International Journal on Emerging Technologies (Special Issue on NCRIET-2015) **6**(2): 387-390(2015)

ISSN No. (Print) : 0975-8364
ISSN No. (Online) : 2249-3255

Data Aggregation and Routing In Wireless Sensor Networks: Optimal and Heuristic Algorithms

Chamundeshwari Kalyane
CSE, Dept, BKIT, Bhalki

(Corresponding author: Chamundeshwari Kalyane)

(Published by Research Trend, Website: www.researchtrend.net)

ABSTRACT: A fundamental challenge in the design of Wireless Sensor Networks (WSNs) is to maximize their lifetimes especially when they have a limited and non replenishable energy supply. To extend the network lifetime, power management and energy-efficient communication techniques at all layers become necessary. In this paper, we present solutions for the data gathering and routing problem with in-network aggregation in WSNs. Our objective is to maximize the network lifetime by utilizing data aggregation and in-network processing techniques. We particularly focus on the joint problem of optimal data routing with data aggregation en route such that the above mentioned objective is achieved. We present Grid-based Routing and Aggregator Selection Scheme (GRASS), a scheme for WSNs that can achieve low energy dissipation and low latency without sacrificing quality. GRASS embodies optimal (exact) as well as heuristic approaches to find the minimum number of aggregation points while routing data to the Base Station (BS) such that the network lifetime is maximized. Our results show that, when compared to other schemes, GRASS improves system lifetime with acceptable levels of latency in data aggregation and without sacrificing data quality.

Keywords: Wireless Sensor Networks, Data Aggregation and Routing, Exact and Heuristic Solutions, Hierarchical Clustering.

I. INTRODUCTION

Wireless Sensor Networks (WSNs) is a class of wirelessad hoc networks in which sensor nodes collect, process, and communicate data acquired from the physical environment to an external Base-Station (BS) [1]. Future WSNs are envisioned to revolutionize a maintenance free and fault tolerant Aside from the task of efficient design of data aggregation algorithms, the task of finding & maintaining routes in WSNs is also nontrivial [2], especially when it includes the selection of aggregation points and routing through those points. Many routing and data dissemination with aggregation protocols have been proposed for WSNs (a comprehensive survey of the routing techniques in WSNs can be found in [2]). In [4], C. In tanagonwiwat et. al. proposed a popular data aggregation paradigm for WSNs, called Directed Diffusion (DD) where aggregation is used to reduce communication costs. In [6], Heinzelman, et. al. introduced a hierarchical clustering algorithm for WSNs, called Low Energy Adaptive Clustering Hierarchy (LEACH). LEACH is a cluster-based protocol where Cluster Head (CH) nodes compress data arriving from nodes that belong to the respective cluster, and send an aggregated packet to the BS. Following these protocols, many other studies focused on the routing problem [7]-[22]. Among these, [18] introduced a linear programming formulation to solve the optimal routing problem in WSNs. The objective was to maximize the network life time, and the life time was defined as the network operational time until one of the nodes fails. Necessary, but solution under arbitrary traffic generation processes were introduced. A heuristic solution was also introduced. However, no aggregation was assumed. In [19], aggregation was take into account, but only full aggregation was considered. That is, regardless of the number packets be aggregated, a single packet will always be produced. This simplifies the problem of aggregation significantly. A special case of partial aggregation, in which the aggregated data size is In this paper, we present a novel data aggregation and routing scheme, called Grid-based Routing and Aggregator Selection Protocol (GRASS). GRASS embodies optimal (exact) as well as heuristic approaches to find the minimum number of aggregation points while routing data to the BS such that the network lifetime is maximized.

