ISSN 1512-1232

Reducing Power Consumption in Wireless Sensor Networks by using Terrain Based Probabilistic Forwarding Scheme (TBPFS)

S.Anandamurugan*, Dr.C.Venkatesh**

*Senior Lecturer, Kongu Engineering College, Erode, Tamil Nadu, India
** Principal, Surya Engineering College, Erode, Tamil Nadu, India

Abstract

In this paper, we consider the problem of discovery of information in a densely deployed Wireless Sensor Network (WSN), where the initiator of search is unaware of the location of target information. We propose a protocol: Increasing Ray Search (IRS), an energy efficient and scalable search protocol. The priority of IRS is energy efficiency and sacrifices latency. The basic principle of this protocol is to route the search packet along a set of trajectories called rays that maximizes the likelihood of discovering the target information by consuming least amount of energy. The rays are organized so that if the search packet travels along all these rays, then the entire terrain area will be covered by its transmissions while minimizing the overlap of these transmissions. In this way, only a subset of total sensor nodes transmits the search packet to cover the entire terrain area while others listen. We believe that query resolution based on the principles of area coverage provides a new dimension for conquering the scale of WSN. We compare IRS with existing query resolution techniques for unknown target location such as Round Robin Search. We show by simulation that, performance improvement in total number of transmitted bytes, energy consumption, and latency with terrain size.

Keywords: *Wireless sensor networks, energy efficiency, scalability, search, querying, terrain size.*

1 INTRODUCTION

Wireless Sensor Network (WSN) [1] consists of a large number of tiny, battery-operated, possibly mobile, self-adjusting nodes with limited on-board processing, environmental sensing, and wireless communication capabilities. Of late, WSNs received significant attention from researchers as WSNs find applications spanning vast and varied areas such as habitat monitoring, object tracking, military systems, industrial automation, and home automation. Due to the advancements in chip technology, the cost of sensor nodes is gradually decreasing, thereby making the deployment of large-scale dense WSNs feasible.

Apart from sensor nodes, a typical WSN consists of one or more sink nodes. Sink nodes are powered and are storage points for most of the data emerging from environmental sensing of sensor nodes. The authors of [2], based on how data are gathered, categorize WSNs as follows:

- PUSH/CONTINUOUS COLLECTION: Sensor nodes periodically sense environment and send data to the sink node.
- PULL/QUERYING: Sensor nodes sense environment and store the information locally. On need basis, the sink node queries for the required information.
- PUSH-PULL: This paradigm involves both PUSH and PULL. Sensor nodes push the sensed events to different sensor nodes in the network in a predetermined way that is used by the search initiator for finding the target information.

In PULL paradigm, WSN can be considered as a distributed database [3] and on need basis, the sink node sends queries for data collection. Some of the factors that influence the usage of PUSH-PULL approaches are the rate of occurrence of events, the query rate, the type of events sensed, and available memory resources on sensor nodes. If the query rate is low and the rate of

ISSN 1512-1232

occurrence of events is high, or event type is audio or video, then it is clearly not feasible to store the events in multiple sensor nodes as they may consume the memory completely.

In this paper, we focus on PULL [3] and UNSTRUCTURED [4] WSNs, where the sink node sends simple and one-shot queries [5] for unique data. In UNSTRUCTURED WSNs, the search initiator (source node) has no clue about the location of target information. In the existing proposals, the cost of search (total number of transmitted bytes) increases with an increase in the sensor node density. This limits the scalability of the protocols especially for densely deployed WSNs. As we move toward dense WSN deployments [6], [7], the issue of scalability is one of the primary concerns and we address this issue by applying the principles of area coverage. We propose Increasing Ray Search (IRS) which are energy efficient query resolution protocols applicable to simple one-shot queries for unique data in UNSTRUCTURED WSNs. Unique data indicates that only one sensor node in the given WSN is capable of resolving the query. We refer to IRS protocols as IRS *variants.*

An example application where the proposed protocols are applicable is acoustic sensing and identification [8] by sensor nodes. Let us suppose that a group of sensor nodes sensed a bird chirp[9], and by using an election algorithm[10], one of them stored the voice data. Due to memory constraints of sensor nodes, it is not feasible to store the voice data in multiple sensor nodes as the number of events occurring might be large, thereby filling up the memory completely.

IRS variants operate by dividing the terrain into very narrow rectangular sections called *rays.* Each ray is characterized by a source and a destination point where the source point is the sink node and the destination point is a point on the circumference of the circular terrain. *IRS variants* search the rays in decreasing order of unexplored area covered. The *unexplored area* is the area not covered by any of the earlier searched rays. IRS sends the query packet on each ray one after the other until either the query is resolved or all the rays are explored. In each ray by exploiting the localization capabilities of sensor nodes, query packet starts from the source point of ray (sink node) and travels to the destination point bisecting the ray via beacon-less forwarding to cover the entire area of the ray. When the target node receives the query packet, a response packet is sent back to the sink node. For a fixed terrain, the number of transmissions required to cover the entire terrain area is constant, and because of this, the cost of *IRS variants* is independent of node density for a given terrain size. Hence, *IRS variants* are highly advantageous for densely deployed WSNs. We show that the cost of *IRS variants* is independent of node density via simulation. Since IRS searches rays sequentially one after the other, the latency incurred will be very high.

The remainder of the paper is organized as follows: In Section 2, we list the assumptions of the proposed *IRS variants.* Related work and existing query resolution protocols such as Expanding Ring Search, Random walk, and variants of Gossip search techniques are detailed in Section 3. In Section 4, we describe the IRS *variants* in detail. Simulation results are presented in Section 5. We then conclude the paper in Section 6.

