sink ask some specific information by flooding, known as directed diffusion model. The main idea behind directed diffusion model is to combine data coming from different source and en-route them by eliminating redundancy, minimizing the number of data transmission; thus maximizing network lifetime. In [3] author proposed a new low-control-overhead data dissemination scheme, which they called as Pseudo-Distance Data Dissemination (PDDD), for efficiently disseminating data from all sensor nodes to mobile sink. Some assumption are made, they are: (1) All source nodes maintain routes to mobile sink node, (2) No periodically messaging for topological changes due to mobile sink node, (3) All link are bi-directional and no control messages are lost, (4) Mobile sink nodes have unlimited battery supply, so no need to care about battery efficiency of sink node, and (5) Network partitioning is not considered. Data dissemination process is influenced by directed diffusion [4]. In [2] author proposed an energy-aware spanning tree algorithm for data aggregation, referred as E-Span. ESpan is a distributed protocol in which source node that has highest residual energy is chosen as root. Other source nodes choose their parent based on residual energy and distance to the root. The protocol uses configuration message to exchange information of node i.e., residual energy and distance to reach the root. Each node performs single-hop broadcast operation to send packets. Single-hop broadcast refers to the operation of sending a packet to all single-hop neighbors [2].

ii. Cluster-Based Approach

This scheme is to organize the network in hierarchical manner. In cluster-based approach, whole network is divided in to some clusters. Each cluster has a cluster-head (CH) which is selected among cluster members. Cluster-heads aggregates data received from cluster members locally and then transmit the result to sink. The advantages and disadvantages of the cluster-based approaches are very much similar to tree-based approaches. In [5] author proposed a maximum lifetime data aggregation (MLDA) algorithm which finds data gathering schedule provided location of sensors and the base station, data packet size, and energy available in each of the sensor. A data gathering schedule specifies how data packet are collected from sensors and transmitted to base station in each round. A schedule can be considered of as a collection of aggregation trees. In [6], they proposed heuristic-greedy clustering-based MLDA based on MLDA algorithm. In this nodes partitioned the network in to cluster and referred each cluster as super-sensor. Super sensor compute maximum lifetime schedule for these super sensors and then use this schedule to construct aggregation trees for the sensors. In [8] they present energy efficient and secure pattern based data aggregation protocol which is designed for clustered environment. In conventional

International Journal of Advances in Computer Networks and its Security – IJCNS
Volume 4 : Issue 1 [ISSN 2250 – 3757]
Publication Date: 09 January 2014

method data is aggregated at cluster-head and cluster-head eliminate redundancy by checking the data content. The protocol says that instead of sending raw data to cluster head, the cluster members send corresponding data codes to cluster-head for data aggregation. If cluster members send the same pattern code then only one of them is finally selected for sending actual data to cluster-head. Author presents a pattern comparison algorithm, for pattern matching

III. Energy Efficient Data Aggregation for Wireless Sensor Networks (EEDA)

In this approach we have taken the advantages of cluster based approach. At the beginning all the sensor nodes which are deployed on a given area forms a cluster using existing cluster based algorithm like used in [9]. There after each node sends its data to the cluster head. The cluster head compresses the data using bitmask based techniques [10] and sends it to the sink node using routing algorithm like AODV.

A. *Network Model*

We assume a wireless sensor network model which is appropriate for data gathering applications such as water vapor pressure monitoring system in the paddy field. The dew point temperature is calculated from thermodynamic equations stored in EPROM of the sensor. The sensor nodes are deployed in a paddy field of 20 acres uniformly and we have taken the reading of the nodes at different temperature ranging from 20^0 to 40^0 Celsius. We observed that the data sensed by the neighboring nodes are almost same or nearest value between 6 to 8mV. And the corresponding humidity is 25 to 30 g/M^3 [13] (correlated data in terms of bit position) as shown in the figure 2. The network model used has the following properties. First, a sink node without energy constraint is the root of the network topology and located on the top of it. Second, a large number of energy-constrained sensor nodes (e.g., MICA Motes) are deployed uniformly in the paddy field and they are equipped with power control capabilities. Third, each sensor node has the capabilities of sensing, aggregating and forwarding data and it can send fixed-length data packets to the cluster head node periodically. Finally, the sensor nodes can switch into sleep mode or a low power mode to preserve their energy when they do not need to receive or send data [7].

