Volume 2, Issue 12, December 2012 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

A Design Approach for Data Collection in Wireless Sensor Networks with a Mobile Base Station

D.B.Shanmugam [1], G.Anitha [2], K.Vinoth Kumar[3], C.Karthi[4], M.Hema[5]
[1] Asst.Prof, Department of MCA, Sri Balaji Chockalingam Engg College., Arni.
[2,3,4,5] Lecturer, Department of MCA, Sri Balaji Chockalingam Engg College., Arni.

Abstract ---Recent research shows that significant energy saving can be achieved in wireless sensor networks with a mobile base station that collects data from sensor nodes via short-range communications. However, a major performance bottleneck of such WSNs is the significantly increased latency in data collection due to the low movement speed of mobile base stations. To address this issue, we propose a rendezvous-based data collection approach in which a subset of nodes serves as the rendezvous points that buffer and aggregate data originated from sources and transfer to the base station when it arrives. This approach combines the advantages of controlled mobility and in-network data caching and can achieve a desirable balance between network energy saving and data collection delay. We propose two efficient rendezvous design algorithms with provable performance bounds for mobile base stations with variable and fixed tracks, respectively. The effectiveness of our approach is validated through both theoretical analysis and extensive simulations.

Keywords--- wireless sensor networks, Steiner Minimum Tree, Mobile Sink First.

I. INTRODUCTION

Wireless sensor networks, designed to monitor and/or control the surrounding environmental phenomena, have the potential to revolutionize many applications. A sensor network consists of sensor nodes and one or more base stations. Sensor nodes generate, process, and forward data (via intermediate sensor nodes) to base stations. Among the major design challenges in the design of sensor networks is the efficient utilization of resources available to sensor nodes such as scarce bandwidth and limited energy supply. Energy efficiency and bandwidth utilization should be improved at all layers of the communication protocol stack. Addressing these issues at the routing layer is very important due to the significant amount of energy consumed in transmitting high volume of data generated by the sensor nodes. Energy is a paramount concern to wireless sensor networks (WSNs) that must operate for an extended period of time on limited power supplies such as batteries. A major portion of energy expenditure of WSNs is attributed to multi-hop wireless communications. Recent research has exploited controlled mobility as a promising approach to reduce communication energy consumption of WSNs. For instance, a mobile base station (BS) may roam about a sensing field and collect data from sensor nodes through short-range communications. The energy consumption of static nodes is thus reduced because fewer numbers of wireless relays are needed in the network

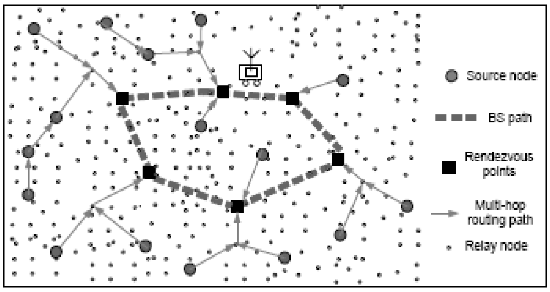

Figure 1: An example of data collection in a 500 × 500 m2 sensing field. The BS moves at 0.5 m/s. It takes the BS about 20 minutes to visit all rendezvous points located within 100 m from the center of field. It takes more than 2 hours to visit 100 source nodes randomly distributed in the field.

The major performance bottleneck of WSNs with a mobile BS is the increased latency in data collection. The typical speed of practical mobile sensor systems , is about 0.1 − 2 m/s. As a result, it takes a mobile BS hour to tour a large sensing field, which cannot meet the delay requirements of many sensing applications. The low movement speed is a fundamental design constraint for mobile BSs because increasing the speed will lead to significantly higher manufacturing cost and power consumption. For instance, the power consumption of the Packbot node is about 60W when moving at 1 m/s, and increases quadratic ally with speed . In this paper, we propose a data collection approach that explores the controlled mobility of BS and the capability of in-network data caching. Specifically, a subset of static nodes in the network will serve as the rendezvous points (RPs) and aggregate data originated from sources.

The BS periodically visits each RP and picks up the cached data. An example of rendezvous based data collection. This approach has several key advantages. First, a broad range of desirable tradeoffs between energy consumption and communication delay can be achieved by jointly optimizing the choices of RPs, motion path of BS and data transmission routes. Second, the use of RPs enables the BS to collect a large volume of data at a time without traveling a long distance, which mitigates the negative impact of slow speed of BS on overall network throughput. Third, mobile nodes communicate with the rest of the network through RPs at scheduled times, which minimizes the disruption to the network topology caused by mobility. This paper makes the following contributions.