2 ASSUMPTIONS

The terrain is considered to be circular. The sink node is static and placed at the center of the circular terrain. We assume that the radius of circular terrain is known. Since we consider a static network, this is a one-time task.

- Sensor nodes are stationary and deployed uniformly in the terrain.
- We assume unit disk model for wireless communication.
- Sensor nodes are aware of their own location coordinates. Since sensor nodes are stationary, assigning location coordinates to sensor nodes is a one-time task and is part of the initial setup of WSN.
- We consider a PULL and UNSTRUCTURED WSN, where the sink node sends simple, one-shot queries for unique data. The occurrence of events in the terrain follows a uniform random distribution.
- To relay the search packet along the rays, we assume that the density of sensor

ISSN 1512-1232

nodes is high. There are several works like [6] and [7] that assume high sensor node density.

3 Related Work

In [14], the authors presented Geographic Hash Table (GHT) approach for data-centric storage in STRUCTURED WSNs. GHT hashes keys into geographic coordinates and stores a key-value pair at the sensor node geographically nearest to the hash of its key. The search initiator directs the query to the target location based on the hashed value of key. GHT is a STRUCTURED search where the search initiator knows where exactly the event is stored (based on the hash function). But, in case of UNSTRUCTURED search, where the search initiator has no idea about the location of the target event information, GHT is not applicable.

Furthermore, in order for GHT to be applicable for a network, all events in the network should be hashed and stored in the corresponding locations. This process of hashing all events incurs additional cost, moreover, this might be a wastage in case the sink node (or search initiator) is not interested in all the events or the interests of the sink node are time varying. In Trajectory-Based Forwarding (TBF) [15], the authors presented a general framework for routing packets via a predefined curve or a trajectory.

They showed that trajectory-based schemes are a viable option for dense ad hoc networks. The authors also demonstrated the applications of TBF to unicast and multicast routing, multipath routing, discovery services, and broadcasting in ad hoc networks. But, the authors of TBF have not presented analysis or simulation results for any specific trajectory to fully understand its benefits quantitatively in terms of energy efficiency and scalability. Furthermore, the discovery mechanism presented in TBF is a generalization of an idea presented in [16] and is very different from the trajectory of IRS *variants.* The idea is that the destination nodes advertise their position along arbitrary lines and the source nodes will replace their flooding phase with a query along a different set of arbitrary lines which will eventually intersect the desired destination's line.

In Acquire mechanism [5], each intermediate node which receives the query packet collects information from its d-hop neighbors and resolves some portion of the query. The packet travels via guided or random path until it gets fully resolved. The Acquire mechanism requires nodes to collect d-hop neighbor information and is applicable to complex one-shot queries for replicated data. In Directed Diffusion protocol [17], nodes advertise their interest for named data and advertisements are distributed throughout the network. The nodes with the relevant information send data to the interested nodes after receiving the advertisements. In a way, Directed Diffusion protocol can be viewed as a publish-subscribe realization for WSNs. In [18], the authors study the performance of Directed Diffusion protocol with respect to sensor node density. They find that due to network-wide flooding used, Directed Diffusion protocol does not scale well with dense WSNs. In [19], the authors propose Geographic Random Forwarding (GeRaF), a forwarding technique based on the geographic locations of relaying nodes, and random selection of relaying nodes via contention among receivers. In GeRaF, the source broadcasts a message to a collection of potential relay nodes. The node that is closest to the destination (i.e., most geographically advantaged) is selected (in a distributed fashion) to serve as the relay node and transmits the message further. GeRaF uses an RTS/CTS-based receiver contention scheme to select the best of many potential forwarders, but prioritizes forwarders based on geographic distance. The forwarding mechanism of IRS has some similarities with that of GeRaF in terms of the selection of relay nodes based on their proximity to the destination node. However, in GeRaF, the forwarding scheme does not consider respecting a trajectory, which involves additional constraints in the selection of forwarding nodes such as they should be closer to the trajectory in addition to the destination point. Furthermore, in IRS, forwarding mechanism is only one aspect, in addition, it is also associated with a trajectory and a way to explore it.

For UNSTRUCTURED WSNs where the sink node is not aware of the location of target information, search proceeds blindly for tracing the target information. The following are most widely used techniques for searching in UNSTRUCTURED WSNs: Expanding Ring Search (ERS)

ISSN 1512-1232

[20], [21], Random walk search [22], [23], and variants of Gossip search [24], [25]. ERS is a prominent search technique used in multihop networks. It avoids network-wide broadcast by searching for the target information with increasing order of *TTL* (Time-To-Live) values. *TTL* limits the number of hops to be searched from the source node. If search fails continuously up to *TTLthreshold* hops, ERS initiates network-wide broadcast. The main disadvantage of this protocol is that it resembles flooding in the worst case. In Random walk search when a node has to forward the search packet, it randomly selects one of its neighbors and forwards the search packet to the selected neighbor. The basic idea here is the random wandering in network in search of the target information until *TTL* (number of hops) is expired or the target information is found. The main disadvantage of Random walk is that the *TTL* required for finding the target information in dense WSNs is large. We show this fact via simulations in Section 6.

In Gossip search technique [24], the source node broadcasts the search packet and all receivers of the search packet either forward with a probability *p (Gossip Probability)* or drop with a probability 1 — *p.* The main disadvantage of Gossip search is that of sending the search packet to most of the nodes even when the target information is located close to the source node.

