Journal of
Applied Sciences

science
alert

Achieving Scalability in Wireless Sensor Network Using Hexagonal Multi-layer Grid Data Dissemination Approach

Mohamed Essam Khedr and Amr Khamis Mahmoud

Department of Electronics and Communications, College of Engineering, AAST, Alexandria, Egypt

Abstract: Many challenges are being faced by wireless sensor networks in general and that of large scale in specific. Location accuracy and scalability in information dissemination are important challenges facing large scale wireless networks. Hexagonal Two Tier Data Dissemination (HTTDD) is an approach that disseminates the data between multiple sources and multiple mobile sinks and targets these two challenges. Instead of propagating the query message from sinks to all sensor nodes in the field to setup the data forwarding information, HTTDD uses a hexagonal grid structure so that only sensor nodes at the grid vertices need to have the forwarding data. Once a sensor node detects an event of interest it becomes a source which proactively builds a hexagonal grid and sets up the forwarding route consisting of the sensor nodes closest to the grid vertices hence called dissemination nodes. The path taken by a query issued to a source from a sink consists of two tiers. The first tier is within the coverage area of the sink and the second tier is the pass made of the dissemination nodes constituting the grid structure. The performance of hexagonal TTDD evaluated through detailed analysis and simulation is presented in this study and compared to the squared TTDD approach to prove the improvement added by HTTDD.

Key words: Wireless sensor networks, hexagonal two tier grid structure, data dissemination, network scalability, location-based routing

INTRODUCTION

Wireless sensor networks (WSN) consist of large number of sensors distributed in space forming self-operated self-managed and self-organized networks. The tasks of such networks are vast and diverge ranging from monitoring the physical behavior of the environment to cooperating in delivering such behavior through the network (Changle *et al.*, 2011). The network ability to apply cooperation and coordination among the sensors shows the fundamental differences between sensor networks and just having sensors deployed in the surroundings. In addition, due to the broadcasting nature of wireless sensor networks, challenges such as energy conservation, efficient bandwidth utilization and network scalability are demanding issues for sustainable operation of such networks (Hua and Yum, 2008).

Concerns about scalability and efficient data dissemination in large scale wireless sensor networks are what drive research in this field especially if there are many mobile sources and sinks. In this study, a source is any sensor node that detects and monitors the occurrence of events of interest and disseminates information about such events. On the other hand, a sink is any sensor node

that gathers this information from the sensor network. Different approaches were developed for achieving good data dissemination in sensor networks. Directed Diffusion (Zhao *et al.*, 2009; Choe and Kim, 2008), Declarative Routing Protocol (Liu *et al.*, 2011) and GRAB (Ye *et al.*, 2005), are some of the well known techniques for data dissemination. Almost all techniques are based on periodic transmission of mobile sinks' location information in the sensor network field in order that all other sensors are kept updated with the route needed for sending further data. This leads to excessive redundant transmission of location information and as a consequence increases number of collisions and consumes more power from battery operated sensors. Wang and Chiang (2011) proposed a grid-based data dissemination protocol with power preservation capability where dissemination nodes with the most residual power are selected for forwarding queries. Therefore, current approaches for data dissemination lack the ability to be scalable and efficient in data dissemination. To solve the scalability issue, Two Tier Data Dissemination "TTDD" was suggested that disseminates the data between multiple sources and multiple mobile sinks over a squared grid structure so that sensor nodes at the grid points only

Corresponding Author: Mohamed Essam Khedr, Department of Electronics and Communications, College of Engineering, AAST, Alexandria, Egypt

need to have the forwarding data (Ye *et al.*, 2002). A query from a sink will traverse two tiers. The first tier is within the coverage area of the sink and the query reaches every node in this area through flooding. The second tier takes over when the query reaches a disseminating node that will forward the query to the other disseminating nodes at the grid vertices. The forwarding process will continue until the query reaches the source or a dissemination node that already receives data from the source. In this process each dissemination node stores the path to the sink to be able to traverse the same two tiers in the reverse order to deliver the data report (Wang and Liu, 2011). Each source constructs its own grid and each sink confines its query to its coverage area only avoiding any excessive energy consumption and securing the network from traffic overload due to global flooding by multiple sinks. When a sink moves away from its cell it sends another query to discover the nearest dissemination node. This dissemination node sends the query to its upstream nodes until it reach to the source which will start forwarding the data back to the sink through the downstream dissemination nodes. Even if the sink is moving continuously the higher tier of the grid will change incrementally and the sink will receive the data continuously. Other sensor nodes that are not employed as dissemination nodes do not maintain any states about the grid which makes TTDD scalable to large number of sources and sinks.

TTDD uses square grid structure to disseminate the data between the sources and the sinks which is not the optimum grid structure preferred in data network such as the mobile networks due to its high overhead in construction.

Our contribution in this paper is a new approach called HTTDD; Hexagonal Two Tier Data Dissemination which is based on the TTDD approach with a modified hexagonal grid scheme. The proposed scheme can use both straight routes and/or diagonal routes following a hexagonal virtual grid. Compared to TTDD, the communication overhead is improved because fewer sensor nodes are employed for data communication and also shows superior performance in delay and state complexity over the squared TTDD (STTDD).

DATA DISSEMINATION PROTOCOLS IN WIRELESS SENSOR NETWORKS

Communication processes in large scale wireless sensor networks are planned around the sensing data rather than the sensor nodes in what is called data centric approach. The distinctive specifications of wireless sensor networks make the problems of data dissemination and scalability challenging and different from conventional wireless ad-hoc networks. In specific, there are three main design challenges.

Scalability: Sensor networks are characterized by thousands or even hundreds of thousands of nodes covering vast areas in the environment while traditional ad-hoc networks have few hundred nodes in general.

Constrained resources: Sensors are simple and inexpensive nodes manufactured with limited battery power, relatively low CPU power and constrained memory size.

Reliability: Environmental obstacles and interferences are common in wireless sensor networks. In addition, sensor nodes are susceptible to failures or damaged unexpectedly.

Flat-architecture data dissemination protocols suffer from congestion in traffic as the network density increases especially as traffic gets closer to the sink. As a consequence, nodes located close to sinks forward more information than nodes farther apart from sinks in the network. As a result, these nodes consume their power faster causing disconnections between sinks and sources in WSN. The disadvantages of the flat-architecture data dissemination can be addressed by hierarchical architecture where sensor nodes are organized in clusters and local interactions between cluster members are managed by a cluster head. The cluster heads can also form another layer of clusters among themselves before reaching the sink in what is called multi-tier approach. Examples of such architecture are the low-energy adaptive clustering hierarchy (LEACH) (Wang *et al.*, 2010), power-efficient gathering in sensor information systems (PEGASIS) (Lindsey and Raghavendra, 2002), real time two hop routing approach (Li *et al.*, 2009) and DAWN (Tang *et al.*, 2010). Location information is essential in many applications of WSNs. To provide location information to sensor nodes, GPS devices can be embedded in sensor nodes. However, GPS is only useful in stationary nodes with a sufficient amount of energy and processing power. Since location information is necessary for most applications in WSN, it is frequent to use this information for data dissemination as well. Location-based data dissemination protocols exploit the location information of sensor nodes to provide efficient and scalable data forwarding.