That is, GRASS jointly addresses the issues of the selection of data aggregation points, and the optimal routing of data from sensors to aggregation points, as well as the routing of the aggregated data to the BS. While solving these two problems separately may simplify the problem, the solution may be far from optimal. Therefore, our proposed solution treats the two problems jointly in
order to reach an optimal solution. Since this joint problem is not trivial, we adopt a hierarchical structure in which each group of sensor nodes elect a cluster head which is responsible for: 1) collecting their sensed data, 2) performing a first level aggregation, and then 3) routing this data to the next aggregator on its way to the BS. This first level of aggregation achieves two benefits. First, it offers the greatest performance benefits in this environment since nodes in a cluster are most likely to generate correlated data, and then it simplifies the routing function since only the cluster head will be in charge of this functionality. Hence, the hierarchical structure facilitates digests of sensor data. Indeed, this is a key issue in the design of GRASS. In GRASS, correlation means that sensors' readings overlap statistically as they monitor the same event. This overlap will be captured in our proposed solutions using aggregation overlap factor. The factor represents linear as well as nonlinear relations among the gathered data. We propose to solve the joint aggregator selection and routing problems in a powerful node, such as the BS, and then dispatch the results to the sensor nodes. Hence, an optimal solution that is obtained by the BS will result in an optimal routing and aggregation strategy. The rest of this paper is organized as follows. The problem description and system model are presented in Section II. Section III presents exact algorithms to solve the problem, and using two definitions of network lifetime. Section IV presents several approximate algorithms for the problem under consideration. Section V presents analysis of energy delay tradeoffs due to our aggregation scheme. The performance evaluation of the proposed scheme is presented in Section VI. We conclude with final remarks in Section VII.

II. THE PROBLEM DESCRIPTION AND SYSTEM MODEL

We consider a network of fixed, homogeneous, and energy-constrained sensor nodes that are randomly deployed in a sensor field (bounded region). Each sensor acquires measurements which are typically correlated with other sensors in its vicinity, and these measurements are to be gathered and sent to the BS for evaluation or decision taking purposes. We assume *periodic sensing* with the same period for all sensors. We also assume that contention between sensors is solved by the MAC layer1. We assume that the information collected by various sensors may be correlated, redundant, and/or of different qualities. Since data correlation in WSNs is strongest among data signals coming from nodes that are close to each other, we believe that the use of a clustering infrastructure will allow nodes that are close to each other to share data before sending it to the BS. Hence, the ideas of fixed cluster-based routing together with application-specific data aggregation techniques will be used. The four basic components of a sensor node are sensing unit, a processing unit, a transceiver unit and a power unit. As per fig 1 sensing devices are usually composed of two subunits: sensors and analog to digital converters (ADCs). The sensor produces analog signals which are converted into the digital signals and are further fed into processing unit. The processing unit consists of a small storage unit, manages the procedures. The nodes are connected to the network by the transceiver. Power unit is the most important component of the sensor node which comprises of secondary power storage units like batteries or a power scavenging unit such as solar cells. There are also other subunits which depend on the application. There are different types of sensors used. Like, seismic, low sampling rate magnetic, thermal, visual, infrared, acoustic and radar, which are able to monitor a wide variety of ambient conditions that include the following [4]: temperature, humidity, vehicular movement, lightning condition, pressure, soil makeup, noise levels, the presence or absence of certain kinds of objects, sand size of an object.

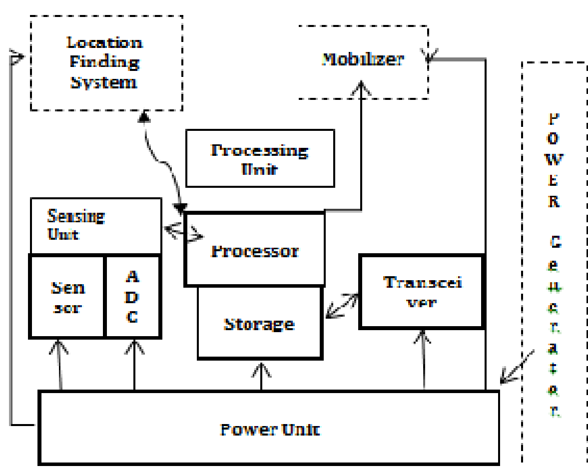

Fig. 1: Sensor Node-Components

III. EXISTING WORK

1. Energy Balanced Routing Method for In-Network Data Aggregation in Wireless Sensor Networks[1].

This paper propose an Enhanced Forward Aware Factor-Energy Balanced Routing Method (EFAF-EBRM) based on Data aggregation technique that has some key aspects such as a reduced number of messages for setting up a routing tree, maximized number of overlapping routes, high aggregation rate, and reliable data aggregation and transmission.

2. Data Aggregation in Wireless Sensor Network: A Survey In this paper authors discuss about data aggregation and its various energy-efficient technique used for data aggregation in WSN.

3. Requirements of Quality of Service in Wireless Sensor Network [13]. In this paper authors define WSNs QoS requirements within a WSNs application, and then analyzing Issues for QoS.