1. We formulate the rendezvous design problem for WSNs with a mobile BS, which aims to find a set of RPs that can be visited by the BS within a required delay while the network cost incurred in transmitting data from sources to RPs is minimized.
2. We develop two efficient rendezvous design algorithms with constant approximation ratios. The first algorithm places RPs on an approximate Steiner Minimum Tree (SMT) of source nodes, which allows the data to be efficiently aggregated at RPs while shortening the data collection tour of BS. The second algorithm is designed for mobile BSs that must move along fixed tracks. Based on the analysis on the optimal structure of connection between sources and a fixed track, we can find efficient RPs within bounded BS tour on the track.
3. Simulation results show that both algorithms can achieve satisfactory performance under a range of settings. The theoretical performance bounds of the algorithms are also validated through simulations. The rendezvous design problems with a variable and fixed BS track are studied. respectively.

II. BASIC APPROACH AND ASSUMPTIONS

In this section, we first provide a brief overview of the problem, and then introduce the network model used in this paper.

In our problem, a set of source nodes generate data samples that must be delivered to the base station (BS) within time interval D. Our objective is to find a tour of the BS that visits a set of nodes referred to as rendezvous points (RPs). The RPs cache the data originated from sources and send to the BS via short-range transmissions when it arrives. The total energy consumption incurred by the network to transmit the data from sources to the RPs should be minimized under the constraint that all data must be delivered to the BS before the deadline D. An important characteristic of this problem is that the BS tour and data transmission routes must be jointly designed in order to find the optimal RP locations. We refer to this problem as rendezvous design in data collection.

The delay bound may be imposed for two different reasons. First, applications often require data to be delivered within certain deadline. For instance, a user may issue the following sliding-window query: “sample seismic data every 10s and archive at the base station every 10 minutes", where the deadline is 10 minutes. Second, the delay bound may also be imposed due to the recharging cycle of the BS. For instance, the battery of Robomote node lasts for about 30 minutes during movement. Although a mobile BS can periodically replenish its energy (e.g., by moving to a fixed docking station), frequent battery recharging should be avoided to reduce the disruptions to normal operation of the network.

III. NETWORK MODEL

According to several empirical studies the speed that data packets are relayed in a WSN is about several hundred meters per second, which is much higher than the speed that a mobile device moves. Therefore, the data collection deadline can be mapped to the maximum allowable length of the BS tour that visits all RPs. We denote the maximum length of the BS tour, L = vmD, where D is a data collection deadline and vm is the average movement speed of BS. We assume that data from different sources can be aggregated at a node before being relayed. Data aggregation has been widely adopted by data collection applications to reduce network traffic. Specifically, we assume the N-to-one aggregation model in which a node can aggregate multiple data packets it received into one packet before relaying it. Such a model is applicable to a number of scenarios such as collecting the maximum or average value of samples from different sensors. We assume that nodes are densely deployed in a region and all nodes use the same transmission power. Accordingly, the total energy consumed by transmitting a data packet along a multi-hop path is proportional to the Euclidean distance between sender and receiver. This assumption is justified by the fact that the Euclidean distance between two nodes in a dense wireless network is approximately proportional to the hop count between the same nodes.

We note that such an energy model is also adopted by several existing power-efficient data communication protocols in WSNs . This assumption also allows the BS to estimate the network energy consumption without knowing the global network topology. We assume that the storage capacity of a node is large enough to buffer the total volume of data generated by the sources within delivery deadline D. Several recent sensor network platforms can integrate. 10 ∼ 100 Mb NAND flash memory with ultra-low power consumption. Finally, nodes and the BS are assumed to know their own physical locations through the GPS units on them or a location service in the network.

IV. OVERVIEW OF THE APPROACH

We investigate two rendezvous design problems in this paper. In the first problem, the BS may freely move within the network deployment region. In the second problem, the motion of the BS is constrained on a fixed track.

Although such limited mobility reduces the contacts with fixed nodes in a network, it significantly simplifies the motion control of BS and improves the system reliability. For instance, several mobile sensor systems (e.g., XYZ and NIMs) are designed to move along fixed cables.

For each rendezvous design problem, we develop an approximation algorithm that is executed by the BS to find a data collection tour, a set of RPs on the tour, and a set of routing trees that are rooted at the RPs and connect all sources. As we assume that the BS does not have the global information about the network except the locations of sources, the RPs found are physical locations at which there may not exist real nodes.

