

Dynamic Forwarding over Tree-on-DAG for Scalable Data Aggregation in Sensor Networks

Kai-Wei Fan, Sha Liu, and Prasun Sinha

Abstract—Computing and maintaining network structures for efficient data aggregation incurs high overhead for dynamic events where the set of nodes sensing an event changes with time. Moreover, structured approaches are sensitive to the waiting time that is used by nodes to wait for packets from their children before forwarding the packet to the sink. Although structureless approaches can address these issues, the performance does not scale well with the network size. We propose *Tree on DAG (ToD)*, a semistructured approach that uses *Dynamic Forwarding* on an implicitly constructed structure composed of multiple shortest path trees to support network scalability. The key principle behind ToD is that adjacent nodes in a graph will have low stretch in one of these trees in ToD, thus resulting in early aggregation of packets. Based on simulations on a 2,000-node network and real experiments on a 105-node Mica2-based network, we conclude that efficient aggregation in large-scale networks can be achieved by our semistructured approach.

Index Terms—Data aggregation, dynamic forwarding, ToD.

1 INTRODUCTION

DATA aggregation is an effective technique for conserving energy in sensor networks. Due to inherent redundancy in raw data collected from sensors, in-network data aggregation can often reduce the communication cost by eliminating redundancy and forwarding only the extracted information from the raw data. As reducing communication energy consumption extends the network lifetime, it is critical for sensor networks to support in-network data aggregation.

Various structured approaches for data aggregation have been proposed for data gathering applications and event-based applications [1], [2], [3], [4], [5], [6]. In data gathering applications, such as environment and habitat monitoring [7], [8], [9], [10], nodes periodically report the sensed data to the sink. As the traffic pattern is unchanging, these structure-based approaches incur low maintenance overhead and are, therefore, suitable for such applications. However, in event-based applications, such as intrusion detection [11], [12] and biological hazard detection [13], the source nodes are not known in advance. Therefore, the approaches that use fixed structures cannot efficiently aggregate data, while the approaches that change the structure dynamically incur high maintenance overhead [3]. *The goal of this paper is to design a scalable and efficient data aggregation protocol that incurs low maintenance overhead and is suited for event-based applications.*

Realizing the shortcomings of structured approaches, an anycast-based structureless data aggregation protocol [14] is proposed. It involves mechanisms to increase the chance of packets meeting at the same node (Spatial Aggregation) at the same time (Temporal Aggregation). As the approach does not guarantee aggregation of all packets, the cost of

forwarding unaggregated packets increases with the scale of the network.

To benefit from the strengths of structured and structureless approaches, we propose a semistructured approach in [15]. In this paper, we further extend its ability to support irregular topology networks. The main challenge in designing such a protocol is to determine the packet forwarding strategy in absence of a preconstructed global structure to achieve early aggregation. Our approach uses a structureless technique locally, followed by *Dynamic Forwarding* on *Tree on DAG (ToD)*, an implicitly constructed packet forwarding structure to support network scalability. After performing local aggregation, nodes dynamically decide the forwarding tree based on the location of the sources. The key principle behind ToD is that adjacent nodes in a graph will have low stretch in at least one of these trees in ToD, thus resulting in early aggregation of packets. This paper makes the following contributions:

- We propose an efficient and scalable data aggregation mechanism that can achieve early aggregation without incurring overhead of constructing a structure.
- We implement the *ToD* on TinyOS and compare it against other approaches on a 105-node sensor network.
- For studying the scalability aspects of our approach, we implement ToD in the *ns2* simulator and study its performance in networks of up to 2,000 nodes.

The organization of this paper is as follows: Section 2 presents background and related work. Section 3 presents the semistructure approach. Section 4 analyzes the performance of ToD in the worst case. The performance evaluation of the protocols using experiments and simulations is presented in Section 5. Finally, Section 6 concludes this paper.

2 RELATED WORK

Data aggregation has been an active research area in sensor networks for its ability to reduce energy consumption.

• The authors are with the Department of Computer Science and Engineering, The Ohio State University, 395 Dreesse Lab, 2015 Neil Ave., Columbus, OH 43210.
E-mail: {fank, liusha, prasun}@cse.ohio-state.edu.

Manuscript received 5 Sept. 2007; revised 3 Mar. 2008; accepted 13 Mar. 2008; published online 2 Apr. 2008.

For information on obtaining reprints of this article, please send e-mail to: tmc@computer.org, and reference IEEECS Log Number TMC-2007-09-0267. Digital Object Identifier no. 10.1109/TMC.2008.55.

Some works focus on how to aggregate data from different nodes [16], [17], [18], some focus on how to construct and maintain a structure to facilitate data aggregation [1], [2], [3], [4], [5], [6], [19], [20], [21], [22], [23], and some focus on how to efficiently compress and aggregate data by taking the correlation of data into consideration [23], [24], [25], [26], [27]. As our work focuses on how to facilitate data aggregation without incurring the overhead of constructing a structure, we briefly describe the structure-based as well as structureless approaches in current research.

In [6], the authors propose an aggregation tree construction algorithm to simultaneously approximate the optimum trees for all nondecreasing and concave aggregation functions. The algorithm uses a simple min-cost perfect matching to construct the tree. Other works, such as Steiner Minimum Tree (SMT) and Multiple Shared Tree (MST) for multicast algorithms which can be used in data aggregation [22], [23], [28], build a structure in advance for data aggregation. In addition to their complexity and overhead, they are only suitable for networks where the sources are known in advance. Therefore, they are not suitable for networks with mobile events.

In addition, fixed tree structures also have the long stretch problem. A stretch of two nodes u and v in a tree T on a graph G is the ratio between the distance from node u to v in T and their distance in G . Long stretch implies that packets from adjacent nodes have to be forwarded many hops away before aggregation. This problem has been studied as Minimum Stretch Spanning Tree (MSST) [29] and Minimum Average Stretch Spanning Tree (MAST) [30]. It has been shown that for any graph, the lower bound of the average stretch is $O(\log(n))$ [30], and is $O(n)$ for the worst case [31].

To achieve efficient aggregation for dynamic event scenarios, DCTC is proposed [3]. DCTC assumes that the distance to the event is known to each sensor and uses the node near the center of the event as the root to construct and maintain the aggregation tree dynamically. However, it involves heavy message exchanges that might offset the benefit of aggregation in large-scale networks. From the simulation results in [3], the energy consumption of tree expansion, pruning, and reconfiguration in DCTC is about 33 percent of the data collection.

Data Aware Anycast (DAA) [14] is the first proposed structureless data aggregation protocol that can achieve high aggregation without incurring the overhead of structure approaches. DAA uses anycast to forward packets to one-hop neighbors that have packets for aggregation. It can efficiently aggregate packets near the sources and effectively reduce the number of transmissions. However, it does not guarantee the aggregation of all packets. As the network grows, the cost of forwarding packets that were unable to be aggregated will negate the benefit of energy saving resulted from eliminating the control overhead.

In order to get benefit from structureless approaches even in large networks, scalability has to be considered in the design of the aggregation protocol. In this paper, we propose a scalable semistructured protocol, ToD, that can achieve efficient aggregation even in large networks.

3 SCALABLE DATA AGGREGATION

The goal of our protocol is to achieve aggregation of data near the sources without explicitly constructing a structure

for mobile event scenarios. Aggregating packets near the sources is critical for reducing the number of transmissions. Aggregating without using an explicit structure reduces the overhead of construction and maintenance of the structure. In this section, we propose a highly scalable approach that is suitable for large sensor networks.

Our protocol has two phases, DAA [14] and *Dynamic Forwarding*. In the first phase, packets are forwarded and aggregated using DAA. Since DAA does not guarantee that all packets will be aggregated, in the second phase, the leftover unaggregated or partially aggregated packets are forwarded using *Dynamic Forwarding* on a structure, ToD, for further aggregation. First, we briefly describe the DAA protocol in this section. For detail protocol description, please refer to [14].

3.1 Data Aware Anycast

DAA [14] is a structureless protocol that aggregates packets by improving the *spatial* and *temporal* convergence. Spatial convergence and temporal convergence are two necessary conditions for aggregation. Packets have to be transmitted to the same node at the same time to be aggregated. Without explicit message exchanges in structureless aggregation, nodes do not know where they should send packets to and how long they should wait for aggregation. Therefore, improving spatial or temporal convergence is critical for improving the chance of aggregation.

Spatial convergence is achieved by using anycast to forward packets to nodes that can achieve aggregation. Anycast is a routing scheme whereby packets are forwarded to the best one, or any one, of a group of target destinations based on some routing metrics. By exploiting the nature of wireless radio transmission where all nodes within the transmission range can receive the packet, nodes are able to tell if they currently have packets that can be aggregated with the transmitting packet. Anycast mechanism allows the sender to forward packets to any one of these nodes that have packets for aggregation.

Temporal convergence is used to further improve the aggregation. In mobile event-triggered networks, nodes are unable to know which nodes are triggered and have packets to transmit in advance. Therefore, nodes cannot know if they should wait for their upstream nodes and how long they should wait for aggregation. A naive approach of using a fixed delay depending on the distance to the sink may make the detection delay very high. *Randomized Waiting* is a simple technique for improving temporal convergence, in which nodes wait for a random delay before transmitting their packets. It has been shown that the randomized waiting can efficiently increase the chance of aggregation in [14].

With DAA, aggregation can always happen in the network as long as two neighboring nodes have packets that can be aggregated. Packets are forwarded in a "random walk" fashion with one constraint that they are forwarded to nodes having packets for aggregation. When packets reach a node that does not have any neighbor that has packets for aggregation, they will be forwarded toward the sink through geographic routing, but may return to DAA if they reach a node that can achieve further aggregation by DAA.