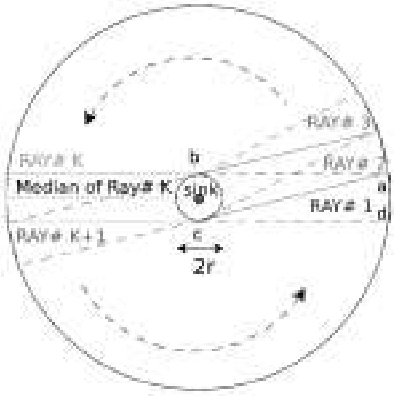

Fig. 1. Circular terrain divided into rays.

The following are the various Gossip protocols which are proposed for dense ad hoc and sensor networks: In [26], the authors presented Gossip schemes suitable to different types of Mobile Ad Hoc Networks (MANETs) such as sparse, dense, and delay-tolerant networks. Instead of fixing the *Gossip Probability,* nodes calculate it based on techniques such as number of over heard packets of the same search packet and distance from the current node to the node which relayed the search packet. In [27], the authors propose new heuristics to reduce the overhead of naive Gossip protocol and to adapt the *Gossip Probability* based on coverage area and/or topology information. Since the *Gossip Probability* is calculated based on the area coverage, the authors show that this Gossip variant is very efficient in terms of reducing overhead.

4 PROTOCOL DESIGN

The basic principle of *IRS variants* is that if a subset of the total sensor nodes transmit the search packet by suppressing the transmissions of remaining sensor nodes, such that the entire circular terrain area is covered by these transmissions, the target node which is also in this terrain will definitely receive the search packet. The selection of subset of nodes which transmit the search packet and suppression of transmissions from the remaining nodes are performed in a distributed way. However, if the search packet is broadcasted to the entire circular terrain, even though we find the target information, the number of messages required will be large. To minimize the number of message transmissions, *IRS variants* divide the circular terrain into narrow rectangular regions called *rays* such that if all these regions are covered, then the entire area of circular terrain will be covered. In IRS, the rectangular regions are covered one after the other until the target information is found or all of them are explored, whereas in k-IRS, the rectangular regions are covered simultaneously. Each ray is formed by dividing the circumference of the circular terrain into arcs of length equal to twice the transmission radius of sensor nodes and attaching the two end points of arc

ISSN 1512-1232

to the two end points of transmission diameter of the sink node as illustrated in Fig. 1. For example, *Ray#1* is formed by joining a to b and *d* to c. The width of each ray is equal to twice the transmission radius of sensor nodes.

TABLE 1. Search Packet Fields of IRS

(rX, rY)	Location coordinates of Relayed Node
(dX,dY)	Location coordinates of Destination Point
SeqNo.	Sequence number of search packet, required to avoid duplicate forwards.
Target	Target event information
Angle	Angle constraint for relaying

TABLE 2. IRS Terminology

Sink Node	Node which initiated the search
Relayed Node(R)	Node which relayed the search packet after receiving from previous Relayed or Sink Node
Current Node(C)	Node which received the search packet from Relayed Node
Destination Point (D)	Destination point of current ray.
$Dist_{r,d}$	Distance between Relayed Node and Destination Point
$Dist_{c,d}$	Distance between Current Node and Destination Point
$Dist_{c,m}$	Distance between Current Node and Median of the ray.

TABLE 3. Effect of W_1 and W_2 on T_{wait}

Node#	W_1	W_2	$Dist_{c,d}$	$Dist_{c,m}$	T_{wait}
1	0.01	0.02	10.0	10.0	0.30
2	0.01	0.02	08.0	12.0	0.32
3	0.01	0.02	12.0	08.0	0.28

The *Median* of a ray is the line joining the midpoint of the arc and the sink node. The *Median* of *Ray#K* is shown as dotted line in Fig. 1.

The fields of the search packet used by *IRS variants* are listed in Table 1 and the terminology used in explaining *IRS variants* is listed in Table 2. When one of the IRS *variants* is initiated for searching the target information, the sink node broadcasts search packet by embedding the information of first ray in it, with *Angle* = 30°. A node which receives the search packet is referred as *CurrentNode.* All *CurrentNodes* evaluate the following two conditions to check whether they are eligible to forward the search packet or not:

1. $Dist_{c;}d < Dist_{r;}d$.
2. *ffDRC < Angle.*

A node which satisfies both these conditions is referred as *EligibleNode.* The first condition makes sure that the *Eligi-bleNode* is closer to *DestinationPoint* compared to *RelayedNode.* The second condition makes sure of the following: 1) The *EligibleNodes* are closer to the *Median* of ray and 2) When the *Angle* = 30°, all nodes in the *EligibleNodes* set are in the transmission range of each other [28]. An *EligibleNode* has to wait for a time proportional to its proximity to the *DestinationPoint* and the *Median* of ray before relaying the packet. When an *EligibleNode* relays the search packet, all other *EligibleNodes* which receive this packet, drop the packet which should be relayed by them. The time to wait before relaying is given by

