IOSR Journal of Electronics & Communication Engineering (IOSR-JECE)
ISSN(e) : 2278-1684 ISSN(p) : 2320-334X, PP 41-46
www.iosrjournals.org

Data Collection Capacity in Arbitrary Wireless Sensor Networks

VIDYA.P[1], IEEE MEMBER Ms.Shalini. L[2]

[1]*(MTECH CSE STUDENT,MCET, ANAD, TRIVANDRUM)*

[2]*(ASST.PROFESSOR, MCET, ANAD, TRIVANDRUM)*

ABSTRACT : *Data collection is a fundamental function provided by wireless sensor networks. The performance of wireless network can be measured by how efficiently collect sensing data from all sensor nodes. Previously, the study of data collection capacity has concentrated on large-scale random networks. But in most of the practical sensor applications, the sensor network is not uniformly deployed and the number of sensors may not be as huge as in theory. Therefore, it is necessary to study the capacity of data collection in an arbitrary network. In this paper a simple BFS tree based method is used which can lead to order-optimal performance for any arbitrary sensor networks. We then study the capacity bounds of data collection under a general graph model, where two nearby nodes may be unable to communicate due to barriers or path fading, and discuss performance implications. From this study among this two BFS tree based method find which one is better.*

Keywords - *Capacity, data collection, arbitrary networks, wireless sensor networks.*

1. INTRODUCTION

A wireless sensor network consists of a set of sensor devices which spread over a geographical area. The ultimate goal of sensor networks is often to collect the sensing data from all sensors to a sink node and then perform further analysis at the sink node. Thus, data collection is one of the most common services used in sensor network applications. In this paper, I study some fundamental capacity problems arising from data collection in wireless sensor networks.

I consider a wireless sensor network where *n* sensors are *arbitrarily* deployed in a finite geographical region. Each sensor measures independent field values at regular time intervals and sends these values to a sink node. The union of all sensing values from *n* sensors at a particular time is called a *snapshot*. The task of data collection is to deliver these snapshots to a single sink. The performance of data collection in sensor networks can be characterized by the rate at which sensing data can be collected and transmitted to the sink node. In particular, the theoretical measure that captures the limits of collection processing in sensor networks is the capacity of many-to-one data collection.

Data collection capacity reflects how fast the sink can collect sensing data from all sensors with interference constrain. It is critical to understand the limit of many-to-one information flows and devise efficient data collection algorithms to improve the performance of wireless sensor networks.

Capacity limits of data collection in random wireless sensor networks have been studied in the literature [1]–[6]. In [1], [2], Duarte-Melo *et al.* first introduced the many-to-one transport capacity in dense and random sensor networks under protocol interference model. El Gamal [3] studied the capacity of data collection subject to a total average transmitting power constraint where a node can receive data from multiple source nodes at a time. However, all the above research shares the common assumption where large number of sensor nodes are either located on a grid structure or randomly and uniformly distributed in a plane. Such assumption is useful for simplifying the analysis and deriving nice theoretical limitations, but may be invalid in many practical sensor applications.

Arbitrary network means the sensor network can be deployed in any way and it can form any network topology. For arbitrary sensor network under protocol interference model, this paper proposes a data collection method based on Breadth First Search (BFS) [1] tree and this method can achieve collection capacity [2] of which matches the theoretical upper bound.

Since the disk graph model is idealistic, also consider a *general graph mode* [4]. In the general graph model, two nearby nodes may be unable to communicate due to various reasons such as barriers and path fading. Thus the study shows that a greedy scheduling algorithm on BFS tree can achieve capacity as in the theoretical limits. These two BFS tree based method is compared by analyzing how many time slots are needed to collect all the sensing data at the sink node.