3.2 Dynamic Forwarding over ToD

In DAA [14], packets were destined to the sink when no further aggregation can be achieved locally. This incurs

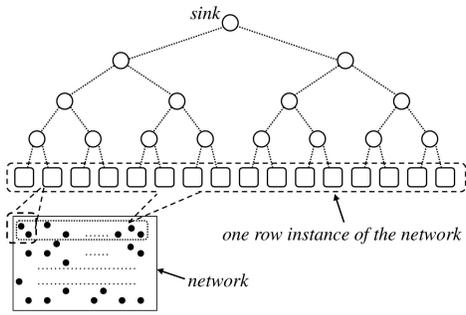


Fig. 1. We illustrate the ToD construction from one row's point of view to simplify the discussion.

higher transmission cost when the distance from sources to the sink is longer. To remedy this problem, we forward packets on a structure when no further aggregation can be achieved instead of forwarding them to the sink directly. Constructing a structure dynamically with explicit message exchanges incurs high overhead. Therefore, we use an implicitly computed structure to forward packets. However, using a fixed structure has the long stretch problem as described in Section 2. Thus, we design a dynamic forwarding mechanism over *ToD* to avoid the long stretch problem.

3.2.1 ToD in One Dimensional Networks

For illustrating the concept of ToD, we first describe the construction of ToD for a 1D (a single row of nodes) network, as shown in Fig. 1.

First, we define a cell as a square with side length Δ , where Δ is greater than the maximum diameter of the area an event can span. Sensing devices equipped on sensor nodes, such as PIR, magnetometer, and acoustic sensors, usually have maximum sensing range for a type of events [32]. The density of gas, biochemical substance, and

radioisotope of nuclear radiation degrades with distance and can be detected within certain range from the source [33]. Therefore, for a specific application, the maximum event size can be determined.

The network is divided into cells. These cells are grouped into clusters, called *F-clusters* (first-level clusters). The size of the *F-clusters* must be large enough to cover the cells an event can span, which is two when we only consider 1D cells in the network. All nodes in *F-clusters* send their packets to their cluster-heads, called *F-aggregators*. Nodes in the *F-cluster* can be multiple hops away from the *F-aggregator*. The formation of the clusters and the election of the aggregators are discussed later in Section 3.2.3. Each *F-aggregator* then creates a shortest path to the sink. Therefore, the structure is a shortest path tree (SPT), where the root is the sink and the leaves are *F-aggregators*. We call this tree an *F-Tree*. Fig. 2a shows the construction of the *F-Tree*.

In addition to the *F-clusters*, we create the second type of clusters, *S-clusters* (second-level clusters) for these cells. The size of an *S-cluster* must also be large enough to cover all cells spanned by an event, and it must interleave with the *F-clusters* so it can cover adjacent cells in different *F-clusters*. Each *S-cluster* also has a cluster-head, *S-aggregator*, for aggregating packets. Each *S-aggregator* creates a shortest path to the sink, and forms a second SPT in the network. We call it *S-Tree*. The illustration of an *S-Tree* is shown in Fig. 2b. For all sets of nearby cells that can be triggered by an event, either they will be in the same *F-cluster*, or they will be in the same *S-cluster*. This property is exploited by *Dynamic Forwarding* to avoid the long stretch problem discussed earlier.

After the *S-Tree* is constructed, the *F-aggregators* connect themselves to the *S-aggregators* of *S-clusters* which its *F-cluster* overlaps with, as shown in Fig. 2c. For example, in Fig. 2c, the *F-aggregator* *F4* connects to *S-aggregators* *S3* and *S4* because its *F-cluster* overlaps with *S-clusters* 3 and 4. Thus, the combination of *F-Tree*

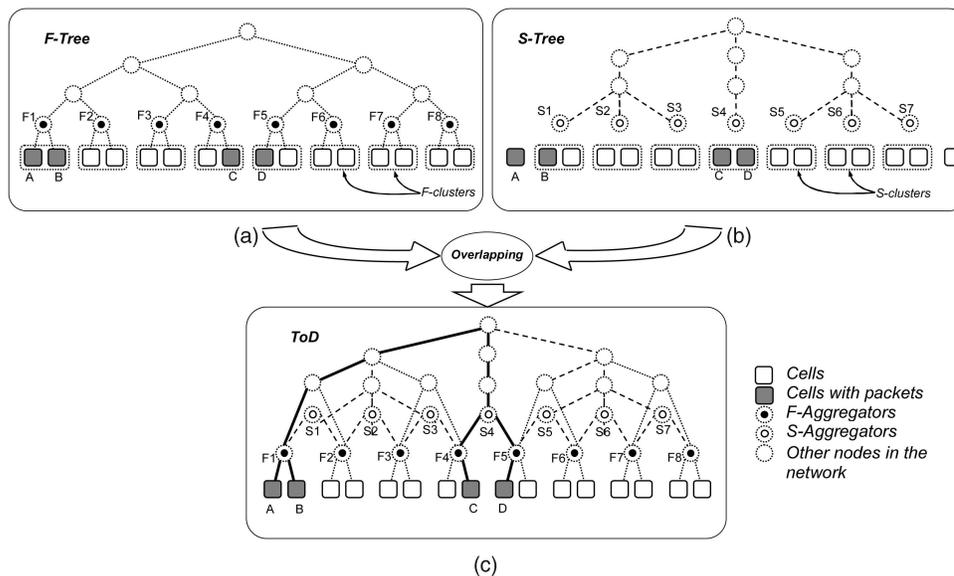


Fig. 2. The construction of *F-Tree*, *S-Tree*, and *ToD*. (a) Leaf nodes are cells. Pairs of neighbor cells define *F-clusters*. Each *F-cluster* has an *F-aggregator*, and *F-aggregators* form the *F-Tree*. (b) Each pair of adjacent cells not in the same *F-cluster* form an *S-cluster*. Each *S-cluster* has an *S-aggregator*, and *S-aggregators* form the *S-Tree*. Nodes on the network boundary do not need to be in any *S-cluster*. (c) *ToD* is created by connecting each *F-aggregator* to two *S-aggregators* of *S-clusters* which its *F-cluster* overlaps with. *F-aggregator* in *ToD* uses *Dynamic Forwarding* to forward packets to the root, or through an *S-aggregator* in the *S-Tree* based on where the packets come from.

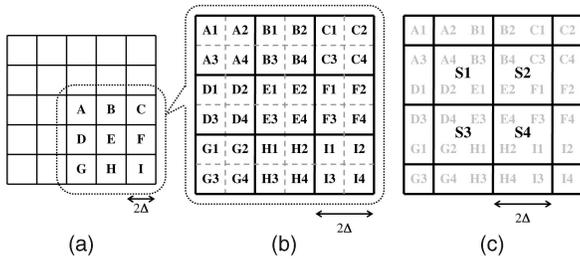


Fig. 3. Grid clustering for a 2D network. (a) The network is divided into 5×5 F-clusters. (b) Each F-cluster contains four cells. For example, the F-cluster *A* in (a) contains cell *A1*, *A2*, *A3*, and *A4*. (c) The S-clusters have to cover all adjacent cells in different F-clusters. Each S-cluster contains four cells from four different F-clusters.

and S-Tree creates a Directed Acyclic Graph, which we refer to as the ToD.

Nodes first use the DAA to aggregate as many packets as possible. When no further aggregation can be achieved, nodes forward their packets to their F-aggregators through geographic routing for further aggregation. If an event only triggers nodes within a single F-cluster, its packets can be aggregated at the F-aggregator, and be forwarded to the sink using the F-Tree. However, in case the event spans multiple F-clusters, the corresponding packets will be forwarded to different F-aggregators. As we assumed that the event size is not larger than the size of a cell, an event on the boundary of F-clusters will only trigger nodes in cells on the boundary of the F-clusters. By the construction of S-clusters, adjacent cells on the boundary of F-clusters belong to the same S-cluster. Thus, F-aggregators can exploit the information collected from received packets to select the S-aggregator that is best suited for further aggregation.

Consider the example in Fig. 2c. If the event spans *A* and *B*, *F1* knows that no other F-cluster will have packets for aggregation since the maximum number of cells an event can span is two; hence, it can forward the packets using the F-Tree. If the event spans two cells in two different F-clusters, for example, *C* and *D*, *F4* will only receive packets from *C*, and *F5* will only receive packets from *D*. *F4* can know either the event happens only in *C*, or it spans *D* as well. Consequently, *F4* can forward packets to *S4*, the S-aggregator of its overlapped S-clusters covering *C*, and so is *F5*. Therefore, these packets can be aggregated at *S4*.

Note that we do not specifically assign cells on the boundary of the network to any S-cluster. They do not need to be in any S-cluster if they are not adjacent to any other F-cluster, or they can be assigned to the same S-cluster as its adjacent cell.

The ToD for 1D network has the following property.

Property 1. For any two adjacent nodes in ToD in 1D network, their packets will be aggregated either at an F-aggregator, or will be aggregated at an S-aggregator.

Proof. See the Appendix. \square

3.2.2 ToD in Two Dimensional Networks

Section 3.2.1 only demonstrates the construction for one row of nodes to illustrate the basic idea of dynamic forwarding, and it works because each cell is only adjacent to one (or none, if the cell is on the boundary of the network) of the

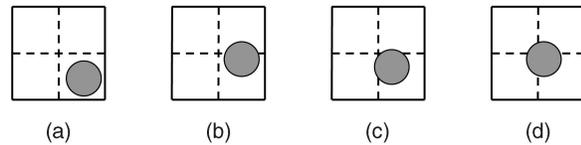


Fig. 4. The possible numbers of cells an event may span in 2×2 cells, which are (a) one, (b) two, (c) three, and (d) four. The four cells in each case are any instance of four cells in the network. They may be in the same F-cluster or different F-clusters.

F-clusters. Therefore, if an event spans two cells, the two cells are either in the same F-cluster or in the same S-cluster, and the F-aggregator can conjecture whether to forward the packets to the S-aggregator, or to the sink directly. When we consider 2D scenarios, a cell on the boundary of an F-cluster might be adjacent to multiple F-clusters. If an event spans multiple F-clusters, each F-aggregator may have multiple choices of S-aggregators. If these F-aggregators select different S-aggregators, their packets will not be aggregated. However, the ideas presented in 1D networks can be extended for 2D networks. But instead of guaranteeing that packets will be aggregated within two steps as in the 1D case (aggregating either at an F-aggregator or an S-aggregator), the ToD in 2D guarantees that the packets can be aggregated within three steps.

We first define the cells and clusters in two dimensions. For the ease of understanding, we use grid clustering to illustrate the construction. As defined before, the size of a cell is not less than the maximum size of an event, and an F cluster must cover all the cells that an event might span, which is four cells in 2D grid clustering. Therefore, the entire network is divided into F-clusters, and each F-cluster contains four cells. The S-clusters have to cover all adjacent cells in different F-clusters. Each F-cluster and S-cluster also has a cluster-head acting as the aggregator to aggregate packets. Fig. 3 shows a 5×5 network with its F-clusters and S-clusters.