This issue can be addressed in the following two ways. First, the BS may find a real node near each RP through the network. For instance, it may send an area any cast message addressed to the physical location of an RP. The message will be delivered to a node in the vicinity of the intended location, which may serve as the RP. Alternatively, the BS may travel along the calculated tour and recruit nodes to serve as RPs.

V. TWO-PHASE COMMUNICATION PROTOCOL

In this section, we develop a centralized communication protocol to support the MASP data collection scheme.

Concept Summary

Considering the strong computation capabilities of the mobile sink, we let most of the highly resource-consuming computations be executed by the mobile sink in the communication protocol that consists of two main phases: discover phase and data collection phase.

Discover Phase

The main tasks of the discover phase include learning the topology information and assigning the members to their corresponding subsinks. To complete the tasks, the discover phase is performed through three different rounds described below where the "round" has the same meaning as the earlier "round" described.

1. Round 1. In this round, the mobile sink transmits broadcast messages continuously. All nodes receiving the broadcast messages from the mobile sinks are automatically selected as subsinks. Then the subsinks start building the shortest path trees (SPTs) rooted from themselves in entire network. As a result, each node obtains the shortest hop information from themselves to all subsinks and then send the related hop information to the corresponding subsink. The latter will transmit the hop information to the mobile sink in Round 2. Another important task in Round 1 is that the mobile sink needs to record the time when each node enters and leaves its communication range. For the mobile sink, the data collection processes in the forward direction and the reverse direction are symmetrical. So only the time records in the forward direction are needed.

2. Round 2. In this round, the subsinks send the shortest hop information collected in Round 1 to the mobile sink when it passes by. In some areas with very dense deployment of sensor nodes, the communication durations of the subsinks may overlap in the case that more than one subsink is located simultaneously within the communication range of the mobile sink. Different overlapping time partition or time slot allocation methods do not theoretically change the total amount of data collected by the mobile sink but may indirectly affect the energy consumption of entire network. The mobile sink calculates the length of the communication time allocated to each subsink according to some rules that will be discussed. Here, we get the shortest hop matrix and the Mreq information that is necessary for the MASP calculation.

.

3. Round 3. In Round 3, the mobile sink traverses the trajectory again to broadcast the results of member assignment to the monitored area. The broadcast message consists of the list of the mapping relation between each member and its destination subsink. Each node receiving the broadcast message will get a subsink as its destination. Then the node will delete its own item in the broadcast message and rebroadcast it. Finally, the optimized member assignment information will be disseminated to the entire network.

VI. DATA COLLECTION PHASE

In this phase, all nodes start collecting data from the monitored area formally. The members send the sensed data or forward data to the destination subsinks according to the routing table built in Round 1 of the discover phase. To deal with the network dynamics caused by the node failure or node addition, existing on-demand routing protocols, may be used to find the closest valid subsink as the temporary destination for one node, when it can not reach its subsink successfully. Subsinks precache all data from their members and themselves before the mobile sink enters into their communication range. During the actual data collection, we adopt a handoff method to partition the overlapping time which is consistent with the one used in Round 2. In order to load balance the data originated from members, a round bin scheme is used to transmit data at the subsinks. In this paper, the constraints on the network throughput are expressed by the number of members required by the subsinks. However, the mobile sink may not collect the expected amount of data due to interference between transmission and reception on the subsinks. The communication resources (including time, bandwidth, etc.) between the subsinks and the mobile sink are more precious than those in a network where complete paths between the member nodes and the subsinks always exist. Based on this observation, we propose a Mobile Sink First (MSF) scheme to process the communication between the subsinks and the mobile sink in order to avoid the interference. In the MSF scheme, the subsink will stop receiving the sensed information from its downstream nodes and make use of all time resource and bandwidth resource to transmit data to the mobile sink when it's current subsink's turn for transmission. In the meanwhile, the downstream nodes of current subsink need to buffer data from their child nodes. After the mobile sink moves away, the downstream nodes will start transmitting buffered data to the current subsink. Next, we will analyze the feasibility of the MSF scheme by explaining the data buffered on the child nodes can be received by the subsink without affecting normal data collection.