$$T_{wait} = W_1 * Dist_{c,d} + W_2 * Dist_{c,m},$$

where *W1* and W_2 are the weight factors and by substituting them with appropriate values, the weights of $Dist_{c;}d$ and $Dist_{c;m}$ can be adjusted. The nodes with low $Dist_{c;}d$ and $Dist_{c;}m$ values will have lesser waiting time compared to the other nodes and among them, by giving more weight to $Dist_{c;m}$, the node which is closer to the *Median* of ray is given more preference to relay the search packet. The values of *W1* and *W2* are instrumental in choosing the path which a ray can take. In our work, to completely cover the terrain area with broadcast transmissions of the search packet, the trajectory should be respected. Therefore, we give more weight to $Dist_{c;m}$ than to $Dist_{c;d}$, so that the node closer to *Median* times-out before the other nodes with a high probability, and thereby, trajectory is respected. Table 3 shows an example of how T_{wait} is affected based on the values of *W1* and *W2.* From the values of first and second rows, we can observe that due to higher W_2 value, the T_{wait} of *Node* 1 is lesser than that of *Node* 2 even though *Node* 1 has higher $Dist_{c;d}$ than *Node* 2. From the values of first and third rows, we can observe that due to higher W_2 value, the T_{wait} of *Node* 3 is lesser than that of *Node* 1, even though *Node* 3 has a higher $Dist_{c;d}$ than *Node* 1. This example demonstrates the fact that, by carefully selecting *W1* and W_2 values, the sensor nodes which are closer to the *Medians* of rays can be given preference over the other nodes. It should be noted that the above procedure does not guarantee that nodes closer to *Median* always time-out before the other nodes, however, by using the above procedure, this phenomenon happens with a very high probability.

In order to reduce the waiting time of sensor nodes, we map the T_{wait} values to an interval $[t_{min}, t_{max}]$. The value of *tmax* should be high enough to avoid relaying of the search packet by multiple *EligibleNodes* and at the same time, it should not be too high as this might result in high-latency values. We will discuss more about this interval in Section 5. Based on the values of *W1* and W_2, the *EligibleNode* which is closer to the *Median* of current ray and among them the one closer to the *DestinationPoint* will have lesser waiting time compared to the other *EligibleNodes.* The *EligibleNode* with smaller waiting time relays the search packet while others drop it. This continues until the search packet reaches the *DestinationPoint* or there is no other node to relay the search packet further.

The node *R* which is the *RelayedNode* relays the search packet and all the nodes in its transmission range which receive the search packet are called *CurrentNodes.* Based on the first condition, the nodes which are farther to the *DestinationPoint D* compared to the node *R* are filtered out from the *CurrentNodes,* i.e., nodes which are toward the right side (toward *DestinationPoint D)* of node *R* in the figure will go ahead to evaluate the second condition and the other nodes will drop the search packet. Based on the second condition, each node which passed the first condition will check the angle made by the line joining the *RelayedNode* and itself with the line segment *RD.* For example, in the figure, the angle made by the line joining *R* and node *C* with the line segment *RD* is *j3.* The measured angle should be less than *Angle (a* in the figure). Based on this condition, more nodes are filtered from the *CurrentNodes* set. Now, the remaining nodes are called *EligibleNodes.* When the *Angle* is set to 30°, based on the principles of geometry, the *4XRY* becomes equilateral and the *EligibleNodes* set is formed such that if one of them relays the search packet, all other nodes in the set will receive it and drop the same search packet which should be relayed by them. Based on the time-out value, the node *C* relays the search packet and all other *EligibleNodes* will receive this search packet and drop the packets which should be relayed by them.

The idea behind forwarding via the *Median* of ray is to keep track of which areas of terrain are covered and also to cover the entire region of ray. The suppression scheme mentioned above is one of the key factors for achieving scalability with the proposed protocol. The protocol used for forwarding the search packet is similar to the position-based beacon-less routing [28] with some modifications. The *RelayedNode* waits for a time-out t_{max} to overhear the transmission by one of the *EligibleNodes.* If the *RelayedNode* overhears the transmission by an *EligibleNode* within the time t_{max}, then it drops the search packet from its cache, and if there are no nodes eligible to relay the search packet, the *RelayedNode* rebroadcasts the packet without angle constraint, i.e., *Angle* =360° and drops the search packet from its cache, and no more retransmissions are performed after this. When the *Angle* is set to 360°, there might be multiple *EligibleNodes* relaying the search packet, since some of them may not overhear other *EligibleNode* transmissions. This is a compromise as

ISSN 1512-1232

there are no *EligibleNodes* to relay the search packet. Under high node density, the probability of not finding *EligibleNodes* with 30° angle constraint is very low and we validate this via simulations in Section 5. When the target node is in the region of the current ray, it receives the search packet and responds to the sink node by sending a response packet. The sink node continues to search until all the rays are explored or the target information is found.

4.1 Ordering of the Rays

We sort the rays in decreasing order of unexplored area covered by them. This ordering is used by IRS *variants* while searching for the target information. After the division of circular terrain into rays, the area covered by each ray is equal. However, the area covered by a ray which is not covered by any of the rays previously searched is not same for all the rays. We call this area as unexplored area. For example, in Fig. 1, let us assume that rays are ordered according to ray number. Clearly, the unexplored area covered by *Ray#1* is more than the unexplored area covered by *Ray#2.* At any point in the order, the next ray is the one which covers the maximum unexplored area of all the remaining unsearched rays. In this way, the rays are ordered in decreasing order of unexplored area covered by them. In Fig. 1, the next ray in the order after *Ray#1* is definitely not *Ray#2,* as there are other rays which cover more unexplored area than *Ray#2. Ray#K* covers the most unexplored area compared to all other unsearched rays. There may be multiple rays that cover the most unexplored area at a given point in the order, in this case, one of them is selected as the next ray to be searched. This pattern of ray ordering or ray growth is called Greedy Ray Growth (GRG), as it tries to maximize the probability of finding the target node in rays searched as early as possible. The numbers on rays in this figure indicate the order of search. We quantify the advantages of using GRG over sequential ray growth via simulations in Section 5.