B. *Data Aggregation Algorithm*

To avoid unnecessary communications overheads and achieve energy efficient data aggregation for WSNs, we present an algorithm for data aggregation in WSNs. The main goal of the proposed algorithm is to provide data aggregation capability to the sensor nodes and they use it during data transmissions to the sink node. So we use Dictionary based technique to compress the data [10].

C. *Dictionary based compression*

The basic idea of this is to take advantage of commonly occurring data sequences by using a dictionary. The repeating data are replaced with a code word that points to the index of the dictionary that contains the pattern. One of the major challenges in bitmask-based code compression is how to determine (a set of) optimal mask patterns that will maximize the matching sequences while minimizing the cost of bitmasks. A 2 bit mask can handle up to 4 types of mismatches while a 4-bit mask can handle up to 16 types of mismatches. Clearly, applying a larger bitmask will generate more matching patterns; however, this may not result in better compression. For our application we can use a 2 bit mask since this application data is correlated to each other, the hamming distance is 1-2 bits and data sensed by the sensors are 8 bit in length. Figure 2 shows the generic encoding scheme used by bit masking compression technique and operation involved is very simple i.e. XOR.

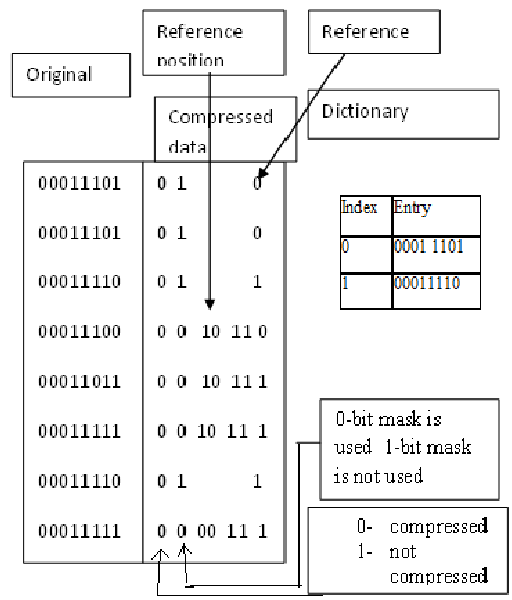

Figure 2. Example for Bitmask based compression

D. *Data Forwarding*

Once the cluster head completes the compression of the data received from the subordinate node, it can send it to the sink using the format given in the figure 3. The data field of the format Figure 1 is modified as shown in Figure 3 to transmit the compressed data to the sink. Data here refers to the compressed data from the different node which is of variable bits. And padding of 0-8bits are used to make the total length multiple of 8 bits.

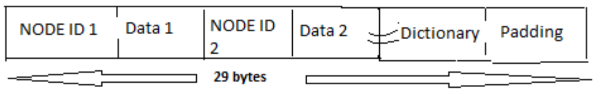

Figure 3. Packet format to send the data from CH to the sink

E. Dictionary selection in sensor networks

Among the available data at the cluster head, it can select most repeating two data set as dictionary data. In the Figure 2 shows that the first and the third occur multiple times so these two data can be selected as dictionary data. Other data's are correlated to this dictionary data. Figure 4 shows the algorithm to for compressing the data in the CH.

N is the number of data received
K1 and K2 dictionary data
Upon receiving all the data from the subordinate nodes
For (i=0;i<n;i++)
Select most frequently occurring two data sets and use it as dictionary references say k1 and K2 with the index 0 and 1 respectively
For (i=0;i<n;i++)
Hamming distance1= data[i] XOR K1
Hamming distance 2=data[i] XOR K2
If (hamming distance1 > hamming distance2 &&
Hamming distance<3)
Then Select index as 0 or index as 1 depending on
Small distance.
First position =0
If (hamming distance=0)
Then Second bit =1 and third bit= index for
Which the hamming distance=0
Else second bit= 0 third and fourth bits=
Hamming difference, and fifth and sixth bits=
Position of the hamming difference Last bit=
Reference
End if
Else first bit= 1 and
Data= original 8 bit data sensed.
End if