Since the size of a cell (one side of the square cell) must be greater or equal to the maximum size of an event (diameter of the event), an event can span only one, two, three, or four cells as illustrated in Figs. 4a, 4b, 4c, and 4d. If an event only spans cells in the same F-cluster, packets can be aggregated at the F-aggregator. Therefore, we only consider scenarios where an event spans cells in multiple F-clusters.

Fig. 5 shows four basic scenarios that an F-aggregator may encounter when collecting all packets generated in its F-cluster. All other scenarios are only different combinations of these four scenarios. If packets originate from three or four cells in the same F-cluster, the F-aggregator knows

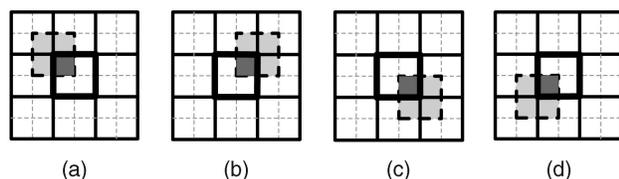


Fig. 5. All possible scenarios in an F-aggregator's point of view. Each case shows 3×3 F-clusters, and the aggregator of the center F-cluster is making the decision. The dark grayed squares are cells that generate packets, and the light grayed squares represent the corresponding S-cluster of the dark grayed cells.

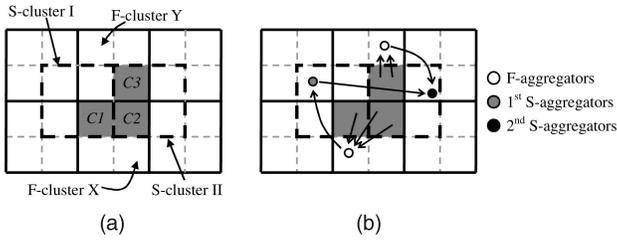


Fig. 6. (a) The F-aggregators have two choices for S-aggregators if they receive packets from two cells. (b) Depending on how many cells generate packets in its F-cluster, one F-aggregator sends packets to two S-aggregators, while the other F-aggregator sends packets to only one S-aggregator. We assume that the sink is located at bottom left of the network.

that no other nodes in other F-clusters have packets, and it can forward the packets directly to the sink. If only one or two cells generate packets, it is possible that other F-clusters also have packets. We assume that the region spanned by an event is contiguous. So, simultaneous occurrences of scenarios of Figs. 5a and 5c, or Figs. 5b and 5d, are impossible in the F-cluster. However, such scenarios are possible in presence of losses in a real environment where packets from third or fourth cluster are lost. In such cases, the F-aggregator can just forward the packets directly to the sink because no other F-cluster will have packets from the same event.

When the F-aggregator collects all packets within its cluster, it knows which cells the packets come from and can forward the packets to best suited S-aggregator for further aggregation. For example, if the packets only come from one cell as in Fig. 5a, the F-aggregator can forward the packet to the S-aggregator of the S-cluster that covers that cell. However, if packets come from two cells in an F-cluster, the two cells must be in different S-clusters. For example, Fig. 6a, where the F-aggregator X (F-aggregator of F-cluster X) receives packets from two cells, is the combination of Figs. 5a and 5b. It is possible that the F-aggregator Y may receive packets from cells as in Figs. 5c and 5d, or both. Since the F-aggregator X does not know which case the F-aggregator Y encounters, it does not know which S-aggregator to forward packets to. To guarantee the aggregation, the F-aggregator X forwards the packet through two S-aggregators that cover cells $C1$ and $C2$; therefore, packets can meet at least at one S-aggregator. If both F-aggregators receive packets from two cells in its cluster, to guarantee that the packets can meet at least at one S-aggregator, these two F-aggregators must select the S-aggregator deterministically. The strategy is to select the S-aggregator that is closer to the sink. If the packets meet at the first S-aggregator, it does not need to forward packets to the second S-aggregator. The S-aggregator only forwards packets to the second S-aggregator if the packets it received only come from two cells in one F-cluster. We will present a simplified construction later (in Section 3.2.3) for the selection of S-aggregators.

To guarantee that the packets can meet at least at one S-aggregator, the second S-aggregator must wait longer than the first S-aggregator. Therefore, if the S-aggregator receives packets from only one cell, it waits longer to wait for possible packets forwarded by the other S-aggregator because it could be the second S-aggregator of the other F-aggregator. Fig. 6b shows an example of one F-aggregator

sending packets to the first S-aggregator and then the second S-aggregator, while the other F-aggregator sends packets directly to the second S-aggregator. As long as the second S-aggregator waits sufficiently longer than the first S-aggregator, the packets can be aggregated at the second S-aggregator.

The ToD for 2D networks has the following property.

Property 2. For any two adjacent nodes in ToD, their packets will be aggregated at the F-aggregator, at the first S-aggregator, or at the second S-aggregator.

Proof. See the Appendix. \square

Although in this section we assume that the size of an event is smaller than the size of a cell, our approach can still work correctly and perform more efficiently than DAA even if the size of the event is not known in advance. This is because the nodes will use Dynamic Forwarding over ToD only at second phase where the aggregation by DAA is no longer achievable. Therefore, at worst, our approach just falls back to DAA. Section 5.1 shows that in experiments, ToD improves the performance of DAA by 27 percent even if the size of the event is greater than the size of a cell.

3.2.3 Clustering and Aggregator Selection

In this paper, we use grid clustering to construct the cells and clusters. Although other clustering methods, such as clustering based on hexagonal or triangular tessellation, can also be used, we do not explore them further in this paper. In principle, any clustering would work as long as they satisfy the following conditions. First, the size of a cell is greater than or equal to the maximum size of an event. Second, the F-cluster and S-cluster must cover the cells that an event may span, and the S-cluster must cover the adjacent cells in different F-clusters.

As opposed to defining an arbitrary clustering, using grid clustering has two advantages. First, the size of the grid can be easily determined by configuring the grid size as a network parameter. Second, as long as the geographic location is known to the node, the cell, F-cluster, and S-cluster it belongs to can be determined immediately without any communication. Geographic information is essential in sensor networks; therefore, we assume that sensor nodes know their physical location by configuration at deployment, a GPS device, or localization protocols [34], [35]. As a consequence, all the cells, F-clusters, and S-clusters can be implicitly constructed.

After the grids are constructed, nodes in an F-cluster and S-cluster have to select an aggregator for their cluster. Because the node that acts as the aggregator consumes more energy than other nodes, nodes should play the role of aggregator in turn in order to evenly distribute the energy consumption among all nodes. Therefore, the aggregator selection process must be performed periodically. However, the frequency of updating the aggregator can be very low, from once in several hours to once in several days, depending on the capacity of the battery on the nodes. Nodes can elect themselves as the cluster-head with probability based on metrics such as the residual energy, and advertise to all nodes in its cluster. In case two nodes select themselves as the cluster-head, the node ID can be used to break the tie.

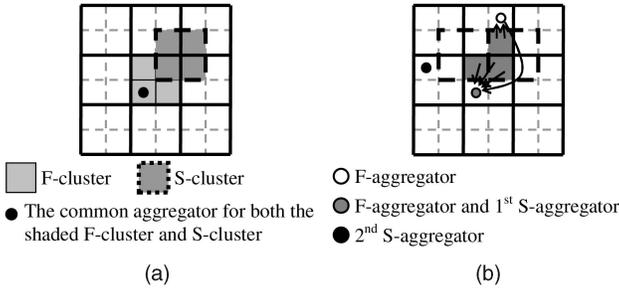


Fig. 7. (a) The S-cluster selects the F-cluster closest to the sink among its overlapped F-clusters, assuming that the sink is located at the bottom-left corner of the network. (b) The white F-aggregator selects the F-cluster containing the gray F-aggregator as the aggregating cluster.

The other approach is that nodes use a hash function to hash current time to a node in their cluster as the aggregator. Nodes have to know the address of all nodes in its F-cluster and sort them by their node id. A hash function hashes the current time to a number k from 1 to n , where n is the number of nodes in its cluster, and nodes use the k th node as the aggregator. Because the frequency of changing the aggregator could be low, the time used could be in hours or days; therefore, the time only needs to be coarsely synchronized, and the cluster-head election overhead can be avoided.

However, the *Dynamic Forwarding* approach requires that each F-aggregator knows the location of S-aggregators of S-clusters that its F-cluster overlaps with. To simplify the cluster-head selection process and avoid the overhead of propagating the update information, we delegate the role of S-aggregators to F-aggregators. We choose an F-cluster, called *Aggregating Cluster*, for each S-cluster, and use the F-aggregator of the *Aggregating Cluster* as its S-aggregator. The *Aggregating Cluster* of an S-cluster is the F-cluster which is closest to the sink among all F-clusters that the S-cluster overlaps with, as shown in Fig. 7a, assuming that the sink is located on the bottom-left corner. When an F-aggregator forwards packets to an S-aggregator, it forwards them toward the aggregating cluster of that S-aggregator. When packets reach the aggregating cluster, nodes in that F-cluster know the location of their F-aggregator and can forward packets to it.