VII. DISTRIBUTED SOLUTION

In earlier sections, we propose a centralized solution to the MASP optimization problem based on a Genetic Algorithm and a corresponding communication protocol to implement the MASP, in which three movement rounds are needed for the mobile sinks to perform the discover phase. Although most of the time-consuming and energy-consuming tasks are executed by the mobile sink, more movement rounds in the discover phase lead to a lower efficiency of total system. To address this issue, we propose a distributed approximate solution in which the mapping relationships between the members and the subsinks are calculated by each member independently, instead of by the mobile sink in a centralized way.

First, the sensor nodes close to the trajectory are selected as the subsinks and several shortest path trees rooted from all the subsinks are established. At the same time, the mobile sink records the time when each subsink enters and leaves its communication range and calculates the length of time allocated to each subsink and the MReqs information according to the overlapping time partitioning methods. In next round, the mobile sink broadcasts the result of Mreqs calculation to the subsinks that continues to disseminate into the entire network. Finally, each member node knows the shortest hop information from itself to all subsinks and the number of members required by each subsink. Then Algorithm 3 is executed by each member to choose a subsink. In Algorithm 3, a weight value α. is introduced to calculate the priority of each subsink, which reflects the probability of being selected as the final destination. Clearly, the weight value is the most important parameter affecting the system performance. Small values of α. denote that the hop information has a higher priority than Mreqs in choosing the subsink, which may not maximize the total amount of data but bring less energy consumption. we will conduct some simulation experiments to show the impacts of different value of α. .

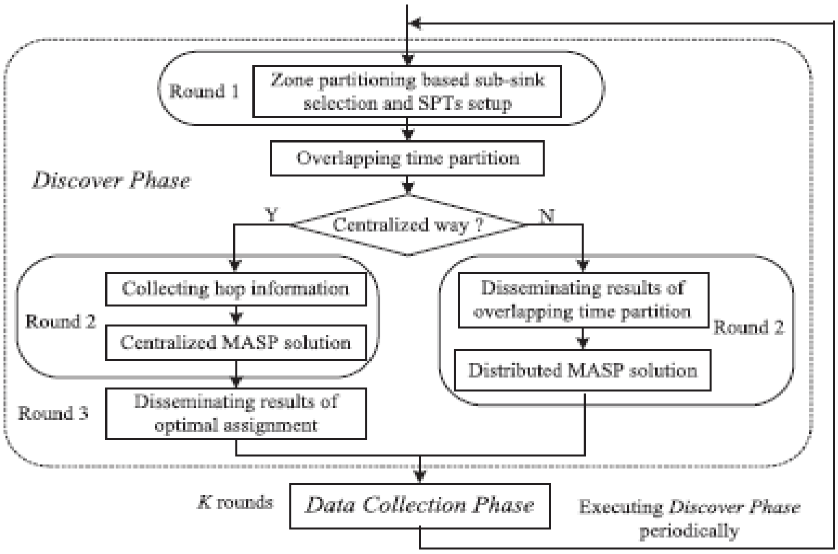

Fig.2.Flowchart of the proposed two-phase communication protocol.

For convenience, the above distributed solution is called MASP-D (MASP-Distributed), which needs only two rounds to perform the discover phase and can get higher efficiency in calculating the optimal assignment compared with the earlier centralized solution of MASP. MASP-D adopts simple rules to choose the subsinks instead of complex genetic algorithm. To deal with the network dynamics caused by node failure or node addition, MASP-D can be re executed to obtain a new available subsink independently on specific nodes.

There are several effective schemes designed to improve the performance and efficiency of the proposed communication protocol. In order to provide a macroscopic insight about how those schemes work together, Fig.2 summarizes the two-phase communication protocol in which our proposed schemes can be integrated and work together according to actual demands, like the number of zones, different overlapping time partition methods, centralized MASP or distributed MASP. In general, the discover phase in the centralized MASP solution consists of three movement rounds of the mobile sink while the one in MASPD has only two rounds. And the discover phase is periodically re executed to deal with the dynamic topology change every K rounds in the data collection phase.