IV. OBJECTIVE FUNCTION

Minimize _The objective function minimizes the maximum zone failure rate Subject to: The following two constraints find the maximum power consumption over all LA nodes, while respecting the maximum power consumption limit of any LA node,_ ≥ Pi/Ei , ∀ i (19) _ ≤ 1/Tmin (20). The following constraint ensures that the power consumed at each LA node is sufficient to send the amount of required traffic to the other nodes .Pi ≥ X g □j ,if Fi j=1Xg ijw T + Xj, if Fji=1Xg jiwR ∀g , i (21) The following two constraints ensure that if K is the number of members of group g sending data from node i to node j, then the value of Ig, k ij is 1 for all the values of k that are greater than or equal to K; otherwise it will be 0. |Sg| Xk=0

The following two constraints together give the exact amount of traffic sent by node i to node j, once node I has aggregated the data coming from members of group g: The following three constraints ensure that the aggregated traffic streams will not be split on the way to the BS:

Mg ij ≥ X s∈Sg T s,g ij /Q , ∀g, , i, j, if Fij = 1 (28)

Mg ij ≤ X s∈Sg T s,g ij , ∀g, , i, j, if Fij = 1 (29)

X j, if Fij=1 Mg ij ≤ 1 , ∀g , i (30)

The guarantee of a minimum lifetime of an LA node is highly dependent on the determination of the actual routing of the data traffic, i.e., we must find the power consumed by each LA source node when participating in routing data over G. The following additional set of constraints are required for performing route computations. The following two constraints ensure that for the traffic from source s to the base station, 0, no traffic is going in (going out) the source s (destination 0), respectively)

X i,if Fis=1,s6=i T s,g is = 0 ; X j,if F0j=1,j6=0 T s,g 0j = 0 ∀ g, s ∈ Sg

The following two constraints ensure that the traffic from s and 0 is originating (terminating) at s (0), respectively

X j,if Fsj=1,s6=j T s,g sj = 1 X i,if if Fi0=1,s6=0 T s,g i0 = 1 ∀g, s ∈ Sg

The following constraint preserves the continuity of connection traffic on one of multiple possible routes

X i,if Fix=1,i6=x,s6=i T s,g ix = X j,if Fxj=1,j6=x,j6=s T s,g xj ∀ x, g, s ∈ Sg, (1 ≤ x ≤ n, x 6= s) (31)

Finally, we point out that the number of variables used in this formulation is O(n 2χ + χE maxg |Sg|), which is based on the number of X g ij and T s,g ij (or I g,k ij), respectively

A. Genetic Algorithms Approach We developed a genetic algorithm strategy to solve both the RSP1 and RSP2 problems. The major step in GA is to find an efficient way to represent the solution. The detailed operation of GA is well-known and can be found in [25]. The ith cell in the string contains the route number that will be used by the ith source of group g, which has an integer value between 0 and Ri max, Ri max denotes the number of routes from the ith source to BS. We assume that this number is known and can be found by any route discovery technique, e.g. reactive protocols. Hence, the individual is a routing structure for each group. First, we generate an initial population of randomly created, and thus different, routing structure for each group g. The generated population is subjected to the typical GA evolution process.. Due to the lack of space, we omit the details of the genetic algorithms approach, which are available in [42]. Results based on this algorithm

V. PERFORMANCE EVALUATION

The performance of the algorithms of GRASS were tested with various experiment a scenarios which were simulated using the NS-2 simulator [44]. Each experiment corresponds to a random placement of sensors in a fixed network area. We assume a single base-station attempting to gather information from a number of data sources in the network area. The location of the base-station can be arbitrarily chosen. Unless stated otherwise, the BS is located at lower edge of the grid (0,0). We randomly place sensor nodes in a 50m×50m square field while always insuring that the initial distribution of sensor nodes always results in a connected graph, as will be explained below. We also experimented with larger sensor fields to test the performance of various heuristics in large networks.