Algorithm 1: Aggregating clusters selection

```

// Cells of sources can be represented by bit-mask of four
// bits. Topleft cell is bit 1, topright cell is bit 2,
// bottomleft is bit 3 and bottomright cell is bit 4.
1: procedure AggregatingCluster(Packet p)
2:  $cells \leftarrow$  cells of sources in  $p$ 
3: if  $cells = topleft$  then
4:    $1st\_fcluster \leftarrow 2nd\_fcluster \leftarrow$  left F-cluster
5: else if  $cells = topright$  then
6:    $1st\_fcluster \leftarrow 2nd\_fcluster \leftarrow$  my F-cluster
7: else if  $cells = bottomleft$  then
8:    $1st\_fcluster \leftarrow 2nd\_fcluster \leftarrow$  bottom-left F-cluster
9: else if  $cells = bottomright$  then
10:   $1st\_fcluster \leftarrow 2nd\_fcluster \leftarrow$  bottom F-cluster
11: else if  $cells = (topleft | topright)$  then
12:   $1st\_fcluster \leftarrow$  my F-cluster
13:   $2nd\_fcluster \leftarrow$  left F-cluster
14: else if  $cells = (bottomleft | bottomright)$  then

```

```

15:   $1st\_fcluster \leftarrow$  bottom F-cluster
16:   $2nd\_fcluster \leftarrow$  bottom-left F-cluster
17: else if  $cells = (topleft | bottomleft)$  then
18:   $1st\_fcluster \leftarrow$  left F-cluster
19:   $2nd\_fcluster \leftarrow$  bottom-left F-cluster
20: else if  $cells = (topright | bottomright)$  then
21:   $1st\_fcluster \leftarrow$  my F-cluster
22:   $2nd\_fcluster \leftarrow$  bottom F-cluster
23: else
24:   $dst \leftarrow$  sink
25: end if

```

Algorithm 1 shows the pseudocode for selecting the first and second aggregating clusters based on where the packets come from. The pseudocode only shows the simplified procedure and does not show the aggregating cluster selection code when the F-cluster is located at the boundary of the network, where the left, bottom, or bottom-left F-cluster does not exist. If the next aggregating cluster is the left or bottom F-cluster that does not exist, the F-aggregator will choose itself as the next aggregator. If the next aggregating cluster is the bottom-left F-cluster that does not exist, the F-aggregator chooses the bottom or left F-cluster if either of them exists, or selects itself if none of them exists. For a general algorithm irrespective of sink location, please refer to our technical report [36].

Now the role of S-aggregators is passed on to the F-aggregators, and the F-cluster selected by an S-aggregator is the one closer to the sink. When an F-aggregator wants to forward packets to both S-aggregators, it selects the F-cluster that is closer to itself as the aggregating cluster of the first S-aggregator (could be itself) to reduce the number of transmissions between aggregators, as shown in Fig. 7b. This selection does not affect the property that packets will eventually be aggregated at one aggregator because the S-clusters that cover the cells in two F-clusters are the same; therefore, the selected aggregating clusters will be the same.

The benefits of using this approach are fivefold. First, no leader election is required for S-clusters, which eliminates the leader election overhead. Second, nodes only need to know the F-aggregator of their F-clusters, which is very scalable. Third, when the F-aggregator changes, the change does not need to be propagated to other F-clusters. Fourth, if nodes choose the aggregator by hashing current time to get a node ID of the aggregator in its cluster, only nodes within the same F-cluster need to be synchronized with each other. And last, since the *Aggregating Clusters* of S-clusters are statically computed, there is no overhead for computing the *Aggregating Clusters*.

3.2.4 ToD in Irregular Topology Networks

In *ToD*, F-aggregators forward packets to their *Aggregating Clusters* using *Dynamic Forwarding* rules. These aggregating clusters are selected implicitly based on their relative locations to F-aggregators. However, in real deployment, the deployed field may not be fully covered by sensors because of obstacles or randomness of deployment. These uncovered regions are referred to as voids. If an F-aggregator selects an aggregating cluster residing in a void, its packets cannot be forwarded for further aggregation. To address this problem, the *Dynamic Forwarding* rules must take voids into consideration. If the selected

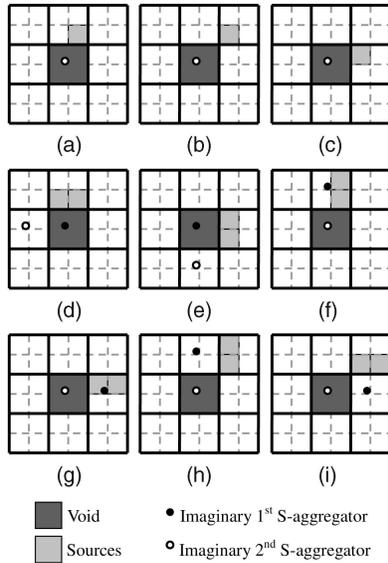


Fig. 8. The possible scenarios for an F-aggregator selecting an aggregating cluster within a void. Case 1: (a) to (c). (a) and (c) Are isomorphic. Case 2: (d) and (e). (d) and (e) Are isomorphic. Case 3: (f) to (i). (f) and (g), and (h) and (i) Are isomorphic, respectively.

aggregating cluster is located within a void, an alternate cluster should be used as a substitute.

In this section, we assume that nodes know if there are nodes in their eight neighboring F-clusters because the aggregating clusters selected by an F-cluster are always its adjacent F-clusters. This can be achieved by periodic *hello* message. After deployment, each node broadcasts a beacon containing its F-cluster ID and an 8-bit vector (initially all 0) indicating which neighboring F-cluster has nodes. If a node receives a beacon from a neighboring F-cluster, it updates the 8-bit vector and propagates this information to all nodes in its F-cluster. Therefore, nodes can know whether there are nodes in their neighboring F-clusters.

Also, we assume that the voids do not split nodes in one F-cluster. This guarantees that nodes in one F-cluster can communicate with each other without routing through nodes in other F-clusters. Although *ToD* still works if nodes in one F-cluster are split into two or more connected components, packets may be aggregated at several aggregators which limits the chance of further aggregation.

The only scenarios that the dynamic forwarding rules must be modified are when the selected aggregating cluster is within a void. There are three possibilities. The first one is when an F-aggregator receives packets from only one of its cells where only one aggregating cluster will be selected, as case 1 shown in Figs. 8a, 8b, and 8c. The second and third possibilities are when an F-aggregator receives packets from two of its cells where two aggregating clusters will be selected, and either the first one, which is closer to the source F-cluster (case 2, Figs. 8d and 8e), or the second one, which is farther away from the source F-cluster (case 3, Figs. 8f, 8g, 8h, and 8i) is in a void. For case 2 where the first aggregating cluster is within a void, the F-aggregator can send the packets directly to the sink because no other F-clusters will have packets (the only F-cluster that the event may span is within the void). Therefore, we only discuss cases 1 and 3.

For cases 1 and 3, if the selected aggregating cluster is within a void, F-aggregators will use the top-right F-cluster

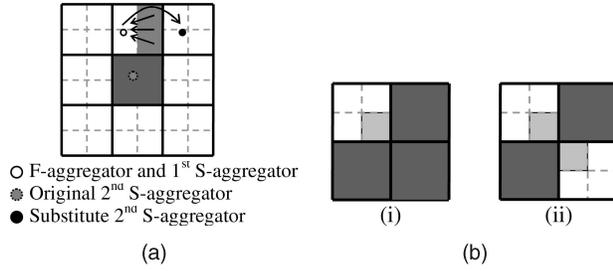


Fig. 9. (a) If the selected aggregating cluster is within a void, the top-right F-cluster will be selected as the substitute aggregating cluster. (b) The scenarios where the selected aggregating cluster and its substitute aggregating cluster are both in a void.

of the original aggregating cluster as the substitute (assume that the sink is at the bottom-left of the network), as shown in Fig. 9a. After the packets are forwarded to the F-aggregator which is also the first S-aggregator, they are supposed to be forwarded to the original second S-aggregator. Since the second aggregating cluster is within a void, the F-aggregator will first wait a short period for possible packets from neighbor cells. If it does not receive packets from other cells, the packets will then be forwarded to the top-right F-cluster instead for further aggregation.

If the substitute aggregating cluster is also within a void as in Fig. 9b.i, the packets will be forwarded to the sink directly. However, the scenario in Fig. 9b.ii can happen in reality. In this case, if the two cells with sources are connected directly without routing around the voids, we use one of these two F-clusters, say the bottom-right F-cluster, as the aggregating cluster; otherwise, the packets will be forwarded to the sink because routing around the voids to aggregate packets may consume more energy than just forwarding them to the sink.

Property 3. *The modified dynamic forwarding rules guarantee that packets can be aggregated to one aggregator in the presence of voids if nodes in one F-cluster and in neighboring F-clusters can communicate with each other without routing through other F-clusters.*

Proof. See the Appendix. \square

4 PERFORMANCE ANALYSIS

In this section, we show that the maximum distance between any two adjacent nodes in *ToD* only depends on the cell size, and is independent of the network size. We ignore the cost from the aggregator to the sink since for perfect aggregation, only one packet will be forwarded to the sink from the aggregator; therefore, the cost is comparatively small. Compared to the lower bound $O(\sqrt{n})$ [30] of the grid network, *ToD* can achieve constant factor even in the worst case.

The worst case in *ToD* is illustrated in Fig. 10 where only two adjacent nodes, u and v , in the corner of two different F-clusters generate packets, and their F-aggregators, f_u and f_v , are located at the opposite corner. We assume a dense deployment of sensor nodes; therefore, the distance between two nodes can be transferred to the cost of transmitting a packet between these nodes. Fig. 10 is the worst case since if more nodes are generating packets in one cluster, it will only amortize the cost of sending packets from the F-aggregator to the S-aggregator, and

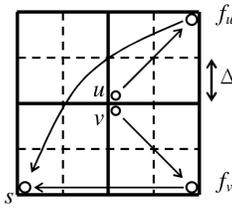


Fig. 10. The worst-case scenario for ToD.

more nodes in multiple F-clusters generating packets will only lower the average distance.

We assume that the length of one side of the cell is Δ , and two nodes are adjacent if their distance is less than a unit of distance. Therefore, in Fig. 10, the distance that packets from u and v have to be forwarded before they are aggregated at s is the sum of distances between u to f_u to s and v to f_v to s , and is $(2\Delta\sqrt{2} + 4\Delta\sqrt{2}) + (2\Delta\sqrt{2} + 4\Delta) = 8\Delta\sqrt{2} + 4\Delta$. Therefore, in the optimal approach, only one transmission is required because u and v are adjacent. In ToD, $8\Delta\sqrt{2} + 4\Delta$ number of transmission is required for the worst case.

However, since we use *DAA* as the aggregation technique, packets from adjacent nodes will be aggregated immediately. Therefore, for the worst cast to happen, the distance between u and v must be at least 2 units, and our protocol has $4\Delta\sqrt{2} + 2\Delta \simeq 7.66\Delta$ times number of transmissions than optimal. The upper bound is only dependent on the size of a cell, and the size of the cell is dependent on the size of an event. This value is independent of the size of the network and, therefore, is very suitable for large-scale networks.