HEXAGONAL TWO TIER DATA DISSEMINATION STRUCTURE AND ANALYSIS

HTTDD (Mahmoud *et al.*, 2012) was built on the same concept of TTDD where a grid structure is used in order that only sensors located at the grid indices are the ones required to capture and forwarding information. The main and radical differences between TTDD and HTTDD are that the later uses hexagonal grid structure making it more realistic to the circular coverage of the

omnidirectional antenna and second the flexibility in choosing diverse routes that are not supported in TTDD. Upon discovery of an incident the source proactively builds a hexagonal grid structure throughout the sensor field and disseminates captured information over that grid to sensors closest to grid vertices that are called afterwards dissemination nodes. A query from sink goes across two tiers in a virtual hexagonal grid to reach source. The lower tier is within the hexagonal cell of the sink's current location while the higher tier is the dissemination nodes at the grid vertices. The sink floods the query in the first tier to ensure that the nearest dissemination node on the hexagonal grid receives the query. After this phase, the disseminating nodes take over and forward the query in the second tier upstream toward the source. HTTDD uses any off-the-shelf geographical forwarding algorithm to construct and maintain the grid structure with low overhead (Pottie and Kaiser, 2000). In addition, each data source builds its grid structure while queries from multiple mobile sinks are confined within their respective local cells. This ensures the avoidance of excessive energy usage and overloading the network from flooding by multiple sinks. Local flooding of query is performed again when a sink moves out of its current cell which will cause small incremental changes in higher-tier data forwarding leading to no interruption in data flow.

HTTDD grid construction: We assume the sensor field to be a two dimensional plane with uniform random distribution of sensor nodes. The source constructs the grid by dividing the field into hexagonal cells with side of length α and the source is at one corner of the grid as shown in Fig. 1. The source propagates the data announcement message to reach all the corner points where other dissemination nodes reside. For a source at location Ls = (x, y), the dissemination nodes are located at Lp = (xi, yj) such that:

$$xi = \frac{i * \alpha}{2}, \ yj = \frac{j * \sqrt{3}}{2} \qquad (1)$$

for i, j = 0,2,4,6,... or for i, j = 1,3,5

A source S calculates the locations of its three neighboring dissemination nodes (Lp,0):

$$\left(\frac{1}{2}Lp.\text{-}Lp \ Sin60 \right)$$

and

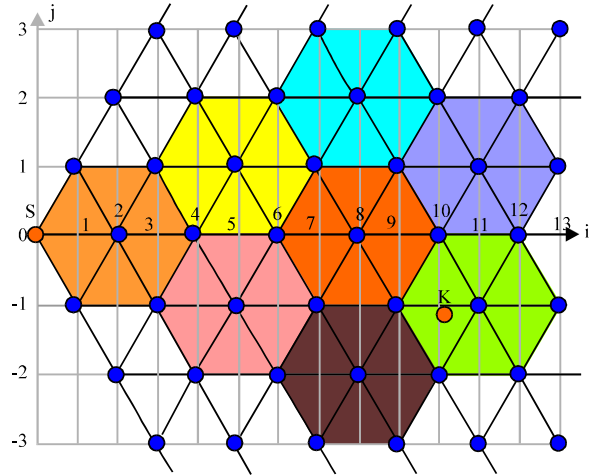$$\left(\frac{1}{2}Lp.\text{-}Lp \ Sin60 \right)$$

Fig. 1: Hexagonal grid structure where I steps in horizontal direction, J steps in vertical direction, S is source and K is destination

given its location (x, y) and cell size α using GPS or any positioning system as shown in Fig. 1.

A dissemination node stores information that is needed in grid construction and maintenance such as the data announcement messages, the disseminating point it is serving and the upstream dissemination nodes' location information. The hexagonal cell size "α" is a crucial parameter as it localizes the impact of sink mobility within the cell which in return stabilizes the higher-tier grid forwarding mechanism. In addition, the choice of the cell size α affects the communication overhead and the state complexity.

Grid construction in HTTDD is on a per-source basis. This means that each source builds its own grid and assigns different sets of dissemination nodes to this grid. This decision will enhance scalability and will provide load balancing and better robustness since at any given time any dissemination node only has states about small number of sources and effectively distributes data load avoiding bottle-necks situations.

Hexagonal two-tier query forwarding: Query and data forwarding in HTTDD is based on the virtual hexagonal grid infrastructure to ensure scalability and efficiency. When a sink requests data, it floods a query in the hexagonal cell it belongs to in order to discover nearby dissemination nodes. Once a query is heard by a local dissemination node, it is up-streamed on the hexagonal grid towards the source, until finally the query reaches the source. As the query is forwarded upstream, each disseminating node along the upstream stores the location information of the downstream

disseminating node from which it receives the query. This state is used to direct data back to the sink as shown in Fig. 2.

An upstream update message traverses the grid to keep the deployed hexagonal grid alive. Dissemination nodes send these upstream update messages periodically in order to keep the grid alive and receive data continuously. This update message is ceased when it is no longer needed, such as in the case when the sink stops sending queries or moves out of the local region. The periodic time value is chosen to be bigger than the inter-arrival time between data messages. This setting balances the overhead of generating periodic upstream update messages and that of sending data to places where it is no longer needed.

In HTTDD, any dissemination node on the query forwarding route maintains at most three states about neighboring dissemination nodes. Sensor nodes that do not participate in query or data forwarding do not keep any state about sinks or sources.

Hexagonal two-tier data forwarding: Upon receiving queries from neighbor disseminating nodes, the source replies back with the data to this disseminating node that will forward this data along the downstream route of the received queries until the data reaches the requested sink. A copy of data is sent along each downstream if a dissemination node is common in multiple down-streams for the same source. For example in Fig. 2 the dissemination node A will send data to both K1 and K2. With the two-tier forwarding as described above, queries and data may take globally suboptimal paths, thus introducing additional cost compared to forwarding along shortest paths. For example in Fig. 2, sink K1 and K2 may find straight line paths to the source if they each flooded their queries across the whole sensor field. We believe that the sub-optimality is well worth the gain in scalability.

Hexagonal grid maintenance: To avoid keeping grid states at dissemination nodes indefinitely, a grid lifetime is included in the data announcement message. Proper grid lifetime values depend on the data availability period and the mission of the sensor network. If the grid lifetime elapsed and the dissemination nodes does not receive any data announcement message to update the grid lifetime the grid will be no longer exists, important issue that have to be taken into consideration is the power of each sensor nodes when the grid is alive, we need to measure the energy level at each sensor node in order to replace it with another one of those who have a duplicated upstream location information to keep the link alive. Each dissemination node on the grid chooses a number of its neighborhood sensor nodes to replicate in them the location of its upstream dissemination node.