4.2 Grouping of the Rays

We now segregate the rays into groups ordered according to the unexplored area covered. All rays in a group cover equal unexplored area. The first ray explored will be part of *Group1* and this will be the only ray in the group. Since, this is the first ray, the unexplored area covered by this ray is maximum compared to any other ray. The unexplored area covered by the first ray searched is same as the total area of the ray which is given by

$$A_1 = (R+r)\ x2r = 2ðnþ1Þr^2 \qquad (1)$$

While deriving the unexplored area covered by a ray, we have to consider the overlap with previous rays. The unexplored area covered by the ray in the first group is same as the total area of the ray. For any of the rays in other groups, the unexplored area covered should exclude the area covered by previous rays. In deriving the unexplored area covered by the rays in second group, the length of the rectangle (ray) should be taken as *ðR — rÞ,* whereas for the ray in first group, the length of the rectangle is *ðRþrÞ.* After the first ray, the next maximum unexplored area is covered by rays which are either 90° or 180° to the first ray, i.e., Rays 2, 3, and 4 in Fig. 4. We add these rays to for which the area covered is given by

$$A_2 = ðR - r)\ \mathrm{x}\ 2r = 2ðn - 1Þr^2. \qquad (2)$$

At this stage, there are four rays in the circular terrain. These four rays divide the circular terrain into four approximately equal sectors. The next ray to maximize the unexplored area is always the one with its *Median* stretching from the sink node to the midpoint of circumference of any one of the newly formed sectors. The four additional rays will create eight new sectors and this process continues until the entire circular terrain area is covered. Based on the above pattern, we can generalize the following:

The maximum number of rays in *Group1 ; Group2; Group3 ; Group4 ; Group5 ;...; etc:* will be 1; 3; 4; 8; 16;...; etc., respectively.

The total number of groups:

$$G= [\log_2 [\pi\eta]] \qquad (3)$$

ISSN 1512-1232

From (1) and (2), we can generalize the unexplored area covered by a ray in $Group_i$ as:

$$A_{i=}2(n-(2i-3))r^2 \qquad (4)$$

4.3 IRS

IRS explores the rays one after the other according to GRG. When IRS is initiated, the sink node sends the search packet to the ray which covers the maximum unexplored area. Then the sink node waits for a time-out *NRwait* before sending the search packet to the next ray. The NR_{wait} value should be carefully estimated based on the radius of circular terrain as a high value of NR_{wait} results in high latency of search and a low value of NR_{wait} results in high cost of search. If the sink node receives acknowledgment from the target node within this time-out value, it stops the search; otherwise it continues with the next ray. IRS explores groups starting with *Group1* and in a single group, no specific ordering is followed and rays are explored sequentially. Due to the conservative nature of IRS in exploring rays, it consumes the least energy of all *IRS variants,* but incurs high latency. Since IRS always chooses the sensor nodes closest to the *Medians* of rays for forwarding the search packet, the energy depletion of these nodes will be high compared to the other sensor nodes. One possible way of alleviating this problem is to find a different set of rays for each search so that the *DestinationPoints* and the *Medians* of rays will be different for each search, thereby, load gets distributed among all the sensor nodes.

In Section 4.2, we have explained about grouping of rays based on the unexplored area covered. Each group contains rays that individually cover equal unexplored area. Let us now consider exploring all rays in a single group together, i.e., in successive iterations, number of rays which are explored in parallel are 1, 3, 4, 8,16 ... etc., until the target information is found or all rays are explored. One can observe that the parallelism of rays grows exponentially. We call this way of exploring rays as exponential-IRS. Intuitively, since the amount of parallelism in exponential-IRS is less than flood-IRS, the cost of the former should be less than that of the latter. Likewise, the latency of exponential-IRS should be more than that of flood-IRS.

In exponential-IRS, the number of groups which are explored in parallel is always one, i.e., at any instance of time, all rays from a single group are explored together. Based on this logic, exponential-IRS can also be called as 1-IRS and flood-IRS can also be called as G-IRS.

5 SIMULATIONS

To validate the results presented in the previous section, we simulate RRS [27] in the ns-2 [32], a popular discrete event network simulator. We refer to IRS protocols as *IRS variants.* In [27], the authors propose a variant of Gossip, Diagonal Area and Copies Coverage-based Probabilistic Forwarding (DACCPF) which includes most of the features of the other Gossip variants for dense WSNs [26] such as probabilistic forwarding, counter-based forwarding, and distance-based forwarding, and in addition to these, it also has area coverage-based probabilistic forwarding which makes it scalable for dense WSNs. In DACCPF, we calculate the distance from the relayed node to the received node using the location awareness of the sensor nodes. The reason for selecting DACCPF compared to the other optimizations proposed in [27] is that DACCPF performs better than the other variants in their simulations in terms of cost.

5.1 Performance Metrics Used

- *Number of transmitted bytes:* Average number of bytes transmitted by all the nodes in the network for finding the target information. As the message formats are not uniform across protocols, we measured the number of bytes transmitted instead of the number of messages transmitted.
- *Number of transmitted and received bytes:* Average number of transmitted and received bytes by all the nodes in the network for finding the target information.
- *Energy consumed:* The total energy consumed by all the nodes in the network for finding the

target information.