On average, the number of transmissions will be much less than $4\Delta\sqrt{2} + 2\Delta$ because first, typically there will be many nodes generating packets. Second, the distance between a node and its F-aggregator is not always $2\Delta\sqrt{2}$, and the distances between the F-aggregators and the S-aggregator are shorter, too. Third, the *DAA* approach can efficiently aggregate packets from adjacent nodes, thereby further reducing the number of transmissions. Therefore, we expect the average distance for nodes generating packets to be much less than the worst case.

5 PERFORMANCE EVALUATION

In this section, we use experiments and simulations to evaluate the performance of our semistructured approach and compare it with other protocols.

5.1 Testbed Evaluation

We conduct experiments on Kansei sensor testbed [37], [38] to show the advantage and practicability of the ToD approach. The testbed consists of 105 Mica2-based motes and each mote is hooked onto a Stargate. The Stargate is a 32-bit hardware device from CrossBow [39] running Linux. The Stargates are connected to the server using wired Ethernet. Therefore, we can program these motes and send messages and signals to them through Stargates via Ethernet connection. The 105 nodes form a 7×15 grid network with 3-ft spacing. The radio signal using default transmission power covers most nodes in the testbed. In our experiments, we do not change the transmission power but

limit nodes only to receive packets from two-grid neighboring nodes, i.e., each node has a maximum of 12 neighbors.

We implement an Anycast MAC protocol on top of the Mica2 MAC layer. The Anycast MAC layer has its own backoff and retransmission mechanisms and we disable the ACK and backoff of the Mica2 MAC module. An event is emulated by broadcasting a message on the testbed to the Stargates, and the Stargates send the message to the Mica2 nodes through serial port. The message contains a unique ID distinguishing packets generated at different time.

When a node is triggered by an event, an event report is generated. If the node has to delay its transmission, it stores the packet in a report queue. Both the application layer and Anycast MAC layer can access the queue; therefore, they can check if the node has packets for aggregation, or aggregate the received packets to packets in the queue.

First, we evaluate the following protocols¹ on the testbed:

- **Dynamic Forwarding over ToD (ToD).** The semi-structured approach we proposed in this paper. *DAA* is used to aggregate packets in each F-cluster, and aggregated packets are forwarded to the sink on ToD.
- **DAA.** The structureless approach proposed in [14].
- **SPT.** Nodes send packets to the sink through the SPT immediately after sensing an event. Aggregation is opportunistic and happens only if two packets are at the same node at the same time.
- **SPT with Fixed Delay (SPT-D).** Same as the SPT approach, but nodes delay their transmission according to their height in the tree to wait for packets from their children.

Due to the scale of the testbed, we only divide the network into two F-clusters in ToD, which forces the smallest cell to have only nine sensor nodes. However, we do not limit the size of an event to be smaller than the cell size. The event size is larger than the cell size in all following experiments.

We use normalized number of transmissions as the metric to compare the performance of these protocols. The normalized number of transmissions is the average number of transmissions performed in the entire network to deliver one unit of useful information from sources to the sink. It can be converted to the normalized energy consumption if we know the sending and receiving cost of one transmission. Here, we do not consider energy consumption on idle listening since all nodes are fully active for all protocols in the experiments and simulations, and the idle energy consumption would be similar for all protocols. To reduce the energy consumption on idle listening, various duty cycling protocols have been proposed. Due to the page limitation, we are unable to describe how to integrate those works.

Fig. 11 shows the normalized number of transmissions for different event sizes. We fixed the location of the event and vary its diameter from 12 to 36 ft where nodes within two grid-hops to six grid-hops of the event will be triggered, respectively, and send packets to the sink located at one corner of the network. We use 6 seconds as maximum delay for all protocols except SPT. For event size less than 12 ft, there are too little nodes been triggered (less than five), and all triggered nodes are within

1. Codes available at <http://www.cse.ohio-state.edu/~fank/research>.

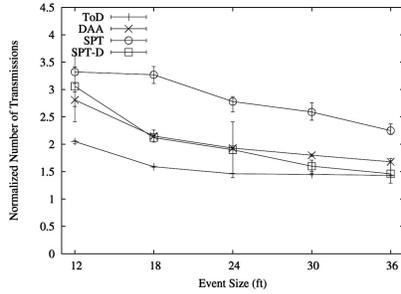


Fig. 11. The normalized number of transmissions for different event sizes from experiments on 105 sensors.

transmission range. Data aggregation is not so interesting in such scenario; therefore, we do not evaluate it.

All protocols have better performance when the size of the event increases because packets have more chances of being aggregated. ToD performs best among all protocols in all scenarios. This shows that DAA can efficiently achieve early aggregation and the Dynamic Forwarding over ToD can effectively reduce the cost of directly forwarding unaggregated packets to the sink in DAA. In SPT-D, when the event size is smaller, the long stretch effect is more significant than in larger event scenario. When event size is large, for example, two-third of nodes in the network are triggered when the diameter of the event is 36 ft, most of the packets can be aggregated to their parent with one transmission. This indicates that in applications where most nodes are transmitting, the fixed structure such as SPT-D is better, but when only a small subset of nodes are transmitting, their performance degrades because of the long stretch problem.

We notice that the variance of some results in SPT and SPT-D is very high. For example, when the event size is 12 ft in diameter, the maximum normalized number of transmissions in SPT-D is 3.41, and the minimum value is 2.41. By tracing into the detail experiment logs, we found that the high variance is because of the different SPTs. The tree is reconstructed for each experiment and, therefore, may change from experiment to experiment. We found that SPT-D always gets better performance in one tree where all sources are under the same subtree, and performs badly in the other tree where sources are located under two or three different subtrees. This further supports our claims that the long stretch problem in fixed structured approaches affects their performance significantly.

The second experiment evaluates the performance of these protocols for different values of maximum delay. We vary the delay from 0 to 8 seconds, and all nodes in the network generate one packet every 10 seconds. Fig. 12 shows the results. As we described, the performance of the structure-based approaches heavily depends on the delay. The SPT-D performs worse than ToD when the maximum delay is less than 5 seconds, and the performance increases as the delay increases. On the contrary, the performance of ToD and DAA does not change for different delays, which is different from results observed in [14]. We believe that this is because with the default transmission power, a large number of nodes are in interference range when nodes transmit. Therefore, even if nodes do not delay their transmissions, only one node can transmit at any given

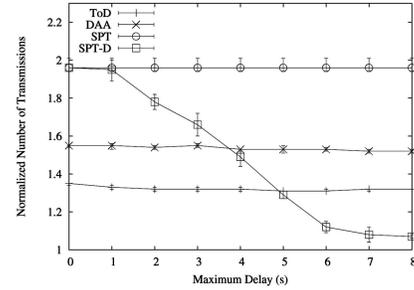


Fig. 12. The normalized number of transmissions for different maximum delays from experiments on 105 sensors.

time. Other nodes will be forced to delay, which has the same effect as the Randomized Waiting.

5.2 Large-Scale Simulation

To evaluate and compare the performance and scalability of ToD with other approaches requires a large sensor network, which is currently unavailable in real experiments. Therefore, we resort to simulations. In this section, we use the *ns2* network simulator to evaluate these protocols. Besides ToD, DAA, and SPT, we evaluate OPT, Optimal Aggregation Tree, to replace the SPT-D protocol.

In OPT, nodes forward their packets on an aggregation tree rooted at the center of the event. Nodes know where to forward packets to and how long to wait. The tree is constructed in advance and changes when the event moves assuming the location and mobility of the event are known. Ideally, only $n - 1$ transmissions are required for n sources. This is the lower bound for any structure; therefore, we use it as the optimal case. This approach is similar to the aggregation tree proposed in [40] but without its tree construction and migration overhead. We do not evaluate SPT-D in simulation in the largest simulation scenario, the network is a 58-hop network. According to the simulation in smaller network, SPT-D gets best performance when the delay of each hop is about 0.64 seconds. This makes nodes closer to the sink have about 36-seconds delay in SPT-D, which is not advisable.

We perform simulations of these protocols on a $2,000 \times 1,200$ m grid network with 35-m node separation; therefore, there are a total of 1,938 nodes in the network. The data rate of the radio is 38.4 Kbps and the transmission range of the nodes is slightly higher than 50 m. An event moves in the network using the random waypoint mobility model at the speed of 10 m/s for 400 seconds. The event size is 400 m in diameter. The nodes triggered by an event will send packets every 5 seconds to the sink located at (0, 0). The aggregation function evaluated here is a perfect aggregation, i.e., all packets can be aggregated into one packet without increasing the packet size.

5.3 Event Size

We first evaluate these protocols on different number of nodes generating the packets. This simulation reflects the performance of each protocol for different event sizes. We study the performance for four mobility scenarios and show the average, maximum, and minimum values of the results.

Fig. 13a shows the result of normalized number of transmissions. ToD improves the performance of DAA and SPT by 30 percent and 85 percent, and is 25 percent higher

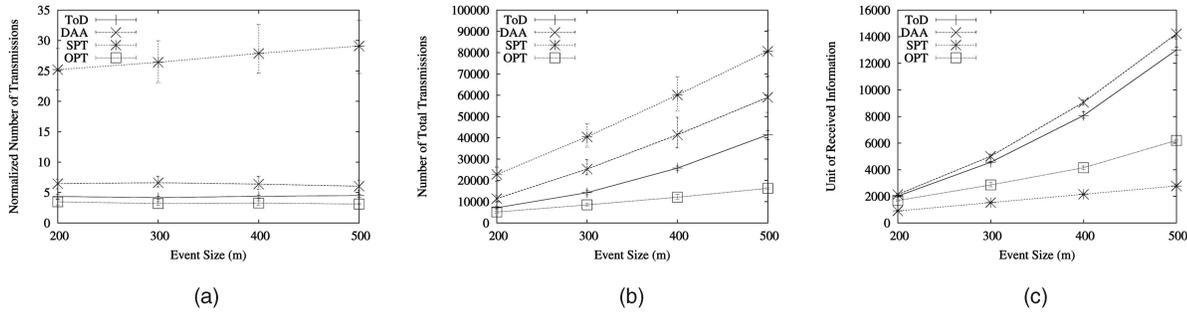


Fig. 13. The simulation results for different event sizes. (a) Normalized number of transmissions. (b) Number of transmissions. (c) Unit of received information.

than OPT. However, OPT has the best performance by using the aggregation tree that keeps changing when event moves, but its overhead is not considered in the simulation. SPT has very poor performance since its aggregation is opportunistic. Except the SPT, the performance of all other protocols is quite steady. This shows that they are quite scalable in terms of the event size.