When this node fails as shown in Fig. 3 and 4, the upstream messages from its downstream dissemination node that needs data will stop at one of these chosen nodes. This node then forwards the update message to the upstream dissemination node according to the stored information route. When data come from upstream later, a new dissemination node will emerge following the same procedure taken by the source when building the grid. A downstream dissemination node that needs data will continue to send upstream update messages to the forwarding state alive. The failure of the immediate
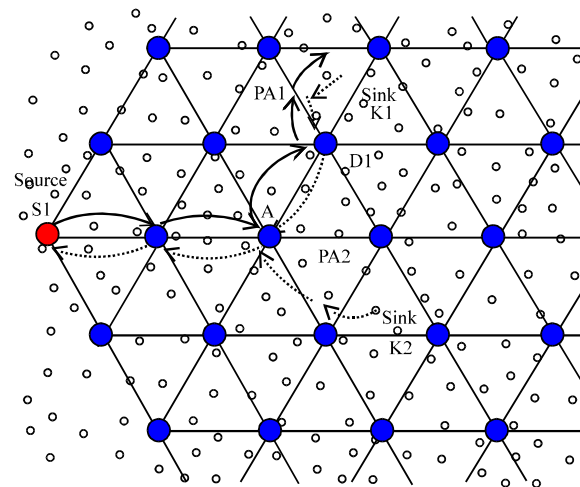

Fig. 2: Two-tier query and data forwarding between Source S and Sink K1, K2. D is disseminating node. A is common disseminating node. PA is local disseminating node
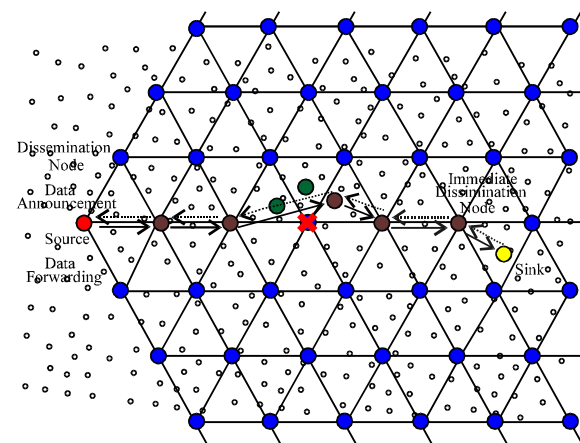

Fig. 3: Upstream Grid Maintenance where red node is the source, brown nodes are the disseminating nodes in the upstream, green nodes are the backup nodes and yellow node is the sink node
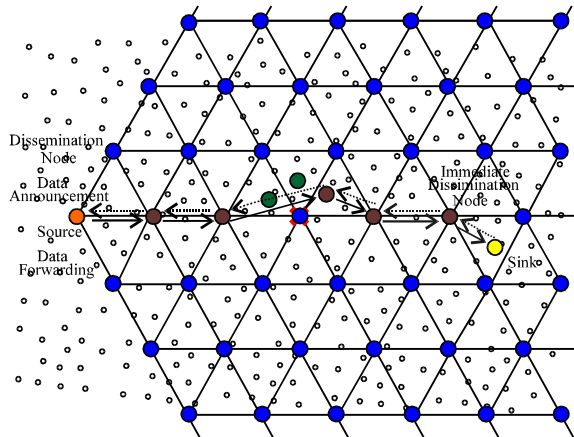
Fig. 4: Downstream Grid Maintenance

dissemination node is detected by a timeout at a sink. When a sink stops receiving data for a certain time, it re-floods a query to locate a new dissemination node. The failures of primary agents or immediate agents are detected by similar timeouts and new ones will be picked.

HTTDD mathematical analysis: Here, we analyze the efficiency and scalability of HTTDD. We measure two metrics: the communication overhead and states complexity. Communication overhead refers to the total messages exchanged between the nodes to accomplish certain task while states complexity refers to number of operations performed by sensor nodes involved in data dissemination mechanism. We compare the performance of HTTDD approach with Squared Two Tier Data Dissemination STTDD approach, our analysis will focus on the worst-case communication overhead of each protocol. We aim at making the analysis simple and easy to follow while capturing the fundamental differences between HTTDD and STTDD.

We have two different scenarios to compare HTTDD with STTDD which are:

- Hexagonal grid vs. squared grid with equal cell side length
- Hexagonal grid vs. squared grid with equal cell area

The analysis is based in a sensor field of area A with N sensor nodes uniformly distributed such that there are approximately $\sqrt{N}$ sensor nodes on each side. There are k sinks in the sensor field moving with an average speed v, while receiving d data packets from a source in a time period of T. Each data packet has a unit size and both the query and data announcement messages have a comparable size L. The communication overhead to flood an area is proportional to the number of sensor nodes in it and to send a message along a route using geographical

forwarding is proportional to the number of sensor nodes in the path. The average number of neighbors within a sensor node's wireless communication range is D.

In HTTDD-equal area scenario, the source divides the sensor field into cells; each has an area α^2. There are:

$$n = N\frac{\alpha^2}{A}$$

sensor nodes in each cell and $\sqrt{N}$ sensor nodes on each side of the squared sensor field. Each sink traverses m cells and m is upper bounded by:

$$1 + \frac{\upsilon T}{\alpha}$$

For stationary sinks, m = 1.

On the other side HTTDD-equal side scenario, the source divides the sensor field into cells; each has an area:

$$\frac{3\sqrt{3}\alpha^2}{2}$$

There are:

$$n = \frac{N}{A} * \frac{3\sqrt{3}\alpha^2}{2}$$

sensor nodes in each cell and $\sqrt{N}$ sensor nodes on each side of the squared sensor field. Each sink traverses m cells and m is upper bounded by:

$$1 + \frac{\upsilon T}{\alpha}$$

and m = 1 for stationary sinks. We analyze the worst-case communication overhead of the HTTDD approach and compare it with STTDD approach. We assume in both approaches that a sink updates its location m times and receives d/m data packets between two consecutive location updates.

Communication Overhead-first scenario: In HTTDD, The overhead for the query to reach the source, without considering query aggregation, is:

$$n_{hex}L + \sqrt{N(1+r^2)}L = 2.6n_{sq}L + \sqrt{N(1+r^2)}L \qquad (2)$$

where, n_{hex} is the number of sensor node inside a hexagonal cell and n_{sq} is the number of sensor nodes inside a squared cell.

$$n_{hex} = \frac{N}{A} * \frac{3\sqrt{3}\alpha^2}{2}, \; n_{sq} = N\frac{\alpha^2}{A}, \; n_{hex} = \frac{3\sqrt{3}}{2}n_{sq} \qquad (3)$$

where, $\sqrt{N(1+r^2)}$ is the average number of sensor nodes along the straight path from the source to the sink $(0 < r \leq 1)$, the other two sides of the triangle represents the two sides of the squared sensor field as shown in Fig. 5a. Similarly, the overhead to deliver d/m data packets from a source to a sink is:

$$\sqrt{N(1+r^2)}\frac{d}{m} \qquad (4)$$

For k mobile sinks, the overhead to receive d packets in m cells is:

$$km\left[n_{hex}L + \sqrt{N(1+r^2)}L + \sqrt{N(1+r^2)}\frac{d}{m}\right] =$$
$$kmn_{hex}L + k\sqrt{(1+r^2)}(mL+d)\sqrt{N} = \qquad (5)$$
$$2.6kmn_{sq}L + k\sqrt{(1+r^2)}(mL+d)\sqrt{N}$$