Algorithm 3. Distributed solution of MASP for each member i

Input: h_{ij} and $r_j^m (j = 1, \ldots, n_s)$, weight $\alpha \in [0, 1]$
Initialize: $x_{ij} = 0 (j = 1, \ldots, n_s)$, denoting the result of assignments; temporary variable $v_s = 0$
Main:
  $\forall j$, calculate, $Pr_{ij} = \alpha * r_j^m + (1 - \alpha)/h_{ij}$
  Let v_t be a temporary rand number, $v_t \sim [0, \sum_j^{n_s} Pr_{ij}]$
  for $j = 1$ *to* n_s
    $v_s = v_s + Pr_j$
    if $v_t \leq v_s$
      break;
    endif
  endloop
  $x_{ij} = 1;$

VIII. CONCLUSION AND FUTURE WORK

In this paper, we proposed an efficient data collection scheme called MASP for wireless sensor networks with path-constrained mobile sinks. In MASP, the mapping between sensor nodes and subsinks is optimized to maximize the amount of data collected by mobile sinks and also balance the energy consumption. MASP has good scalability to support sensor networks with low density and multiple mobile sinks. A heuristic based on genetic algorithm and local search is presented to solve the MASP optimization problem.

In addition, we design a communication protocol that supports MASP and adapts to dynamic topology changes. To reduce the computational complexity, we develop two practical algorithms, a zone partitioning-based solution and a distributed solution (MASP-D). We also study the impacts of different overlapping time partitioning methods and present an optimal partitioning scheme (OptShareOT). Simulation experiments under OMNET++ shows that MASP improves the energy utilization efficiency and outperforms SPT and static sink methods in terms of total amount of data with almost the same energy consumption. For future work, we plan to validate the proposed schemes on different scenarios with various movement trajectories of mobile sinks. Considering that minimizing the total energy consumption may not lead to the maximum network lifetime, we also plan to study the subsink selection problem with network lifetime maximization as the optimization objective as future work.

REFERENCES

[1] S. Arora. Polynomial time approximation schemes for euclidean traveling salesman and other geometric problems. J. ACM, 45(5), 1998.
[2] M. A. Batalin, M. Rahimi, Y. Yu, D. Liu, A. Kansal, G. S. Sukhatme, W. J. Kaiser, M. Hansen, G. J. Pottie, M. Srivastava, and D. Estrin. Call and response: experiments in sampling the environment. In SenSys, 2004.
[3] E. E. Y. G. D. Bozdag. Mobility-based communication in wireless sensor networks. IEEE Communications Magazine, 44(7), 2006.
[4] D. J. Chang and E. K. Morlok. Vehicle speed profiles to minimize work and fuel consumption. Journal of Transportation Engineering, 131(3), 2005.
[5] O. Chipara, Z. He, G. Xing, Q. Chen, X. Wang, C. Lu J. Stankovic, and T. Abdelzaher. Real-time power-aware routing in sensor networks. In IWQoS, 2006.
[6] K. Dantu, M. Rahimi, H. Shah, S. Babel, A. Dhariwal, and G. S. Sukhatme. Robomote: enabling mobility in sensor networks. In IPSN, 2005.
[7] S. R. Gandham, M. Dawande, R. Prakash, and S. Venkatesan. Energy efficient schemes for wireless sensor networks with multiple mobile base stations. In Globecom, 2003.
[8] R.C. Shah, S. Roy, S. Jain, and W. Brunette, "Data MULEs: Modeling a Three-Tier Architecture for Sparse Sensor Networks," Proc. First IEEE Int'l Workshop Sensor Network Protocols and Applications, pp. 30-41, 2003.
[9] S. Jain, R.C. Shah, W. Brunette, G. Borriello, and S. Roy, "Exploiting Mobility for Energy Efficient Data Collection in Sensor Networks," Mobile Networks and Applications, vol. 11, no. 3, pp. 327-339, 2006.
[10] A. Chakrabarti, A. Sabharwal, and B. Aazhang, "Communication Power Optimization in a Sensor Network with a Path-Constrained Mobile Observer," ACM Trans. Sensor Networks, vol. 2, no. 3, pp. 297-324, Aug. 2006.
[11] L. Song and D. Hatzinakos, "Architecture of Wireless Sensor Networks with Mobile Sinks: Sparsely Deployed Sensors," IEEE Trans. Vehicular Technology, vol. 56, no. 4, pp. 1826-1836, July 2007.
[12] D. Jea, A. Somasundara, and M. Srivastava, "Multiple Controlled Mobile Elements (Data Mules) for Data Collection in Sensor Networks," Proc. First IEEE/ACM Int'l Conf. Distributed Computing in Sensor Systems (DCOSS), pp. 244-257, 2005.

[13] A. Kansal, A. Somasundara, D. Jea, M. Srivastava, and D. Estrin, "Intelligent Fluid Infrastructure for Embedded Networks," Proc.ACM MobiSys, pp. 111-124, 2004.