Figs. 13b and 13c show the total number of transmissions and total units of useful information received by the sink. DAA and ToD have higher number of received packets than OPT due to the ability of structureless aggregation to aggregate packets early and scatter them away from each other to reduce contention. ToD performs better than DAA in terms of the normalized number of transmissions because of its ability to aggregate packets at nodes closer to the source, and thus, it reduces the cost of forwarding packets from sources to the sink. It has slightly lower number of units of received information than DAA. From the simulation logs, we found that most dropped packets in ToD are packets forwarded from sources to their F-aggregators. We believe that the convergecast causes higher contention, thus leading to higher dropping rate.

5.4 Scalability

To evaluate the scalability of a protocol, we limit an event to move only in a bounded region at a certain distance from the sink to simulate the effect of different network sizes. We limit an event to move within a $400 \times 1,200$ m rectangle, and change the distance of the rectangle to the sink from 200 to 1,400 m, as shown in Fig. 14. In order to be fair to all scenarios, we limit the event not to move closer than 200 m to the network boundary such that the number of nodes triggered by the event does not change drastically.

Fig. 15 shows the results of scalability simulations. The performance of ToD and OPT remains steady. This shows that ToD is quite scalable as its performance does not degrade as the size of the network increases. The

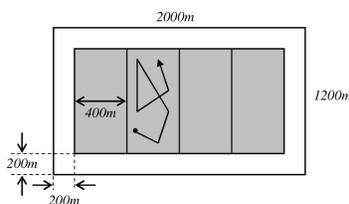


Fig. 14. The simulation scenario for scalability. The event is limited to move only within a small gray rectangle in each simulation.

performance of both DAA and SPT degrades as the size of the network increases. The normalized number of transmissions for DAA and SPT doubled when the event moves from the closest rectangle (to the sink) to the farthest rectangle.

Fig. 15c shows the number of packets received at the sink per event. If all packets can be aggregated near the event and forwarded to the sink, the sink will receive only one packet. Conversely, more packets received at the sink shows that fewer aggregations happened in the network. The cost of forwarding more packets to the sink increases rapidly as the size of the network increases. We can see that, in both DAA and SPT, the sink receives many packets. Although the number of packets received at the sink remains quite steady, the total number of transmissions increases linearly as the distance from the sources to the sink increases.

Ideally, the number of received packets at sink is 1, if all packets can be aggregated at the aggregator. However, the number of received packets at sink is higher than 1 in ToD and OPT. This is because the delay in CSMA-based MAC protocol cannot be accurately predicted; therefore, the aggregator might send the packet to the sink before all packets are forwarded to it. Although the cost of forwarding the unaggregated packets from aggregator to the sink in ToD and OPT also increases when the size of the network increases, the increase is comparably smaller than DAA and SPT because few packets are forwarded to the sink without aggregation. The number of received packets at the sink in ToD is higher when the event is closer to the sink. In ToD, nodes in the same F-cluster as the sink always use sink as the F-aggregator because we assume that the sink is wire powered and there is no need to delegate the role of aggregator to other nodes.

5.5 Aggregation Ratio

In this section, we conduct simulations for different aggregation ratios. Source nodes generate packets with 50 bytes payload. Data are aggregated based on a simple aggregation function where the size of a packet after aggregation is $\max\{50, n \times (1 - \rho)\}$, where n is the number of packets being combined together and ρ is the aggregation ratio. $\rho = 1$ stands for perfect aggregation. The maximum payload of a packet is 400 bytes. Two packets cannot be aggregated if the aggregated size is greater than 400 bytes. As shown in Fig. 16, ToD improves the normalized number of transmissions of DAA, but the improvement decreases as the aggregation ratio decreases. This is because when the aggregation ratio decreases, packet size increases after

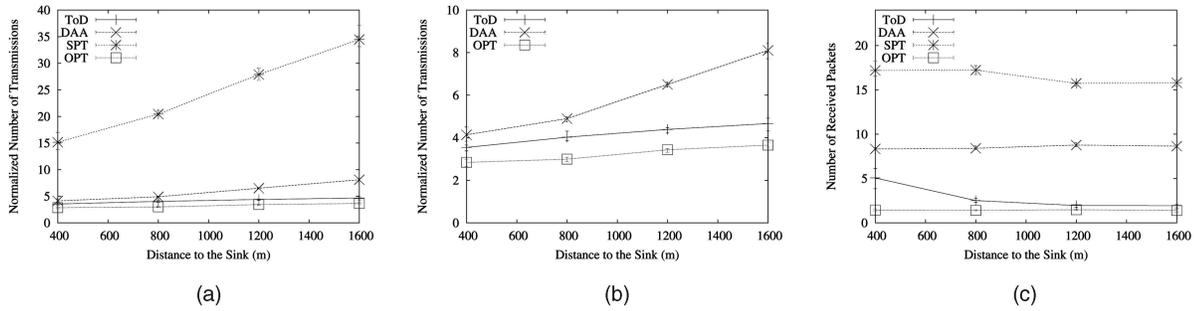


Fig. 15. The simulation results for difference distances from the event to the sink. (a) Normalized number of transmission. (b) Zoom in of (a). (c) Number of received packets.

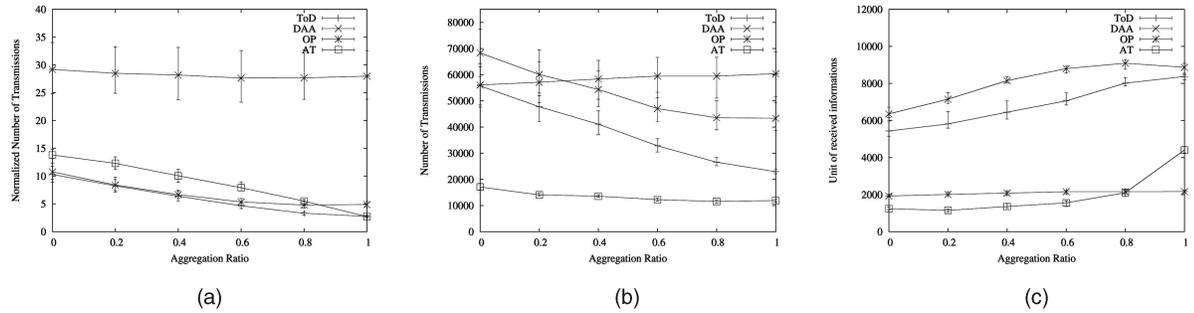


Fig. 16. The simulation results for different aggregation ratio. (a) Normalized number of transmissions. (b) Number of transmissions. (c) Unit of received information.

aggregation. Packets cannot be aggregated anymore when they reach maximum payload even if they meet. Both ToD and DAA perform better than OPT when the aggregation ratio is not 1 because the packet dropping rate in OPT is very high. OPT only receives less than 2,000 units of information, compared to more than 5,000 in ToD and DAA. We believe the high dropping rate is because of the convergecast traffic in OPT. When aggregation ratio decreases, more packets with larger size is forwarded to the root of the aggregation tree, which results in high contention and leads to high dropping rate.

5.6 Cell Size

The above simulations use maximum size of an event as the cell size. This ensures that the Dynamic Forwarding can aggregate all packets at an S-aggregator, and the cost of forwarding the aggregated packets to the sink is minimized. However, large cell size increases the cost of aggregating packets to the aggregator as we use DAA as the aggregation technique in an F-cluster and DAA is not scalable. In this section, we evaluate the impact of the size of a cell on the performance of ToD.

We vary the cell size from 50×50 m to 800×800 m and run simulations for three different event sizes, 200, 400, and 600 m, in diameter. The results are collected from five different event mobility patterns and shown in Fig. 17.

When the size of cell is larger than the event size, the performance is worse because the cost of aggregating packets to F-aggregator increases, but the cost of forwarding packets from S-aggregator does not change. When the size of cell is too small, the cost of forwarding packets to sink increases because packets will be aggregated at different F-aggregators and more packets will be forwarded to the sink without further aggregation. In general, when the size of the F-cluster is small enough to only contain one node, or when the size of the F-cluster

is large enough to include all nodes in the network, ToD just downgrades to DAA.

ToD has the best performance when the cell size is 100×100 m (F-cluster size is 200×200 m) when the event size is 200 m in diameter. When the diameter of an event is 400 and 600 m, using 200×200 m as the cell size has the best performance (F-cluster size is 400×400 m). This shows that the ToD performance can be further optimized by selecting the appropriate cell size. To explore the relation between the event and cell size for optimization will be part of our future work.

5.7 Random Deployment for Irregular Topology

In this section, we evaluated the modified dynamic forwarding rules for irregular topology networks. We create five $1,000 \times 1,000$ m networks and randomly place five circular obstacles with radius ranging from 100 to 200 m, and randomly place 2,000 sensors in these fields. For each deployment, we generate five event moving scenarios as described before. With voids and random deployment, geographic routing may encounter a "local minimum" and has to switch from greedy forwarding to perimeter routing. We implemented the perimeter routing on a GG planar graph [41] to deal with local minimum in greedy forwarding.

To incorporate perimeter routing with anycasting in DAA, when greedy forwarding encounters a local minimum and switches from greedy mode to perimeter mode, the local minimum node specifies the next hop in the perimeter mode in the RTS packet. The next hop node has lower priority to reply a CTS than nodes having packets for aggregation. This allows the packets to be aggregated if neighboring nodes have packets for aggregation even in perimeter mode, and routes packets around the void if they cannot be aggregated.