- *Latency:* Time taken to find the target information, i.e., the time difference between, the time at which the search is initiated by the sink node by transmitting the search packet, and the time at which the search packet is received by the target node.
- *Probability of finding the target information:* Probability of finding the target information is a measure of the success probability of the search protocols. It is also a measure of non determinism of the search protocols.

5.2 Simulation Setup

We consider the terrain to be circular, where the sink node is static and placed at the center of the terrain. Sensor nodes are static and uniformly deployed in the given terrain. The transmission radius (radio range) of sensor nodes is fixed at 30 m. The propagation model used is *TwoRayGround.* The *TTLthreshold* value for ERS is fixed at 3 as this value is found to be optimum [21]. The *Gossip Probability* is fixed at 0.65 as this value is found to gossip the search message to most of the nodes in a given WSN [24]. The *TTL* value for Random walk protocol is set to twice the number of sensor nodes deployed in the terrain. The simulation parameters used for DACCPF are: $k1 = 0.7$, $k2 = 0.175$, $p = 0.7$, and $t = 240$ msec, where k_1 ;k2 ^{ar}e node passivity parameters, t is the upper limit of random wait time for over hearing messages from neighbor nodes, and p is the *Gossip Probability.* The actual forwarding probability is calculated based on DACCPF algorithm and using the above mentioned parameters. The values used for DACCPF are based on the results in [27] and also to keep the packet delivery ratio to a reasonably high value. In *IRS variants,* t_{max} is set to 280 msec, *NRwait* (time between searching rays) is set to t_{max} x n, and n is calculated based on the terrain radius as in (6). We consider an energy model based on the power model of Mica2 mote [33] where the current consumption for reception is 7.0 mA and for transmission *(þ4dBm)* is 11.6 mA with a 3 V power supply. The MAC protocol used for all search techniques is IEEE 802.11. All the graphs for the performance metrics are plotted for 95 percent confidence level. We first consider a terrain of fixed radius and vary node density to study the effect of variation of node density on the performance metrics. Then, we fix density and vary the terrain radius to study the effect of increase in the terrain size on the performance metrics. For density variation scenario, the number of nodes is varied from 250 to 2,000, in increments of 250 and the circular terrain diameter is fixed at 600 m. For terrain size variation scenario, the diameter of the circular terrain is varied from 100 m to 500 m and the density of sensor nodes is fixed to a high value at $p = 0.00884$ nodes/sq.m.

5.3 Simulation Results

5.3.1 *Terrain Size Variation*

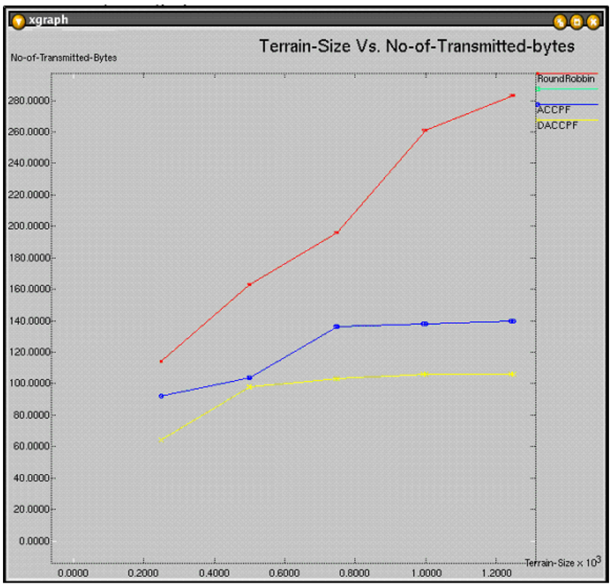

Fig. 2. Effect of terrain size on number of bytes transmitted for finding the target information.

ISSN 1512-1232

From Fig.2, we can infer that the cost of all the search techniques increases with increase in terrain size. However, *IRS variants* consume the least cost of all the search techniques, irrespective of the terrain size. Fig. 3 shows the effect of terrain size on the energy consumption of all nodes in the network. We can observe that the energy consumption of *IRS variants* is the least of all the search techniques. Fig. 4 shows that the latency of IRS is very high compared to the other proposals and the latency increases with terrain size, as the rays grow in length. These observations are in accordance with the analytical results: Under high node densities, 1) the cost and latency of IRS *variants* increase with increase in terrain size.

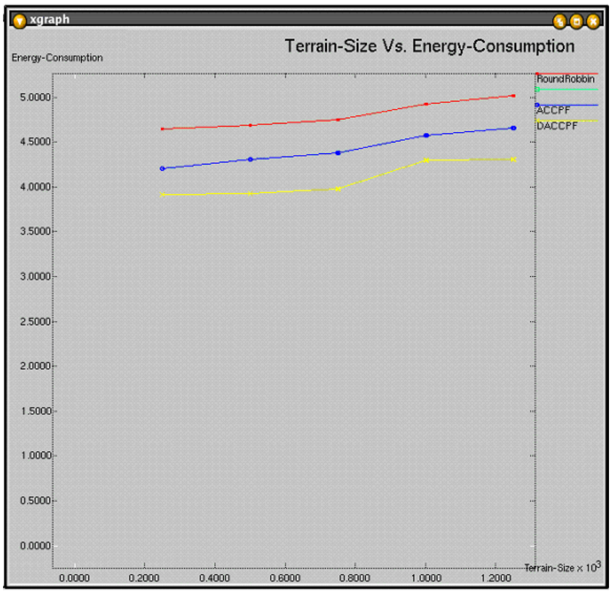

Fig. 3. Terrain size versus total energy consumed by all nodes.