For updating the mission of the sensor network, the overhead will be NL and the overhead to construct the grid is:

$$6\frac{N}{\sqrt{n_{hex}}}L = 3.7\frac{N}{\sqrt{n_{sq}}}L \qquad (6)$$

From Eq. 4, 5 and 6, the total communication overhead of HTTDD becomes:

$$NL + 3.7\frac{N}{\sqrt{n_{sq}}} + 2.6kmn_{sq}L + k\sqrt{(1+r^2)}(mL+d)\sqrt{N} \qquad (7)$$

In STTDD approach, the overhead for the query to reach the source, without considering query aggregation, is:

$$n_{sq}L + \sqrt{2}\left(c\sqrt{N}\right)L \qquad (8)$$
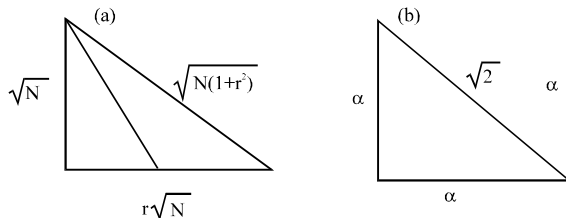
Where:

$c\sqrt{N}$ = Average number of sensor nodes along the straight-line path from the source to the sink $(0 < c \leq \sqrt{2})$ as shown in Fig. 5b

$\sqrt{2}$ = Because a query in STTDD traverses a grid instead of straight-line path, the worst-case path length is increased by a factor of $\sqrt{2}$

The overhead to deliver d/m data packets from a source to a sink is:

$$\sqrt{2}\frac{\left(c\sqrt{N}\right)d}{m} \qquad (9)$$

For k mobile sinks, the overhead to receive d packets in m cells is:

$$km\left[n_{sq}L + \sqrt{2}\left(c\sqrt{N}\right)L + n_{sq}L + \sqrt{2}\left(c\sqrt{N}\right)\frac{d}{m}\right] = \qquad (10)$$
$$kmn_{sq}L + kc(mL+d)\sqrt{2N}$$

The overhead for updating the mission of the sensor network is NL and constructing the grid overhead

$$4\frac{N}{\sqrt{n_{sq}}}L \qquad (11)$$

From Eq. 9-11, the total communication overhead of STTDD approach becomes:

$$NL + 4\frac{N}{\sqrt{n_{sq}}}L + kmn_{sq}L + kc(mL+d)\sqrt{2N} \qquad (12)$$

Figure 6 shows the theoretical HTTDD communication overhead vs. cell size in terms of (n) and with different sink speed (m). The figure indicates that the communication overhead decreases with the increase of the cell size up to a certain size where it will increase with cell size. The reason behind this is that the communication overhead consists of two terms the grid construction overhead term and the query forwarding overhead term. Thus as the cell size increases the grid construction overhead decreases while the query forwarding overhead increases. As a consequence, for small cell size the main term contributing in the communication overhead is the grid construction overhead while for large cell size the main term will be the query forwarding overhead. The figure also indicates that the speed of the mobile sink does not affect the grid construction overhead term while significantly affect the query forwarding overhead term.

Figure 7 shows the theoretical communication overhead of HTTDD with the same cell side length as that

Fig. 5(a-b): Sensor field sides, (a) Hexagonal grid and (b) Squared grid
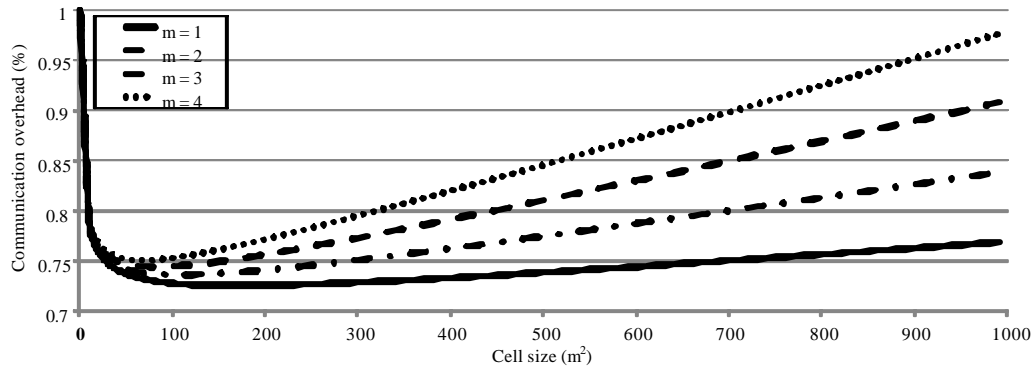
Fig. 6: HTTDD overhead vs. Cell size for different sink moving speeds-First Scenario
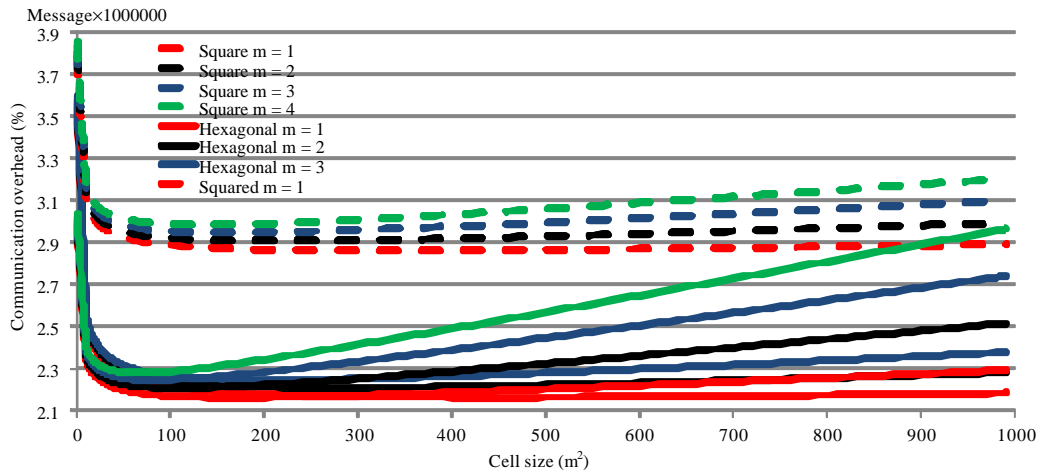

Fig. 7: HTTDD vs. STTDD communication overhead for different sink moving speed-first scenario

of STTDD. As indicated from figure, HTTDD has in general a lower communication overhead compared to STTDD. However as the mobility increases the communication overhead in HTTDD grows faster than in STTDD but still keeping the communication overhead in HTTDD lower than STTDD except in extremely large cell sizes. This is mainly due to the local query flooding.