Figure 4. Algorithm for Cluster Head to Compress Data and Select Dictionary

F. Analytical Models

In this section, we present analytical model for the data aggregation schemes .The energy consumption issue for WSNs is the most important because the lifetime of a sensor node is extremely depends on the available energy of its battery. There are three domains to be considered regarding energy consumption: (i) sensing activity (data collection from the environment), (ii) communication (sending and receiving packets) and (iii) data processing/in-network data aggregation. Although all these activities waste energy, communication is responsible for the bulk of the power consumption which is the main point of attention in many algorithms designed for sensors networks. That is to say, the energy consumed in one bit of data transfer can be used to perform a large number of arithmetic operations in the sensor processor [3].

Consider a homogeneous network of N sensor nodes and a base station or sink node distributed over a region. The location of the sensors and the base station are fixed and known priori. Each sensor produces some information as it monitors its locality. We assume that the whole network is divided into several clusters; each cluster has a cluster-head (CH). The clustering and the selection of cluster-head (CH) can be done by using any existing protocol like LEACH.

This energy model we considered for the sensors is based on the first order radio model described in [11]. A sensor consumes $E_{elect} = 50\ nJ/bit$ to run the transmitter or receiver circuitry and $E_{amp} = 100\ pJ/bit/m^2$ for the transmitter amplifier to achieve an acceptable signal to nose ratio. Thus, the energy consumed by a sensor i in receiving a k-bit data packet from N nodes is given by the following equation (1).

$$E_{Rx} = \mathrm{E}_{elect} * k * \mathrm{N} \quad (1)$$

While the energy consumed in transmitting the data packet to sink node is given by the following equation (2).

$$E_{Tx\ i,j} = \mathrm{E}_{elect} * k + E_{amp} * d_{i,j}^2 * k \quad (2)$$

Where $d_{i,j}$ is the distance between nodes i and j.

After the compression the number of bits required to transmit a packet at CH will reduce by a compression factor C. The value of C will come in between 0 to 1.

$$C = \frac{\mathrm{Compressed}_{dats\ size} + Dictionary_{data\ size}}{\mathrm{Original}_{data\ size}} \quad (3)$$

Apply equation (3) in equation (2), so we will get equation (4) as follows.

$$E_{Tx\ i,j} = (\mathrm{E}_{elect} * k + E_{amp} * d_{i,j}^2 * k) * C \quad (4)$$

From equation (4), the energy required to transmit the data bits using this approach reduce by a factor of *C.*

IV. EXPERIMENTAL ANALYSIS AND RESULTS

To test this approach, we determined energy consumption of all the nodes in the network, which is defined as the sum of energy consumed in transmit, receive, idle and sleep mode of all nodes and considered that as the metric to prove energy efficiency. We used QualNet simulator to test the approach [12]. The nodes are deployed over 500 square meter area, and we have used to monitor the temperature of a 500 square meter area and we have transmitted data of temperature values in between 29 to 32 degree Celsius to the CH by the all the nodes randomly. We also evaluated the battery consumption in the network using battery model option available in QualNet for the conventional and as well as the proposed approach. We got the following result as shown in Figure 5, which clearly shows the energy required for transmitting the data using conventional approach is more than the proposed EEDA. The graph is moving down at simulation time 20,000 indicates that some of the nodes are dying from battery supply but EEDA shows that it is still available indicates more lifetime of the sensor networks.

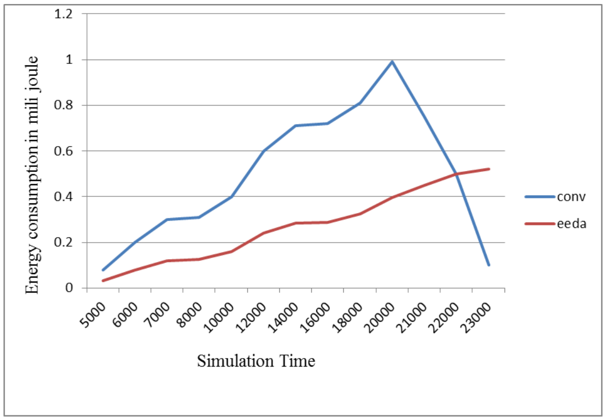

Figure 5. Average energy consumption of the cluster as a function of time.