It is assumed that the sensing range is the same as the transmission range which was set to a default value of 20 meters. The sensor field is divided into the appropriate number of zones, which is 30. We consider four scenarios corresponding to the distribution of sensors in the sensing field, which result in z nonempty zones (clusters) forming a connected virtual graph G. In the four scenarios, n takes values of 6, 8, 10, and 15, respectively. In each nonempty zone, there are on average 10 sensor nodes monitoring the area of that zone. We assume that sensors generate data packets of variable sizes such that the packet size is exponentially distributed with mean value of 1000 bits. In another setting, we fixed the packet size generated by all nodes for the sake of comparison with other schemes. In all settings, the aggregation function was based on taking the packet size with the maximum length. Each sensor i has a battery with finite, non-replenishable energy, which was set to an initial energy of 2 Joules. Whenever a sensor transmits or receives a data packet, it consumes some energy from its battery. The base station has an unlimited amount of energy. The choice of the MAC protocol can completely dominate energy consumptions. We assume that energy-conscious protocols like PAMAS [26] or TDMA-based MAC [6] are used for long-lived sensor networks. Our energy model for the sensors is based on the first order radio models [6], [27] in which a fixed amount of energy is spent in transmitting and receiving a packet in the electronics, and an additional amount proportional to the distance between tow nodes is spent in transmitting a packet. The radios can perform power control and hence use the minimum required energy to reach the intended recipients. Due to attenuation with distance, an energy loss model with d2 ij is used for relatively short distances, where dij is the distance between sensor nodes i and j. More precisely, a radio dissipates Eelec=50 nJ/bit to run. Aggregation versus No-Aggregation: We consider the lifetime of the network without aggregation to be the baseline network lifetime, which is taken as 1. We also define the performance metric L as the ratio of the system lifetime achieved using aggregation to that obtained without using aggregation. We refer to L as the lifetime extension ratio. We performed separate sets of experiments for both 2L and ML aggregation schemes. The results are shown in Table III for the 2L scheme for different values of n and number of MAs M, and in Table IV for the ML scheme for different values of n and number of groups _.As shown in Table III, all schemes with aggregation result in prolonging the lifetime of the sensor network. The value of lifetime extension ratio (L) is the highest with the optimal approach, which can be as large as 5, and sometimes even larger. Out of the group of the approximate approaches, the LBA approach has the best results. However, the GA approach is not very far behind, which makes it a good candidate for use. Table IV shows values for L for different values of LA nodes, n, and when the number of groups.

VI. CONCLUSIONS

In this paper, we studied the maximum lifetime data gathering and routing problem in WSNs. We showed that cluster-based algorithms along with data aggregation and in network processing can achieve significant energy savings in WSNs. This has a direct effect on prolonging the network lifetime. In particular, we developed GRASS (Grid-based Routing and Aggregator Selection Scheme), a scheme for WSNs that combines the ideas of fixed cluster-based routing together with application-specific data aggregation in order to enhance the wireless sensor network performance in terms of extending the network lifetime, while incurring acceptable levels of latency under data aggregation. Within GRASS, we have presented optimal as well as heuristic algorithms that solve the joint problem of optimal routing with data aggregation for the sake of maximizing the network lifetime. Our results show that, when compared to other approaches in the literature, the proposed scheme is able to improve the network lifetime while incurring acceptable levels of latency and without sacrificing quality. Hence, GRASS can attain the energy and latency efficiency needed for wireless sensor networks.

REFERENCES

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cyirci, *Wireless sensor networks: A survey*, Computer Networks, Vol. 38, No. 4, Page(s): 393-422, 2002.

[2] J. N. Al-Karaki, A. E. Kamal, *Routing Techniques in Wireless Sensor Networks: A Survey*, IEEE wireless communications, Volume 11, No. 6, Dec. 2004, pp. 6–28.

[3] Y. Xu, J. Heidemann, and D. Estrin,*Geography-informed energy conservation for ad hoc routing*, Proc. of MobiCom 2001, pp. 70–84.

[4] C. Intanagonwiwat, R. Govindan, and D. Estrin, *"Directed diffusion: a scalable and robust communication paradigm for sensor networks,"* Proc. of ACM MobiCom '00, Boston, MA, 2000, pp. 56-67.

[5] B. Krishnamachari, D. Estrin, and S. Wicker, *Impact of data aggregation in wireless sensor networks*, International Workshop on Distributed Event-Based Systems, Vienna, Austria, July 2002.

[6] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, *Energy efficient communication protocol for wireless microsensor networks,* proceedings of HICSS'00, Jan. 2000.

[7] S. Lindsey, C. Raghavendra, *PEGASIS: power efficient gathering in sensor information systems*, IEEE Aerospace Conference Proceedings, Vol. 3, Page(s): 1125-1130, 2002.

[8] A. Manjeshwar and D. P. Agarwal, *TEEN: a routing protocol for enhanced efficiency in wireless sensor networks*, In 1st International Workshop on Parallel and Distributed Computing- Issues in Wireless Networks and Mobile Computing, April 2001.