Communication overhead-second scenario: The overhead for the query to reach the source, without considering query aggregation, is:

$$n_{sq}L + \sqrt{N(1+r^2)}L \qquad (13)$$

where, $\sqrt{N(1+r^2)}L$ is the average number of sensor nodes along the straight-line path from the source to the sink ($0<r\leq1$).

The overhead to deliver d/m data packets from a source to a sink is $\sqrt{N(1+r^2)}\dfrac{d}{m}$. For k mobile sinks, the overhead to receive d packets in m cells is:

$$km\left[n_{hex}L + \sqrt{N(1+r^2)}L + \sqrt{N(1+r^2)}\dfrac{d}{m}\right] =$$
$$kmn_{hex}L + k\sqrt{N(1+r^2)}(mL+d)\sqrt{N} = \qquad (14)$$
$$kmn_{sq}L + k\sqrt{N(1+r^2)}(mL+d)\sqrt{N}$$

From Eq. 13 and 14, the total communication overhead of HTTDD becomes:

$$NL + 6\dfrac{N}{\sqrt{n_{sq}}}L + kmn_{sq}L + k\sqrt{N(1+r^2)}(mL+d)\sqrt{N} \qquad (15)$$

Figure 8 shows the theoretical communication overhead of HTTDD with the same cell area as the
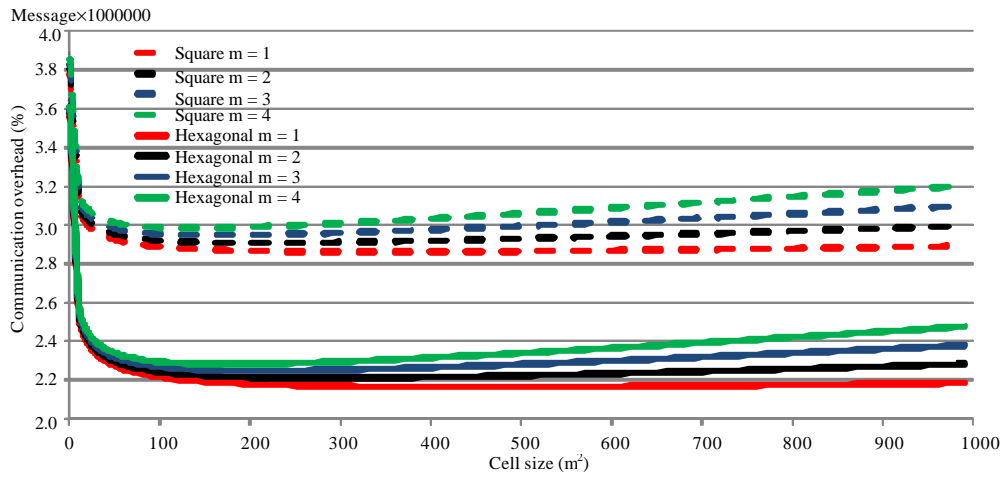
Message×1000000

Fig. 8: HTTDD vs. STTDD communication overhead for different sink moving speed-second scenario
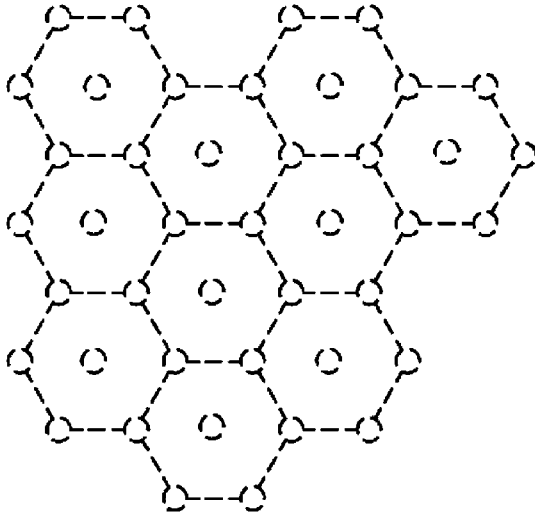

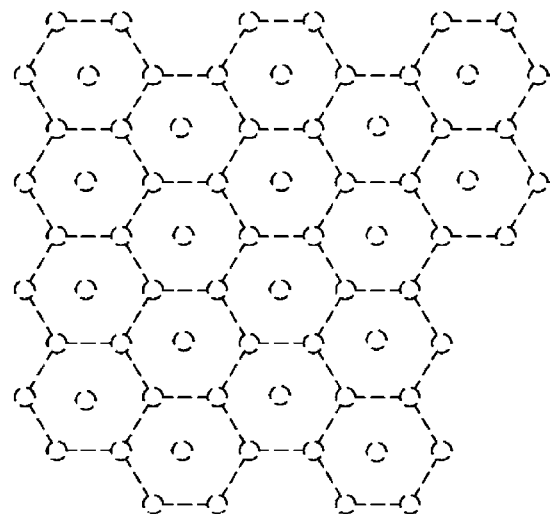Fig. 9: HTTDD grid with 10 cells and 43 dissemination node

Fig. 10: HTTDD grid with 18 cells and 72 dissemination node

STTDD. Compared to Fig. 7, it is clear that the issue of large cells does not exist in this situation and that HTTDD is always better than STTDD for all values of cell size. This is contributed to the fact that the local query flooding overhead is much lower in HTTDD due to the diagonal paths that can be routed between the source and the sink and this will be the dominant term in the overhead calculations.

State complexity: State complexity represents the number of states that have to be maintained in each sensor node during the life of the grid. In HTTDD, only dissemination nodes and their neighbors which duplicate upstream information maintain states for data dissemination. All other sensor nodes do not need to

maintain any state. The state complexities at different sensor nodes are analyzed as follows:

First we need to get the total number of dissemination nodes in a grid, so we assume that the hexagonal grid will be distributed on a squared area field and the grid must cover the entire field.

From Fig. 9 and 10 we can deduce empirically a relationship between the number of cells and the number of dissemination nodes which will be as follow:

Number of dissemination nodes in HTTDD grid is:

$$3.625\frac{N}{n_{hex}} + 6.75$$

Each dissemination node maintains the location of its upstream dissemination node for query forwarding. For
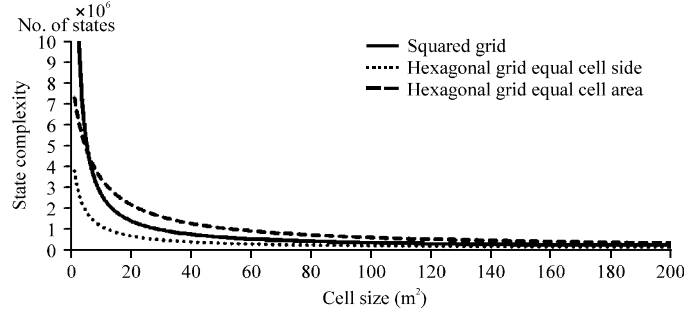
Fig. 11: State complexity for HTTDD vs. STTDD

the case of data forwarding, each maintains locations of at most all the other three neighboring dissemination nodes for data forwarding. We consider data forwarding from s sources to k mobile sinks. In Sink Oriented Data Dissemination, SODD (Jeon *et al.*, 2009) the total number of sensor nodes on data forwarding paths from a source to all sinks is P. Therefore, the number of sensor nodes in STTDD forwarding paths is at most $\sqrt{2}$ P and in general it will be $P\sqrt{(1+r^2)}$. The total number of states maintained for trajectory forwarding in sinks' dissemination nodes are $k(s + 2)$. The total state complexity is:

For STTDD:

$$s\left[b\left(\sqrt{\frac{N}{nsq}}+1\right)^2+3\frac{\sqrt{2}P\,sq}{\sqrt{nsq}}\right]+k\left(s+2\right) \qquad (16)$$

For HTTDD with equal cell side:

$$s\left[b\left(3.625\frac{N}{nhex}+6.75\right)+3\frac{\sqrt{2}P\,sq\sqrt{(1+r^2)}}{\sqrt{nhex}}\right]+k\left(s+2\right) \qquad (17)$$

For HTTDD with equal cell area:

$$s\left[b\left(3.625\frac{N}{nhex}+6.75\right)+3\frac{\sqrt{2}P\,sq\sqrt{(1+r^2)}}{\sqrt{nhex}}\right]+k\left(s+2\right) \qquad (18)$$

Figure 11 shows the state complexity comparison between HTTDD and STTDD. HTTDD with equal cell side has the lowest state complexity because it has the least number of cells covering the space. This is due to the fact that it has the largest cell size among the other approaches. It is also noticed that the STTDD has lower state complexity than HTTDD with equal cell size as it maintains less dissemination nodes.

HEXAGONAL TWO TIER DATA DISSEMINATION SIMULATION RESULTS

Here, we evaluate the performance of HTTDD through simulations. The results confirm the efficiency and scalability of HTTDD to deliver data from multiple sources to multiple mobile sinks. We show that HTTDD has superior performance compared to STTDD. We implemented HTTDD using ns-2. We used the basic greedy geographical forwarding with local flooding to bypass dead ends. To further improve the robustness of HTTDD against dissemination node failures and geographical forwarding failures, we added the following optimization technique: For each source from which a sink receives data, the sink keeps a timer that is reset every time a data packet from that source is received. If the sink does not receive any data from that source for some time, the corresponding timer expires and the sink locally floods a query to locate new immediate dissemination nodes for that source. Simulations show that this technique makes HTTDD more robust to unexpected dissemination node failures.

We used 802.11 DCF as the underlying MAC and we used two metrics to evaluate the performance of HTTDD. The first metric is the success rate that represents the ratio of the number of successfully received packets at a sink to the total number of packets generated by a source, averaged over all source-sink pairs. This metric shows how effective the data delivery is. The second metric is the delay defined as the average time between the moment a source transmits a packet and the moment a sink receives the packet, also averaged over all source-sink pairs. This metric indicates the freshness of data packets and the applicability of HTTDD to real time scenarios.

We simulated the whole network using three different grid scenarios:

* Squared grid
* Hexagonal grid with equal cell area
* Hexagonal grid with equal cell side length

Table 1 shows the simulation parameters used in all scenarios, some of these parameters may be changed in each scenario and this will be mentioned in details.

Table 2 shows the different results of the simulation, describing the expected curves from each simulation run.

Table 1: Simulation parameters

Parameter	Description
Sinks	4
Sources	4
Sensor nodes	200-randomly distributed
Filed area	2000×2000 m²
Simulation run time	200 sec-enough to guarantee the stability of the network.
Network topologies	15 random network topologies-averaged for each simulation
Data packet generation rate	A source generates one data packet each second
Sinks' mobility standard	Random way point model
Query packet length	36 bytes
Each data packet length	64 bytes
Cell size α	600 m
Sink's local query flooding range	1.3 α-to handle irregular dissemination node distribution.
Number of sinks and sources range	Vary from 1, 2, 4, 6 to 8
Sink maximum speed	10 m sec⁻¹ with 5 sec pause time
Sink maximum speed range	Vary from 0, 5, 10, 15 to 20 m sec⁻¹
Node failure	15% randomly chosen nodes fail at t = 20 sec

Table 2: Simulation results' output curves flow chart

|	Curve name	Details
Sink number	Success Ratio vs. Sink Number	1, 2 and 4 sources, Stationary sinks, Equal cell area scenario
	Delay vs. Sink Number	1 source, Stationary sinks, Comparing the three scenarios
Sink mobility	Success Ratio vs. Sink Mobility	1 source, 1 sink, Equal cell area scenario
	Delay vs. Sink Mobility	
Node failure	Success Ratio vs. Node Failure Ratio	1 source, 1 sink, Equal cell area scenario
	Delay vs. Node Failure Ratio	
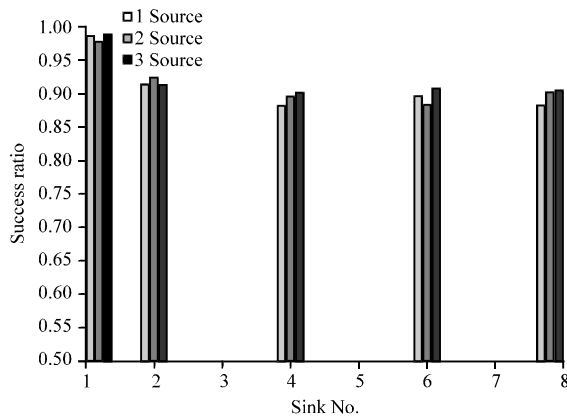
Fig. 12: Success ratio vs. sink number with different number of sources for HTTDD-equal cell area
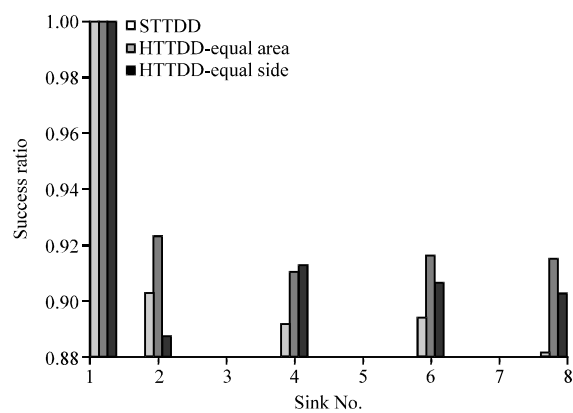

Fig. 13: Success ratio vs. sink number with different grid structures

Impact of the numbers of sinks and sources: Figure 12 shows the success ratio for HTTDD. For a specific number of sinks, the success ratio tends to decrease as the number of source increases. In the 8-sink case, the success ratio decreases from about 0.92 to about 0.89 as the number of sources increases to 4. This is because more sources generate more data packets which lead to more contention-induced losses which is a common problem in wireless networks as explained by Fu *et al.* (2003). Overall, the success ratio shows that HTTDD delivers most of the data packets successfully from multiple sources to multiple mobile sinks and the delivery quality does not degrade much as the number of sources or sinks increases.