2. RELATED WORKS

Capacity of data collection in random wireless sensor networks has been investigated in [1]–[6]. Duarte-Melo *et al.*[1], [2] first studied the many-to-one transport capacity inrandom sensor networks under protocol interference model.They showed that the overall capacity of data collection is $\Theta(W)$. El Gamal [3] studied data collection capacity subjectto a total average transmitting power constraint. They relaxed the assumption that every node can only receive from one source node at a time. It was shown that the capacity of random networks scales as $\Theta(\log nW)$ when n goes to infinity and the total average power remains fixed.

Their method uses antenna sharing and channel coding. Barton and Zheng [4] also investigated data collection capacity under more complex Recently, Chen *et al.* [6] have studied data collection capacity with multiple sinks. They showed that with k sinks the capacity increases to $\Theta(kW)$ when $k = O(\ n \log n)$ or $\Theta(\ nW \log n)$ when $k = \Omega(\ n \log n)$. Liu *et al.* [5] lately introduced the capacity of a more general some-to-some communication paradigm in random networks where there are $s(n)$ randomly selected sources and $d(n)$ randomly selected destinations. They derived the upper and lower bounds for such a problem.

However, all research above shares the standard assumption that a large number of sensor nodes are either located on a grid structure or randomly and uniformly distributed in a plane. Such an assumption is useful to simplify the analysis and derive nice theoretical limits, but may be invalid in many practical sensor applications.

3. NETWORK MODELS AND COLLECTION CAPACITY

3.1 Basic Network Models

In this paper, I focus on the capacity bound of data collection in arbitrary wireless sensor networks. For simplicity, start with a set of simple and yet general enough models. Later, we will relax them to more realistic models. I consider an arbitrary wireless network with n sensor nodes v_1, v_2,…. , v_n and a single sink v_0. These n sensors are arbitrarily distributed in a field. At regular time intervals, each sensor measures the field value at its position and transmits the value to the sink. Here first adopt a fixed data-rate channel model where each wireless node can transmit at W bits/second over a common wireless channel. Here I will refer to a network where the nodes are randomly placed following a uniform distribution as a randomly deployed network or a random network. In such a network we have no direct control over the exact location of the nodes. I will refer to a network where we can determine the exact locations of the nodes as an arbitrary network. Note that an arbitrary network is thus a particular instance of the random network with a very low probability of occurring. TDMA scheduling is used at MAC layer. Under the fixed data-rate channel model, and assume that every node has a fixed transmission power P. Thus, a fixed transmission range r can be defined such that a node v_j can successfully receive the signal sent by node v_i.

Fig.1.shows the example of sink is situated deterministically at the center of this field. It is the ultimate receiver of all data generated by sources in the network. The effect of positioning the sink closer to the edge of the network is discussed in the next section. Throughout the paper W will refer to the transmission capacity of the channel in a flat network. In a hierarchical network W will refer to the transmission capacity of the channel used within clusters. W 0 will refer to the transmission capacity of the channel used from the heads to the sink.

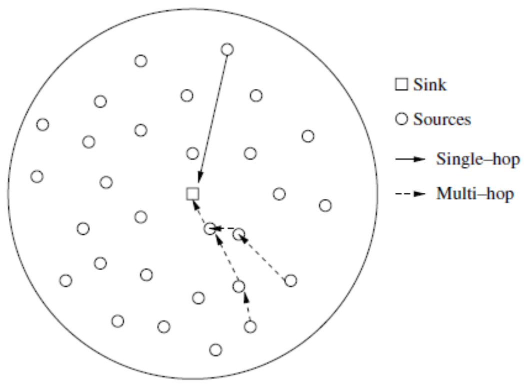

Fig. 1. Many-to-one network scenario.

3.2 CAPACITY OF DATA COLLECTION

In this paper, formally define delay and capacity of data collection in wireless sensor networks. Recall that each sensor generates a field value with *b* bits at regular time intervals, and tries to transport it to the sink. We call the union of all values from all *n* sensors at particular sampling time a *snapshot* of the sensing data. Then the goal of data collection is to collect these snapshots from all sensors. It is clear that the sink prefer to get each snapshot as quickly as possible. In this paper, assume that there is no correlation among all sensing values and no network coding or aggregation technique is used during the data collection.