To determine the energy gain ratio for this approach we calculated the average energy consumed by this scheme in all the experiments done at different time interval to that of conventional method. Table 1 illustrates energy efficiency of EEDA over the conventional method. We found that the proposed work dissipates just about 40% of that energy which is required by the conventional scheme.

Table 1. Average Energy consumption of a cluster over function of time.

Method	Average Energy consumption (mili Joules)	Energy gain ratio
EEDA	0.199077	0.40
Conventional	0.497692	

V. Conclusion and Future Work

This work studies the energy-efficient cooperative data aggregation problem for wireless sensor networks and we proposed simple compression based techniques to send the data to the sink without compromising the loss of the data. We applied the existence bitmask based techniques with some modification for wireless sensor. The theoretical analysis indicates that we can save energy using EEDA without compromising the quality of the data. The experimental simulations show that the proposed approach can decrease the power consumption by about 40 percent. In this approach we have considered the hamming distance up to 2bits. The future work depends upon application of this approach in wireless sensor network in military like object tracking system, the where data may not be correlated like water vapor sensing application. So we need to work on selection of the proper dictionary values for less correlated data.

References

[1] Mohammad Ilyas and Imad Mahgoub(2005) "Handbook of Sensor Networks: Compact Wireless and Wired Sensing Systems" CRC press Boka Ratton London, ISBN 0-8493-1968-4.

[2] Morais Cordeiro and Dharma Prakash Agrawal(2003) "Ad-hoc & Sensor Networks" by Carlos de 2nd edition, ISBN 9814338893.

[3] M. Lee, and S. Lee, "Data Dissemination for Wireless Sensor Networks", in Proceedings of the 10th IEEE International Symposium on Object and Component-Oriented Real-Time Distributed Computing (ISORC'07).

[4] C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann, and F. Silva, "Directed Diffusion for Wireless Sensor Networking", IEEE/ACM Transactions on Networking, Vol. 11, no. 1, Feb 2003.

[5] K. Dasgupta et al., "Maximum Lifetime Data Gathering and Aggregation in Wireless Sensor Networks", In Proc. of IEEE Networks'02 Conference, 2002, Atlanta,pp.325-349.

[6] K. Dasgupta, K. Kalpakis, and P. Namjoshi, "An Efficient Clustering-based Heuristic for Data Gathering and Aggregation in Sensor Networks", Proc. Of 5th ACM/IEEE Moicom.seattle, WA,pp.271-278.

[7] S. Madden et al., "TAG: a Tiny Aggregation Service for Ad-hoc Sensor Networks," OSDI 2002, Boston, MA, Dec. 2002.

[8] H. Cam, S. Ozdemir, P. Nair, and D. Muthuavinashiappan, "ESPDA: Energy-Efficient and Secure Pattern-based Data Aggregation for Wireless Sensor Networks", in Proceedings of IEEE Sensor- The Second IEEE Conference on Sensors, Toronto, Canada, Oct. 22-24, 2003, pp. 732-736.

[9] Jin Ke-yin "Based on the Improvement of LEACH Protocol for Wireless Sensor Network Routing Algorithm" Intelligent System Design and Engineering Application (ISDEA), 2012 Second International Conference pp.1525-1528.

[10] IEEE transaction on computer-aided design of integrated circuits and analysis VOL. 27, NO. 4, April 2008 Bitmask-Based Code Compression for Embedded Systems Seok-Won Seong and Prabhat Mishra, Member, IEEE.

[11] W. Heinzelman, A.P. Chandrakasan, and H. Balakrishnan, "Energy-Efficient Communication Protocols for Wireless Microsensor Networks", In Proc. Of Hawaiian Intl. Conference on System Science, 2000.

[12] http://web.scalable-networks.com/content/libraries-element-models.

[13] http://www.wmo.int/pages/prog/www/IMOP/.../IOM-87_Module_B2_2.pps