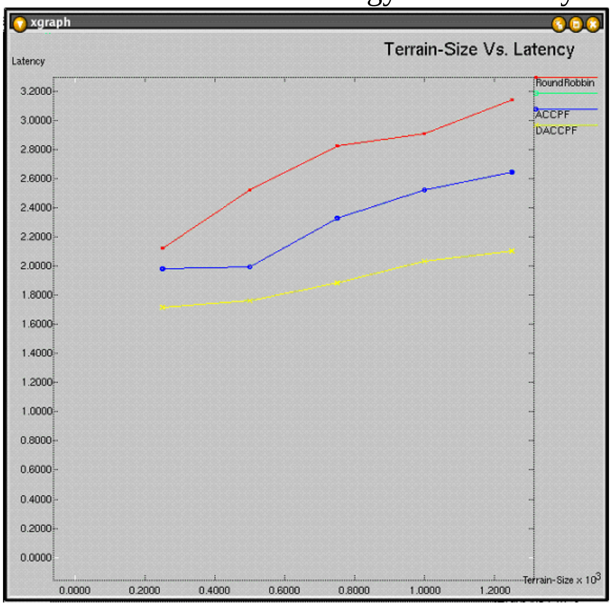

Fig. 4. Effect of terrain size on latency (on logarithmic scale) of search.

5.4 Inference from Simulation

The cost of ERS, variants of Gossip, and Random walk protocols increase with increase in node density and terrain size whereas the cost and latency of IRS *variants* are independent of node density at high node densities, but increase with terrain size. Among the existing proposals, the

ISSN 1512-1232

performance of DACCPF is close to *IRS variants* compared to the other protocols, but DACCPF incurs much more cost than that of *IRS variants.*

6 CONCLUSION

Information discovery in UNSTRUCTURED WSNs via querying is a key aspect. In this paper, we presented IRS that is energy efficient and scalable query resolution protocols for simple, one-shot queries for unique data. The conclusion drawn from the paper is that under high node density, *IRS variants* consume much less cost compared to the existing search techniques such as RRS,ERS, Random walk, and variants of Gossip protocols and it is unaffected by the variation in node density. We believe that query resolution based on the principles of area coverage provides a new dimension for enhancing the scalability of query protocols in WSNs. We validated our claims by simulation. The following are the main advantages of IRS *variants:*

- Energy is the most premium resource in WSNs and *IRS variants* achieve significant energy savings for dense WSNs.
- Cost of search is unaffected by variation in node density, which makes *IRS variants* scalable for highly dense WSNs.

ISSN 1512-1232

REFERENCES

[1] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A Survey on Sensor Networks," *IEEE Comm. Magazine,* vol. 40, no. 8, pp. 102-114, Aug. 2002.
[2] X. Liu, Q. Huang, and Y. Zhang, "Combs, Needles, Haystacks: Balancing Push and Pull for Discovery in Large-Scale Sensor Networks," *Proc. Second Int'l Conf. Embedded Networked Sensor Systems (SenSys '04),* pp. 122-133, Nov. 2004.
[3] R. Govindan, J. Hellerstein, W. Hong, S. Madden, M. Franklin, and S. Shenker, "The Sensor Network as a Database," Technical Report 02-771, Computer Science Dept., Univ. of Southern California, Sept. 2002.
[4] J. Ahn and B. Krishnamachari, "Modeling Search Costs in Wireless Sensor Networks," Technical Report CENG-2007-1, Computer Science Dept., University. of Southern California, 2007.
[5] N. Sadagopan, B. Krishnamachari, and A. Helmy, "The Acquire Mechanism for Efficient Querying in Sensor Networks," *Proc. First IEEE Int'l Workshop Sensor Network Protocols and Applications (SNPA '03),* pp. 149-155, May 2003.
[6] T.Q. Quek, D. Dardari, and M.Z. Win, "Energy Efficiency of Dense Wireless Sensor Networks: To Cooperate or Not to Cooperate," *IEEE J. Selected Areas in Comm.,* vol. 25, no. 2, pp. 459-470, Feb. 2007.
[7] G. Halkes and K. Langendoen, "Crankshaft: An Energy-Efficient MAC-Protocol for Dense Wireless Sensor Networks," *Proc. Fourth European Conf. Wireless Sensor Networks (EWSN '07),* pp. 228-244, Jan. 2007.
[8] J. Cai, D. Ee, B. Pham, P. Roe, and J. Zhang, "Sensor Network for the Monitoring of Ecosystem: Bird Species Recognition," *Proc. Third Int'l Conf. Intelligent Sensors, Sensor Networks and Information (ISSNIP '07),* pp. 293-298, Dec. 2007.
[9] C.-E. Chen, A.M. Ali, and H. Wang, "Design and Testing of Robust Acoustic Arrays for Localization and Enhancement of Several Bird Sources," *Proc. Fifth Int'l Conf. Information Processing in Sensor Networks (IPSN '06),* pp. 268-275, Apr. 2006.
[10] Q. Wang, R. Zheng, A. Tirumala, X. Liu, and L. Sha, "Lightning: A Hard Real-Time, Fast, and Lightweight Low-End Wireless Sensor Election Protocol for Acoustic Event Localization," *IEEE Trans. Mobile Computing,* vol. 7, no. 5, pp. 570-584, May 2008.
[11] Crossbow MICAz Mote, http://www.xbow.com/Products/ productdetails.aspx?sid=164, 2009.
[12] Moteiv Tmote Sky Mote, http://www.sentilla.com/pdf/eol/ tmote-sky-datasheet.pdf, 2009.
[13] G. Werner-Allen, K. Lorincz, M. Welsh, O. Marcillo, J. Johnson, M. Ruiz, and J. Lees, "Deploying a Wireless Sensor Network on an Active Volcano," *IEEE Internet Computing,* vol. 10, no. 2, pp. 18-25, Mar./Apr. 2006.
[14] S. Ratnasamy, B. Karp, L. Yin, F. Yu, D. Estrin, R. Govindan, and S. Shenker, "GHT: A Geographic Hash Table for Data-Centric Storage," *Proc. First ACM Int'l Workshop Wireless Sensor Networks and Applications (WSNA '02),* pp. 78-87, Sept. 2002.
[15] D. Niculescu and B. Nath, "Trajectory Based Forwarding and Its Applications," *Proc. Ninth Ann. Int'l Conf. Mobile Computing and Networking (MobiCom '03),* pp. 260-272, Sept. 2003.
[16] I. Stojmenovic, "A Scalable Quorum Based Location Update Scheme for Routing in Ad Hoc Wireless Networks," Technical Report TR-99-09, SITE, Univ. of Ottawa, Sept. 1999.
[17] C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann, and F. Silva, "Directed Diffusion for Wireless Sensor Networking," *IEEE/ ACM Trans. Networking,* vol. 11, no. 1, pp. 2-16, Feb. 2003.
[18] L. Khelladi and N. Badache, "On the Performance of Directed Diffusion in Dense Sensor Networks," *Proc. Fourth Int'l Conf. Innovations in Information Technology (Innovations '07),* pp. 113-117, Nov. 2007.
[19] M. Zorzi and R. Rao, "Geographic Random Forwarding (GeRaF) for Ad Hoc and Sensor Networks: Energy and Latency Performance," *IEEE Trans. Mobile Computing,* vol. 2, no. 4, pp. 349-365, Oct.-Dec. 2003.