Figure 13 shows a comparison among the three different scenarios of the grid structure in terms of the success ratio. HTTDD with equal cell area is always giving the highest success ratio due to its shortest routes between sources and sinks. The gain in using HTTDD grid with equal cell area is 1.3% less than STTDD and 2.2% less than HTTDD gird with equal cell side.

Figure 14 plots the delay measurements which ranges from 0.02 to 0.038 second. The delay tends to increase when there are more sinks or sources as more sources generate more data packets and more sinks need more local query flooding. Both increase the traffic volume and lead to longer delivery time. However, the delay is quite small even with 4 sources and 8 sinks.
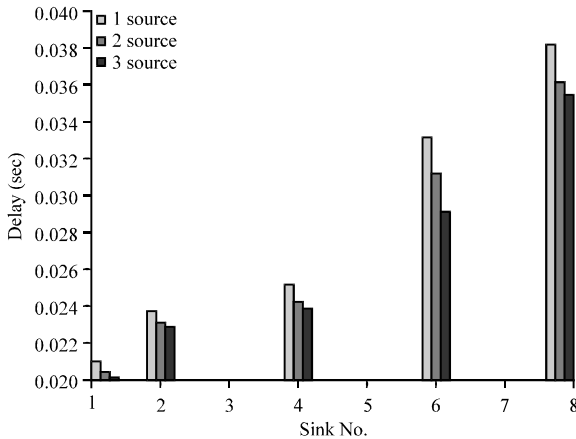
Fig. 14: Delay vs. sink number with different number of sources for HTTDD-equal cell area
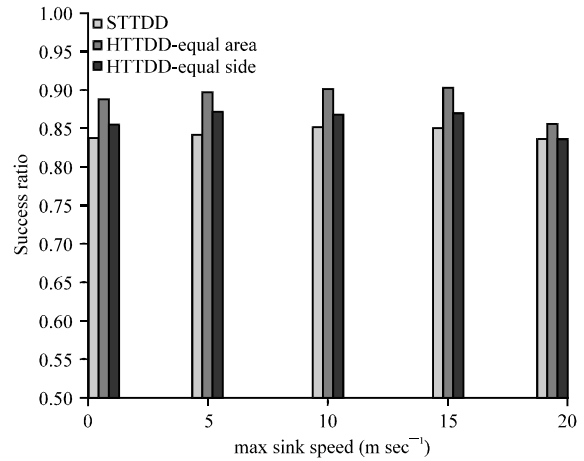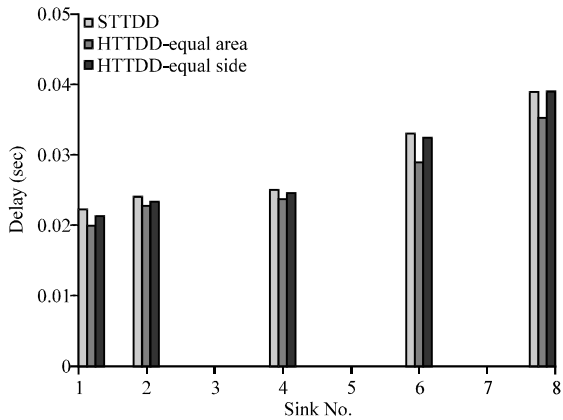

Fig. 15: Delay vs. sink number with different grid structures

Figure 15 shows that HTTDD with equal cell area have the lowest delay compared to STTDD and HTTDD with equal cell side. This is due to shorter routes between the source and the sink compared to STTDD or HTTDD with equal cell side.

Impact of Sinks' Mobility: We next evaluate the impact of sinks' speeds on performance of HTTDD. In the default simulation setting, we vary the maximum speed of sinks from 0 to 20 m sec^{-1} with step size 5 m sec^{-1}.

Figure 16 shows the success ratio as the sinks' speed changes. The success ratio for HTTDD with equal cell area fluctuates around 0.87 as sinks move faster. This shows that sinks react quickly to its location changes and receive data packets from new agents and/or new dissemination nodes even at speeds as high as 20 m sec^{-1}. HTTDD with equal cell side shows lower

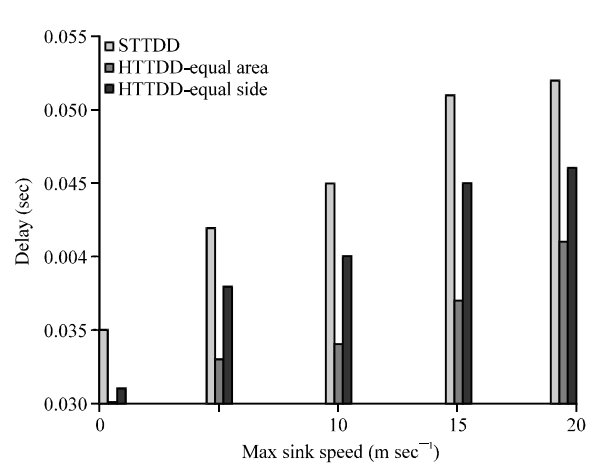Fig. 16: Success ratio vs. max sink speed with different grid structures

Fig. 17: Delay vs. max sink speed with different grid structures

success rate than other grid structures, it fluctuates around 0.83 when increasing the sink speed from 0 to 20 m sec^{-1}. HTTDD with equal cell side shows lower success rate due to its larger cell area.

Comparing the success ratio at speed = 0 m sec will give different values from the stationary case shown in Fig. 13 as we are using 15 random network topologies averaged to get the final result and each run randomly fails 15% of the nodes at t = 20 sec.

Figure 17 plots the delay for data delivery which increases slightly from 0.03 to 0.037 sec as sinks move faster. This shows that high-tier grid forwarding effectively localizes the impact of sink mobility. HTTDD with equal cell area shows lower delay due to shorter path between the source and the sink.
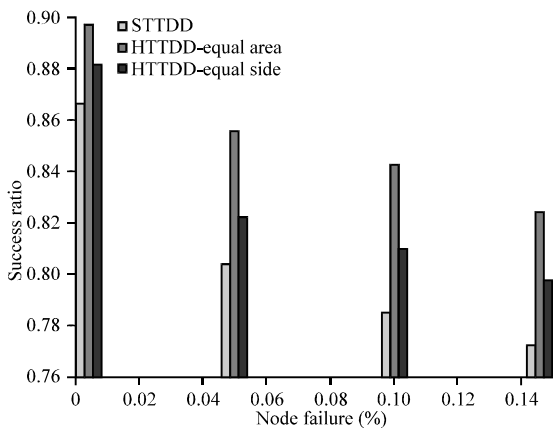
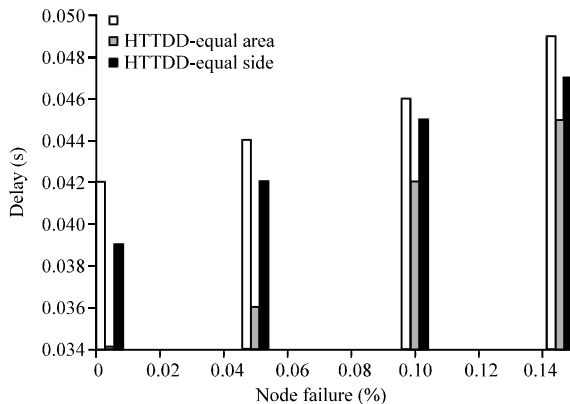Fig. 18: Success ratio vs. node failure with different grid structures