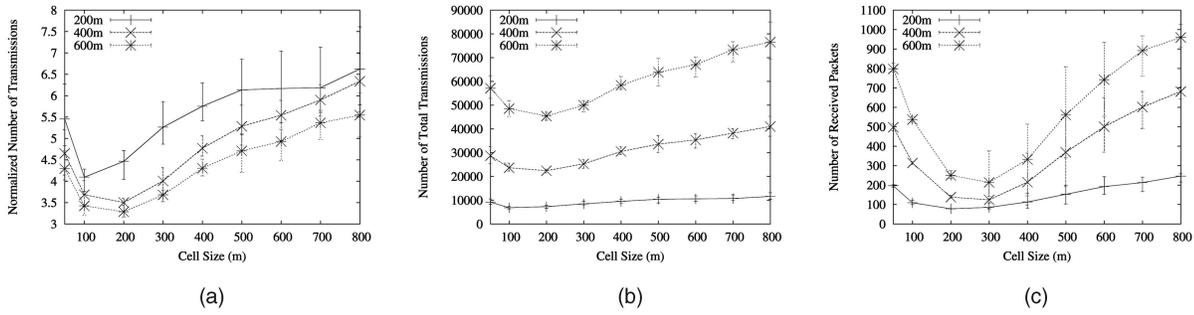


Fig. 17. The simulation results for difference cell sizes. (a) Normalized number of transmissions. (b) Number of transmissions. (c) Number of received packets.

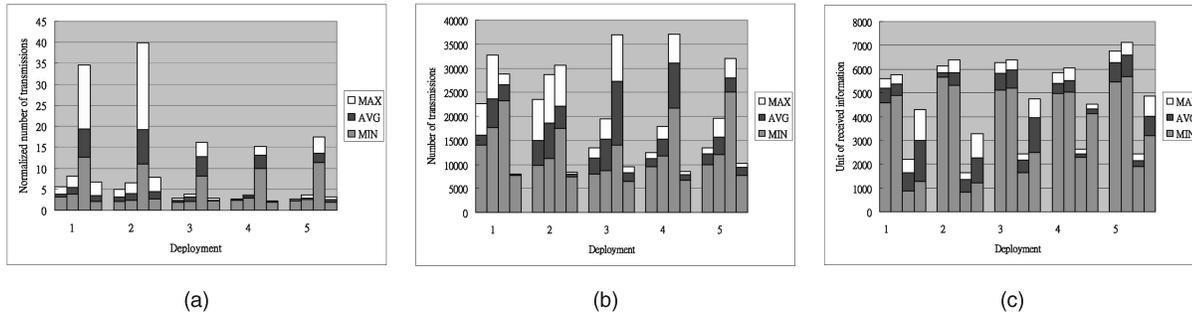


Fig. 18. The simulation results for five different random deployment with randomly generated voids. For each deployment, the results from left to right are ToD, DAA, SPT, and OPT. (a) Normalized number of transmissions. (b) Number of transmissions. (c) Unit of received information.

Fig. 18 shows the results for these five deployments. Due to high variability across different scenarios, we show the results for each scenario rather than averaging over all scenarios. We can see similar results as before. ToD can still improve the normalized number of transmissions compared with DAA and perform close to OPT. ToD uses less number of transmissions than DAA but transmits comparable amount of information.

6 CONCLUSION

In this paper, we have proposed a semistructured approach that locally uses a structureless technique followed by *Dynamic Forwarding* on an implicitly constructed packet forwarding structure, *ToD*, to support network scalability. ToD avoids the long stretch problem in fixed structured approaches and eliminates the overhead of constructing and maintaining dynamic structures. We evaluate its performance using real experiments on a testbed of 105 sensor nodes and simulations on 2,000 node networks. Based on our studies, we find that ToD is highly scalable and it performs close to the optimal structured approach. Therefore, it is very suitable for conserving energy and extending the lifetime of large-scale sensor networks.

APPENDIX

Property 1. For any two adjacent nodes in ToD in 1D network, their packets will be aggregated either at an F-aggregator, or will be aggregated at an S-aggregator.

Proof. There are only three possibilities when an event triggers nodes to generate packets. If only nodes in one cell are triggered and generate the packets, their packets can be aggregated at one F-aggregator since all nodes in

a cell reside in the same F-cluster, and all packets in an F-cluster will be aggregated at the F-aggregator.

If an event triggers nodes in two cells that are in the same F-cluster, the packets can be aggregated at the F-aggregator as well. If an event triggers nodes in two cells that are in different F-clusters, they must be in the same S-cluster because S-clusters and F-clusters are interleaved. Moreover, packets in one F-cluster will only originate from the cell closer to the other F-cluster that also has packets. Therefore, the F-aggregator can forward packets to the S-aggregator accordingly, and packets will be aggregated at the S-aggregator.

Since the cell is not smaller than the maximum size of an event, it is impossible for an event to trigger more than two cells, and this completes the proof. \square

Property 2. For any two adjacent nodes in ToD, their packets will be aggregated at the F-aggregator, at the first S-aggregator, or at the second S-aggregator.

Proof. First, we define the F-aggregator X as the aggregator of F-cluster X and S-aggregator I as the aggregator of S-cluster I , and so forth.

For packets generated only in one F-cluster, their packets can be aggregated at the F-aggregator since all packets in the F-cluster will be sent to the F-aggregator.

If an event triggers nodes in different F-clusters, there are only three cases. First, only one cell in each F-cluster generates packets. In this case, all cells having packets will be in the same S-cluster since the adjacent cells in different F-clusters are all in the same S-cluster. Therefore, their packets can be aggregated at the S-aggregator.

Second, the event spans three cells, C_1 , C_2 , and C_3 , and two of them are in one F-cluster and one of them is in the other F-cluster. WLOG, we assume that C_1 and C_2 are in the same F-cluster, X , and C_3 is in the other

F-cluster, Y . Moreover, $C3$ must be adjacent to either $C1$ or $C2$, and let us assume that it is $C2$. From the ToD construction, we know that $C2$ and $C3$ will be in the same S-cluster, S-cluster II , and $C1$ will be in another S-cluster, S-cluster I . Fig. 6a illustrates one instance of this case. First, the F-aggregator X will aggregate packets from $C1$ and $C2$ because they are in the same F-cluster, and forward the aggregated packets through S-aggregator I to S-aggregator II , or the other way around, because $C1$ is in S-cluster I and $C2$ is in S-cluster II . F-aggregator Y will aggregate packets from $C3$ and forward packets to S-aggregator II because $C3$ is in S-cluster II . Because packets of F-aggregator Y only come from $C3$, they will have longer delay in S-aggregator II in order to wait for packets being forwarded through the other S-aggregator. In the mean time, if F-aggregator X forwards packets to S-aggregator II first, the packets can be aggregated at S-aggregator II . If F-aggregator X forwards packets to S-aggregator I first, S-aggregator I will forward packets to S-aggregator II with shorter delay because the packets come from two cells in one F-cluster; therefore, their packets can also be aggregated at S-aggregator II .

In the third case, the event spans four cells. Two of them will be in one F-cluster and the other two will be in the other F-cluster. WLOG, we assume that cells $C1$ and $C2$ are in F-cluster X and cells $C3$ and $C4$ are in F-cluster Y , and $C1$ and $C3$ are adjacent, $C2$ and $C4$ are adjacent. From the ToD construction, $C1$ and $C3$ will be in one S-cluster, S-cluster I , and $C2$ and $C4$ will be in the other S-cluster, S-cluster II . Because from S-aggregator I and II , F-aggregator X and Y choose one that is closer to the sink as the first S-aggregator, they will choose the same S-aggregator. Therefore, their packets can be aggregated at the first S-aggregator, and this completes the proof. \square

Property 3. *The modified dynamic forwarding rules guarantee that packets can be aggregated to one aggregator in the presence of voids if nodes in one F-cluster and in neighboring F-clusters can communicate with each other without routing through other F-clusters.*

Proof. There are three cases as shown in Fig. 8 where some neighboring F-clusters are in a void. For case 2, since the only F-cluster that the event can span is within a void, the packets can be aggregated at the F-aggregator. Therefore, we only consider cases 1 and 3.

To show that packets will be aggregated at one aggregator, it is sufficient to show that for any possible combination of cases 1 and 3 where an event can span (such as (a, b, c), (a, h), or (f, h)), the modified dynamic forwarding rules will eventually select a common aggregating cluster for these F-clusters.

Assume that the void does not exist. The original dynamic forwarding rules guarantee that packets will be aggregated at one aggregating cluster. This aggregating cluster could be the first aggregating cluster or the second aggregating cluster. If the common aggregating cluster is the first aggregating cluster and it is not within the void, the packets can be aggregated at the first aggregating cluster. If the first aggregating cluster is

within the void, Figs. 8d and 8e are the only possibilities and the packets can be forwarded to the sink directly.

If the common aggregating cluster is the second aggregating cluster and it is not within the void, the packets can be aggregated at the second aggregating cluster. If the second aggregating cluster is within the void, the substitute aggregating cluster will be selected. Because the second aggregating cluster selected by F-aggregators are the same, the substitute aggregating cluster will also be the same, and packets will be aggregated at the substitute aggregating cluster.

If the substitute aggregating cluster is also within the void as Fig. 9b.1, all F-clusters that might be spanned by the event are in the void; therefore, the packets can be aggregated at the F-aggregator. If it is the case as Fig. 9b.2, since nodes can communicate with nodes in neighboring F-clusters directly, the packets can be forwarded to the bottom-right F-clusters for further aggregation. This completes the proof. \square

ACKNOWLEDGMENTS

This work was partially supported by the US National Science Foundation's (NSF) CAREER program (Grant CNS 0546630) and NSF's RI program (Grant CNS 0403342). This material is based upon work supported by the US National Science Foundation under Grants CNS-0546630 (CAREER Award), CNS-0721434, CNS-0721817, and CNS-0403342.