Definition 1: The delay of data collection *D* is the time used by the sink to successfully receive a snapshot, i.e., the time needed between completely receiving one snapshot and completely receiving the next snapshot at the sink.

Definition 2: The capacity of data collection *C* is the ratio between the size of data in one snapshot and the time to receive such a snapshot (i.e., *nb/D*) at the sink.

4. COLLECTION CAPACITY UNDER DISK GRAPH MODEL

Upper Bound of Collection Capacity: It has been proved that the upper bound of capacity of data collection for random networks is *W* [1], [2]. It is obviously that this upper bound also holds for any arbitrary network. The sink *v*0 cannot receive at rate faster than *W* since *W* is the fixed transmission rate of individual link. Therefore, we are interested in design of data collection algorithm to achieve capacity in the same order of the upper bound, *i.e.* £(*W*). We now propose a BFS-based data collection method and demonstrate that it can achieve the capacity of £(*W*) under our network model. Our data collection method includes two steps: data collection tree formation and data collection scheduling.

4.1 DATA COLLECTION TREE – BFS TREE

The data collection tree used by our method is a classical Breadth First Search (BFS) tree rooted at the sink v_0. The time complexity to construct such a BFS tree is $O(|V|+|E|)$. Let *T* be the BFS tree and $v1\ldots\ldots v_m$ be all leaves in *T*. For each leaf v_m there is a path P_i from itself to the root v_0. Let $\partial Pi\ (vj)$ be the number of nodes on path P_i which are inside the interference range of v_m (including *vm* itself).

4.2 BRANCH SCHEDULING ALGORITHM

Now illustrate how to collect one snapshot from all sensors. Given the collection tree *T*, our scheduling algorithm basically collects data from each path *Pi* in *T* one by one. First, we explain how to schedule collection on a single path. For a given path *Pi*, we can use Δi slots to collect one data in the snapshot at the sink. Each node on the path has unit data to transfer. Links with the same color are active in the same slot. After three slots, the leaf node has no data in this snapshot and the sink got one data from its child. Therefore, to receive all data on the path, at most $\Delta i \times |Pi|$ time slots are needed. This scheduling method is *Path Scheduling*.

Algorithm 1 Branch Scheduling on BFS Tree

Input: BFS tree T.

1: **for** each snapshot **do**
2: **for** $t = 1$ to c **do**
3: Collect data on path P_i. All nodes on P_i transmit data towards the sink v_0 using *Path Scheduling*.
4: The collection terminates when nodes on branch B_i do not have data for this snapshot. The total slots used are at most $\Delta_i \cdot |B_i|$, where $|B_i|$ is the hop length of B_i.
5: **end for**
6: **end for**

4.3 CAPACITY ANALYSIS

Now we can analyze the achievable capacity of our data collection method by counting how many time slots the sink needs to receive all data of one snapshot. The data collection method based on path scheduling in BFS tree can achieve data collection capacity of $\Theta(W)$ at the sink.

In Algorithm 1, the sink collects data from all *c* paths in *T*. In each step (Lines 3-4), data are transferred on path *Pi* and it takes at most $\Delta i\ .\ |Bi|$ time slots. Recall that *Path Scheduling* needs at most $\Delta i\ .k$ time slots to collect *k* packets from path *Pi*. Therefore, the total number of time slots needed for Algorithm 1, denoted by τ. the upper bound of data collection capacity is *W*, thus our data collection algorithm is order-optimal.

5. COLLECTION CAPACITY FOR GENERAL GRAPH MODEL

Now assume that the communication graph is a disk graph where two nodes can communicate if and only if their distance is less than or equal to transmission range *r*. However, a disk graph model is idealistic since in practice two nearby nodes may be unable to communicate due to various reasons such as barriers and path fading. Therefore, in this section, consider a more general graph model $G = (V,E)$ where *V* is the set of sensors and *E* is the set of possible communication links. Every sensor still has a fixed transmission range *r* such that the necessary condition for *vj* to receive correctly the signal from vi.