Fig. 19: Delay vs. node failure with different grid structures

Resilience to sensor node failures: We further study how node failures affect HTTDD. In the default simulation setting of 200 nodes, we let up to 15% randomly-chosen nodes to fail suddenly at t = 20 sec. The detailed study of simulation traces shows that under such scenarios, some dissemination nodes on the grid fail. Without any repairing, failures of such dissemination nodes would have stopped data delivery to all the downstream sinks and decreased the success ratio substantially. However, Fig. 18 shows that the success ratio drops mildly. This confirms that our grid maintenance mechanism effectively reduced the impact of node failures.

Because it takes time to repair failed dissemination nodes, the average delay increases slightly as more and more nodes fail, as shown Fig. 19. Overall, HTTDD is quite resilient to node failures in all simulated scenarios.

CONCLUSION

HTTDD design targets efficient data dissemination to multiple, both stationary and mobile sinks in large sensor networks. HTTDD differs from any previous work in three fundamental ways. First, HTTDD demonstrates the feasibility and benefits of building a virtual grid structure to support efficient data dissemination in large-scale sensor fields. A grid structure keeps forwarding states only in the nodes around dissemination points and only the nodes between adjacent grid point's forward queries and data. Depending on the chosen cell size, the number of nodes that keep states or forward messages can be a small fraction of the total number of sensors in the field. Second, this grid structure enables mobile sinks to continuously receive data on the move by flooding queries within a local cell only. Such local flooding minimizes the overall network load and the amount of energy needed to maintain data-forwarding paths. Third, HTTDD design incorporates efforts from both sources and sinks to accomplish efficient data delivery to mobile sinks; sources in HTTDD proactively build the grid structure to enable mobile sinks learning and receiving sensed data quickly and efficiently. HTTDD provided advantages over the STTDD and that was showed through mathematical analysis and simulation model results.

HTTDD can benefit from the cognitive radio concept as in study of Khedr *et al.* (2010) for adapting some important resources (type of modulation, code rate and number of subcarriers). The cognitive engine in HTTDD dissemination nodes controls the type of modulation, code rate and number of subcarriers in order to maintain a good system throughput without wasting available bandwidth.

REFERENCES

Changle, Li, H. Zhang, B. Hao and J. Li, 2011. A survey on routing protocols for large-scale wireless sensor networks. Sensors, 11: 3498-3526.

Choe, J. and K. Kim, 2008. EADD: Energy aware directed diffusion for wireless sensor networks. Proceedings of the International Symposium Parallel and Distributed Processing with Applications, December 10-12, 2008, Sydney NSW, Australia, pp: 779-783.

Fu, Z., P. Zerfos, H. Luo, S. Lu, L. Zhang and M. Gerla, 2003. The impact of multihop wireless channel on TCP throughput and loss. Proceedings of the 20TH Annual Joint Conference of the IEEE Computer and Communications, March 30-April 3, San Francisco, CA., USA., pp: 1744-1753.

Hua, C. and T.S.P. Yum, 2008. Optimal routing and data aggregation for maximizing lifetime of wireless sensor networks. IEEE/ACM Trans. Networking, 16: 892-903.

Jeon, H., K. Park, D.J. Hwang and H. Choo, 2009. Sink-oriented dynamic location service protocol for mobile sinks with an energy efficient grid-based approach. Sensors, 9: 1433-1453.

Khedr, M., H. Shatila and J.H. Reed, 2010. Adaptive resource management for a vague environment using cognitive radio. Proceedings of the 3rd IEEE International Conference on Computer Science and Information Technology, Volume 5, July 9-11, 2010, Chengdu, Sichuan, China, pp: 581-589.

Li, Y., C.S. Chen, Y.Q. Song, Z. Wang and Y. Sun, 2009. Enhancing real-time delivery in wireless sensor networks with two-hop information. IEEE Trans. Ind. Inform., 5: 113-122.

Lindsey, S. and C. Raghavendra, 2002. PEGASIS: Power-efficient gathering in sensor information systems. Aerospace Conf. Proc., 3: 3-1125-3-1130.

Liu, C., R. Correa, X. Li, P. Basu, B.T. Loo and Y. Mao, 2011. Declarative policy-based adaptive mobile ad hoc networking. IEEE/ACM Trans. Networking, 20: 770-783.

Mahmoud, A.K., M.E. Khedr and S. Shaaban, 2012. Hexagonal two tier data dissemination model for large scale wireless sensor networks. Proceedings of the Japan-Egypt Conference on Electronics, Communications and Computers, March 6-9, 2012, Alexandria, Egypt, pp: 138-144.

Pottie, G.J. and W.J. Kaiser, 2000. Wireless integrated network sensors. Commun. ACM, 43: 51-58.

Tang, S., J. Yuan, X.Y. Li, Y. Liu and G.H. Chen *et al.*, 2010. DAWN: Energy efficient data aggregation in WSN with mobile sinks. Proceedings of the 18th International Workshop on Quality of Service, June 16-18, 2010, Beijing, China, pp: 1-9.

Wang, F. and J. Liu, 2011. Networked wireless sensor data collection: Issues, challenges and approaches. IEEE Commun. Surv. Tutorials, 13: 673-687.

Wang, L., J. Liu and W. Wang, 2010. An improvement and simulation of LEACH protocol for wireless sensor network. Proceedings of the 1st International Conference on Pervasive Computing Signal Processing and Applications, September 17-19, 2010, Haerbin, Heilongjiang, China, pp: 444-447.

Wang, N.C. and Y.K. Chiang, 2011. Power-aware data dissemination protocol for gridbased wireless sensor networks with mobile sinks. IET Commun., 5: 2684-2691.

Ye, F., G. Zhong, S. Lu and L. Zhang, 2005. GRAdient broadcast: A robust data delivery protocol for large scale sensor networks. Wireless Networks, 11: 285-298.

Ye, F., H. Luo, J. Cheng, S. Lu and L. Zhang, 2002. A two tier data dissemination model for large scale wireless sensor networks. Proceedings of the 8th Annual International Conference on Mobile Computing and Networking. September 23-28, 2002, Atlanta, GA, USA, pp: 148-159.

Zhao, L., G. Liu, J. Chen and Z. Zhang, 2009. Flooding and directed diffusion routing algorithm in wireless sensor networks. Proceedings of the 9th International Conference Hybrid Intelligent Systems, Volume 2, August 12-14, 2009, Shenyang, Liaoning, China, pp: 235-239.