ISSN 1512-1232

[20] N.B. Chang and M. Liu, "Controlled Flooding Search in a Large Network," *IEEE/ACM Trans. Networking,* vol. 15, no. 2, pp. 436-449, Apr. 2007.
[21] J. Hassan and S. Jha, "Optimising Expanding Ring Search for Multi-Hop Wireless Networks," *Proc. 47th Ann. IEEE Global Telecomm. Conf. (GLOBECOM '04),* pp. 1061-1065, Nov. 2004.
[22] L. Lima and J. Barros, "Random Walks on Sensor Networks," *Proc. Fifth Int'l Symp. Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt '07),* pp. 1-5, Apr. 2007.
[23] H. Tian, H. Shen, and T. Matsuzawa, "Random Walk Routing for Wireless Sensor Networks," *Proc. Sixth Int'l Conf. Parallel and Distributed Computing Applications and Technologies (PDCAT '05),* pp. 196-200, Dec. 2005.
[24] Z.J. Haas, J.Y. Halpern, and L. Li, "Gossip-Based Ad Hoc Routing," *IEEE/ACM Trans. Networking,* vol. 14, no. 3, pp. 479-491, 2006.
[25] A.V. Kini, V. Veeraraghavan, N. Singhal, and S. Weber, "Smartgossip: An Improved Randomized Broadcast Protocol for Sensor Networks," *Proc. Fifth Int'l Conf. Information Processing in Sensor Networks (IPSN '06),* pp. 210-217, Apr. 2006.
[26] R. Friedman, D. Gavidia, L. Rodrigues, A.C. Viana, and S. Voulgaris, "Gossiping on MANETs: The Beauty and the Beast," *ACM SIGOPS Operating Systems Rev.,* vol. 41, no. 5, pp. 67-74, 2007.
[27] H. Ling, D. Mosse, and T. Znati, "Coverage-Based Probabilistic Forwarding in Ad Hoc Routing," *Proc. 14th Int'l Conf. Computer Comm. and Networks (ICCCN '05),* pp. 13-18, Oct. 2005.
[28] M. Uttel and T. Braun, "A Novel Position-Based and Beacon-Less Routing Algorithm for Mobile Ad-Hoc Networks," *Proc. Third IEEE Workshop Applications and Services in Wireless Networks (ASWN '03),* pp. 197-210, July 2003.
[29] A. Rao, S. Ratnasamy, C. Papadimitriou, S. Shenker, and I. Stoica, "Geographic Routing without Location Information," *Proc. Ninth Ann. Int'l Conf. Mobile Computing and Networking (MobiCom '03),* pp. 96-108, Sept. 2003.
[30] R. Kiran Kumar, "Information Discovery in Wireless Sensor Networks: Energy Efficiency and Scalability," master's thesis, Dept. of Computer Science and Eng., Indian Inst. of Technology Madras, Dec. 2008.
[31] MATLAB, http://www.mathworks.com/products/matlab, 2009.
[32] The Network Simulator, http://www.isi.edu/nsnam/ns, 2009.
[33] V. Shnayder, M. Hempstead, B. rong Chen, G.W. Allen, and M. Welsh, "Simulating the Power Consumption of Large-Scale Sensor Network Applications," *Proc. Second Int'l Conf. Embedded Networked Sensor Systems (SenSys '04),* pp. 188-200, Nov. 2004.

Article received: 2009-12-13