A new greedy-based scheduling algorithm is showed which is inspired by [19]. The scheduling algorithm still uses the BFS tree as the collection tree. All messages will be sent along the branch towards the sink *v*0. For *n* messages from one snapshot, it works as follows. In every time slot, it sends each message along the BFS tree from the current node to its parent, without creating interference with any higher-priority message. The priority *ρi* of each packet *pi* is defined as 1*l*(*vi*) . It is clear that packets originated from the children of the sink have the highest priority *ρi* = 1 while packets originated from other nodes have lower priority *ρi* < 1. For two packets with the same priority (on the same level in the BFS tree), ties can be broken arbitrarily. Given a schedule, let v^j_m be the node of packet *pj* in the end of time slot *τ*. The detailed greedy algorithm is given in Algorithm 2.

Algorithm 2 Greedy Scheduling on BFS Tree

Input: BFS tree T.

Compute the priority $\rho_i = 1/l(v_i)$ of each message p_i.
for each snapshot **do**
  while $\exists p_j$ such that $v_j^\tau \neq v_0$ **do**
    for all such p_i in decreasing order of priority ρ_i **do**
      if sending p_i from node v_i^τ will not create interference with any higher-priority messages that are already scheduled for this time slot **then**
        node v_i^τ sends p_i to its parent $par(v_i^\tau)$ in T.
      end if
    end for
    $\tau = \tau + 1$.
  end while
end for

6. DISCUSSIONS

Here used two BFS tree based method which gives order optimal performance. In BFS using branch scheduling it requires less number of time slots to receive all snapshots from all sensor nodes in the network. But in Greedy scheduling method the snapshot will be send according to the priority and it requires more number of time slots than branch scheduling. Thus the data collection capacity of this type of arbitrary network is higher by using the branch scheduling algorithms.

When compared to other methods for analyze the data collection capacity in any network models these BFS tree based method will give order optimal performance.

7. CONCLUSION

In this paper, study the theoretical limitations of data collection in terms of capacity for arbitrary wireless sensor networks. First propose an efficient data collection method to achieve capacity which is order-optimal under protocol interference model. However, when the underlying network model is a general graph, prove that BFS-based method can still achieve capacity of for general graphs. All of our methods can also achieve these results for random networks.

Among these proposed methods ranch scheduling algorithm will give better performance when compared to the other one. These proposed methods can lead to order optimal performance in a arbitrary sensor networks.

REFERENCES

[1] E.J. Duarte-Melo and M. Liu, "Data-gathering wireless sensor networks:Organization and capacity," *Computer Networks*, 43, 519-537, 2003.

[2] D. Marco, E.J. Duarte-Melo, M. Liu, and D.L. Neuhoff, "On the manyto-one transport capacity of a dense wireless sensor network and the compressibility of its data," in *Proc. Int'l Workshop on Information Processing in Sensor Networks*, 2003

[3] H.E. Gamal, "On the scaling laws of dense wireless sensor networks:the data gathering channel," *IEEE Trans. on Information Theory*, vol.51, no. 3, pp. 1229–1234, 2005.

[4] B. Liu, D. Towsley, and A. Swami, "Data gathering capacity of large scale multihop wireless networks," in *Proc. of IEEE MASS*, 2008.

[5] S. Chen, Y. Wang, X.-Y. Li, X. Shi, "Order-optimal data collection inwireless sensor networks: Delay capacity," in *IEEE SECON*, 2009.

[6] X.-Y. Li, J. Zhao, Y.W. Wu, S.J. Tang, X.H. Xu, X.F. Mao, "Broadcastcapacity for wireless ad hoc networks," in *IEEE MASS*, 2008.

[7] www.google.com