REFERENCES

- [1] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "An Application-Specific Protocol Architecture for Wireless Microsensor Networks," *IEEE Trans. Wireless Comm.*, vol. 1, pp. 660-670, Oct. 2002.
- [2] C. Intanagonwiwat, D. Estrin, R. Govindan, and J. Heidemann, "Impact of Network Density on Data Aggregation in Wireless Sensor Networks," *Proc. 22nd Int'l Conf. Distributed Computing Systems (ICDCS '02)*, pp. 457-458, July 2002.
- [3] W. Zhang and G. Cao, "DCTC: Dynamic Convoy Tree-based Collaboration for Target Tracking in Sensor Networks," *IEEE Trans. Wireless Comm.*, vol. 3, pp. 1689-1701, Sept. 2004.
- [4] M. Ding, X. Cheng, and G. Xue, "Aggregation Tree Construction in Sensor Networks," *Proc. 58th IEEE Vehicular Technology Conf. (VTC '03)*, vol. 4, pp. 2168-2172, Oct. 2003.
- [5] H. Luo, J. Luo, and Y. Liu, "Energy Efficient Routing with Adaptive Data Fusion in Sensor Networks," *Proc. Third ACM/SIGMOBILE Workshop Foundations of Mobile Computing*, Aug. 2005.
- [6] A. Goel and D. Estrin, "Simultaneous Optimization for Concave Costs: Single Sink Aggregation or Single Source Buy-at-Bulk," *Proc. 14th Ann. ACM-SIAM Symp. Discrete Algorithms (SODA)*, 2003.
- [7] *Networked Infomechanical Systems*, <http://research.cens.ucla.edu>, 2007.
- [8] *Center for Embedded Networked Sensing at UCLA*, <http://research.cens.ucla.edu>, 2007.
- [9] J. Polastre, "Design and Implementation of Wireless Sensor Networks for Habitat Monitoring," Master's thesis, Univ. of California at Berkeley, Spring 2003.
- [10] A. Mainwaring, R. Szewczyk, J. Anderson, and J. Polastre, *Habitat Monitoring on Great Duck Island*, <http://www.greatduckisland.net>, 2007.
- [11] A. Arora, P. Dutta, and S. Bapat, *Line in the Sand: A Wireless Sensor Network for Target Detection, Classification, and Tracking*, OSU-CISRC-12/03-TR71, 2003.
- [12] *ExScal*, <http://www.cast.cse.ohio-state.edu/exscal/>, 2007.
- [13] *Chemical/Bio Defense and Sensor Networks*, Sentel Corporation, <http://www.sentel.com/html/chemicalbio.html>, 2007.

- [14] K.W. Fan, S. Liu, and P. Sinha, "On the Potential of Structure-Free Data Aggregation in Sensor Networks," *Proc. IEEE INFOCOM '06*, Apr. 2006.
- [15] K.W. Fan, S. Liu, and P. Sinha, "Scalable Data Aggregation for Dynamic Events in Sensor Networks," *Proc. Fourth ACM Conf. Embedded Networked Sensor Systems (SenSys '06)*, Nov. 2006.
- [16] C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann, and F. Silva, "Directed Diffusion for Wireless Sensor Networking," *IEEE Trans. Networking*, vol. 11, pp. 2-16, Feb. 2003.
- [17] S. Madden, M.J. Franklin, J.M. Hellerstein, and W. Hong, "TAG: A Tiny AGgregation Service for Ad-Hoc Sensor Networks," *Proc. Fifth Symp. Operating Systems Design and Implementation (OSDI '02)*, pp. 131-146, Dec. 2002.
- [18] S. Madden, R. Szewczyk, M.J. Franklin, and D. Culler, "Supporting Aggregate Queries over Ad-Hoc Wireless Sensor Networks," *Proc. Fourth IEEE Workshop Mobile Computing Systems and Applications (WMCSA '04)*, pp. 49-58, June 2004.
- [19] S. Lindsey, C. Raghavendra, and K.M. Sivalingam, "Data Gathering Algorithms in Sensor Networks Using Energy Metrics," *IEEE Trans. Parallel and Distributed Systems*, vol. 13, pp. 924-935, Sept. 2002.
- [20] J. Wong, R. Jafari, and M. Potkonjak, "Gateway Placement for Latency and Energy Efficient Data Aggregation," *Proc. 29th Ann. IEEE Int'l Conf. Local Computer Networks (LCN '04)*, pp. 490-497, Nov. 2004.
- [21] B.J. Culpepper, L. Dung, and M. Moh, "Design and Analysis of Hybrid Indirect Transmissions (HIT) for Data Gathering in Wireless Micro Sensor Networks," *ACM SIGMOBILE Mobile Computing and Comm. Rev.*, vol. 8, pp. 61-83, Jan. 2004.
- [22] H.F. Salama, D.S. Reeves, and Y. Viniotis, "Evaluation of Multicast Routing Algorithms for Real-Time Communication on High-Speed Networks," *IEEE J. Selected Area in Comm.*, vol. 15, Apr. 1997.
- [23] R. Cristescu, B. Bererull-Lozano, and M. Vetterli, "On Network Correlated Data Gathering," *Proc. IEEE INFOCOM '04*, vol. 4, pp. 2571-2582, Mar. 2004.
- [24] A. Scaglione and S.D. Servetto, "On the Interdependence of Routing and Data Compression in Multi-Hop Sensor Networks," *Proc. ACM MobiCom '02*, pp. 140-147, Sept. 2002.
- [25] A. Scaglione, "Routing and Data Compression in Sensor Networks: Stochastic Models for Sensor Data That Guarantee Scalability," *Proc. IEEE Int'l Symp. Information Theory (ISIT '03)*, p. 174, June 2003.
- [26] R. Cristescu and M. Vetterli, "Power Efficient Gathering of Correlated Data: Optimization, NP-Completeness and Heuristics," *Summaries of MobiHoc 2003 Posters*, vol. 7, pp. 31-32, July 2003.
- [27] S. Pattern, B. Krishnamachari, and R. Govindan, "The Impact of Spatial Correlation on Routing with Compression in Wireless Sensor Networks," *Proc. Third Int'l Symp. Information Processing in Sensor Networks (IPSN '04)*, pp. 28-35, Apr. 2004.
- [28] Y. Zhu, K. Sundaresan, and R. Sivakumar, "Practical Limits on Achievable Energy Improvements and Useable Delay Tolerance in Correlation Aware Data Gathering in Wireless Sensor Networks," *Proc. IEEE Second Ann. IEEE Comm. Soc. Conf. Sensor and Ad Hoc Comm. and Networks*, Sept. 2005.
- [29] L. Cai and D. Corneil, "Tree Spanners," *SIAM J. Discrete Math.*, vol. 8, 1995.
- [30] N. Alon, R.M. Karp, D. Peleg, and D. West, "A Graph Theoretic Game and Its Application to the K-Server Problem," *SIAM J. Computing*, vol. 24, 1995.
- [31] D. Peleg and D. Tandler, "Low Stretch Spanning Trees for Planar Graphs," Technical Report MCS01-14, Math. and Computer Science, Weizmann Inst. Science, 2001.
- [32] A. Arora et al., "ExScal: Elements of an Extreme Scale Wireless Sensor Network," *Proc. 11th IEEE Int'l Conf. Embedded and Real-Time Computing Systems and Applications (RTCSA '05)*, Aug. 2005.
- [33] Millimeter-Wave Radar Detection of Chemicals, Gases, and Radiation, Argonne Nat'l Laboratory, http://www.anl.gov/techtransfer/pdf/Profile_MMWave_9-16-04.pdf, 2007.
- [34] N. Bulusu, J. Heidemann, and D. Estrin, "GPS-less Low Cost Outdoor Localization for Very Small Devices," *IEEE Personal Comm.*, special issue on Smart Spaces and Environments, vol. 7, Oct. 2000.
- [35] D. Moore, J. Leonard, D. Rus, and S. Teller, "Robust Distributed Network Localization with Noisy Range Measurements," *Proc. Second ACM Conf. Embedded Networked Sensor Systems (SenSys '04)*, pp. 50-61, Nov. 2004.
- [36] K.W. Fan, S. Liu, and P. Sinha, "Dynamic Forwarding over Tree-on-DAG for Scalable Data Aggregation in Sensor Networks," Technical Report, OSU-CISRC-02/08-TR8, Dept. CSE, OSU, Feb. 2008.
- [37] E. Ertin, A. Arora, R. Ramnath, M. Nesterenko, V. Naik, S. Bapat, V. Kulathumani, M. Sridharan, H. Zhang, and H. Cao, "Kansei: A Testbed for Sensing at Scale," *Proc. Fourth Symp. Information Processing in Sensor Networks (IPSN/SPOTS Track)*, 2006.
- [38] A. Arora, E. Ertin, R. Ramnath, W. Leal, and M. Nesterenko, "Kansei: A High-Fidelity Sensing Testbed," *IEEE Internet Computing*, special issue on Large-Scale Sensor Networks, Mar. 2006.
- [39] Crossbow, Crossbow, <http://www.xbow.com>, 2007.
- [40] W. Zhang and G. Cao, "Optimizing Tree Reconfiguration for Mobile Target Tracking in Sensor Networks," *Proc. IEEE INFOCOM '04*, vol. 4, pp. 2434-2445, Mar. 2004.
- [41] K. Gabriel and R. Sokal, "A New Statistical Approach to Geographic Variation Analysis," *Systematic Zoology*, pp. 259-278, 1969.



Kai-Wei Fan received the BS and MS degrees in computer science and information engineering from National Chiao Tung University, Hsinchu, in 1997 and 1999, respectively. He is currently working toward the PhD degree in the Department of Computer Science and Engineering, The Ohio State University, Columbus. He was a senior engineer and a project manager in network security division in a start-up company from 1999 to 2004. His research interests include wireless sensor networks and mesh networks, where he focuses on design and implementation of energy-efficient protocols.



Sha Liu received the BS degree in statistics and the MS degree in computer science from the University of Science and Technology of China, in 2001 and 2004, respectively. He is currently a PhD student in computer science in the Department of Computer Science and Engineering, The Ohio State University, Columbus. His research interests include low latency routing, wake-up scheduling, and data aggregation in wireless sensor networks.



Prasun Sinha received the BTech degree from IIT Delhi, in 1995, the MS degree from Michigan State University, in 1997, and the PhD degree from the University of Illinois, Urbana-Champaign, in 2001. He worked at Bell Labs, Lucent Technologies as a member of technical staff from 2001 to 2003. Since 2003, he has been an assistant professor in the Department of Computer Science and Engineering, The Ohio State University, Columbus. His research interests include design of network protocols for sensor networks and mesh networks. He served on the program committees of various conferences including IEEE INFOCOM (2004-2007) and ACM MobiCom (2004-2005). He has won several awards including Ray Ozzie Fellowship (UIUC, 2000), Mavis Memorial Scholarship (UIUC, 1999), and Distinguished Academic Achievement Award (MSU, 1997). He received the prestigious NSF CAREER award in 2006.

► For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/publications/dlib.