Spatial-Temporal Sink-Proxy of the Intelligent-Sink for On-demand Information Retrieval in Sensor Network

Rui Teng
National Institute of Information and Communications Technology (NICT), Japan
teng@nict.go.jp

Bing Zhang
National Institute of Information and Communications Technology (NICT), Japan
zhang@nict.go.jp

ABSTRACT
This paper introduces the usage and selection of spatial-temporal sink proxy for on-demand information retrieval in sensor networks. A sink proxy is a sensor node selected among sensor nodes by the sink node for local query distribution and data collection. Such a sink node, termed the intelligent sink, utilizes the correlation of query content and individual sensor nodes to specify the corresponding sensor nodes of user interest, and then selects the sink proxy according to the locations of sensor nodes. The proposed scheme achieves highly distributed and low energy consumption among sensor nodes. The evaluation results reveal that the proposed sink-proxy scheme reduces the energy consumption of information retrieval in the network and obtains balanced energy consumption among sensor nodes.

Keywords
Sensor networks, information retrieval, intelligent sink, sink proxy

1. INTRODUCTION
On-demand information retrieval from sensor networks consists of two phases of operation: data query and data collection. A user initiates the process by sending a query message to sensor networks. Upon receiving the query message, each sensor node that matches the query delivers its data to the sink node. In contrast to conventional end-to-end communication systems, sensor networks adopt data-centric communication, in which the destination of a query is not described by sensors' identifiers but rather by the users' interest as represented by the low-level naming of data attributes in query contents [1], [2]. For example, a query can be"what is the temperature on the first floors".

Although the existing approaches of data query and collection can reduce energy consumption by utilizing efficient broadcasting and cooperative data collection among sensor nodes, etc., most of these approaches adopt the direct use of sink node[3], [4], [5], [7], [6]. That is, it is at the sink node that a query message is disseminated to sensor nodes and most data are accumulated. Such direct use of sink node leads to the high and unbalanced energy consumption in the case of on-demand information retrieval, and more energy are consumed near the sink node.

In this paper, we introduce a design of intelligent sink node mechanism which utilizes spatial-temporal sink proxies to localize data query and data collection. Rather than only serving as a relay router between user and sensor nodes, the designed intelligent sink node attempts to specify the correlation between query and its corresponding sensor nodes, and performs local on-demand data query and collection. A sink proxy is selected among the sensor nodes corresponding to each query. The intelligent sink node utilizes sink proxies among sensor nodes to localize the data query and data collection in a certain small area, leading to the small and balanced energy consumption among individual sensor nodes. Since the position and ID of sink proxy varies with the time passing according to each query's content, hence we call it "*Spatial-Temporal Sink Proxy*"

We propose a sink proxy selection scheme for minimum energy cost in data query and data collection. The proposed scheme utilizes the comparison between costs of using different candidate sink proxies, and selects a smallest cost one as the sink proxy. Since this calculation of comparison is conducted at the sink node, which is a powerful node, this scheme takes little cost at sensor node.

The rest of this paper is organized as follows. Section 2 introduces the intelligent sink node for sensor networks. Section 3 describes the sink proxy selection scheme, Section 4 presents the simulation results. Section V concludes.

2. INTELIGETN SINK FOR SENSOR NETWORK: FOUNDERMENTALS

2.1 Approach overview
The overview of our proposed intelligent sink mechanism is shown in Figure1. The proposed intelligent sink utilizes a query resolution scheme to specify the correlation between query and its corresponding sensor nodes.

According to the location information of individual sensor node, a sink proxy is selected among the corresponding sensor nodes. By utilizing the sink proxy, data query and collections is localized to a certain local sensor node group. The intelligent sink utilizes an efficient on-demand data query and collection scheme: LDQD (Localized Data Query Distribution) and LSDC (Localized Sensing Data Collection). In the LDQD scheme, instead of broadly disseminating the query to all sensor nodes, a query message is first unicasted to a sink proxy, which is a local node residing in the area where the corresponding sensor nodes reside.

Bionetics'08, November 25-28, 2008, Hyogo, Japan.

Then, query is geocasted to nodes in the query area. To achieve uniform energy distribution among sensor node, the sink proxy spatial-temporal rotates among sensor nodes.

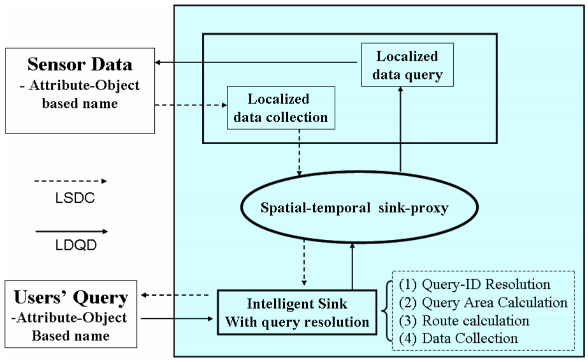


Figure 1 Approach overview of the intelligent sink node

2.2 Query resolution at the sink node

We utilize Attribute-Objects based name (AO name) to represent the query's interest and the property of sensor nodes. That is, a user's query can be described by the interested object that sensor nodes are monitoring, and the sensing data attribute that a user is interested in. In the same way, a sensor node can be described by the object that it is monitoring, and the physical attribute such as temperature, humidity that describes a sensor's type. In the AO name of query, a specification of interested data can also be included, similar to that introduced in [8].

The resolution table contains attributes that a sensor node contains, such as temperature, humidity, etc, and the objects that a sensor is monitoring. Thus, according to a query's name, the sensor nodes corresponding to a query can be specified to a particular sensor node group.

According to the IDs of corresponding sensor nodes, a sink node calculates the query area based on node locations. The query area gives a region where a query message can be distributed. In LDQD, the query area is calculated by a rectangle in which all corresponding sensor nodes reside. Such a query area is calculated by the following formula:

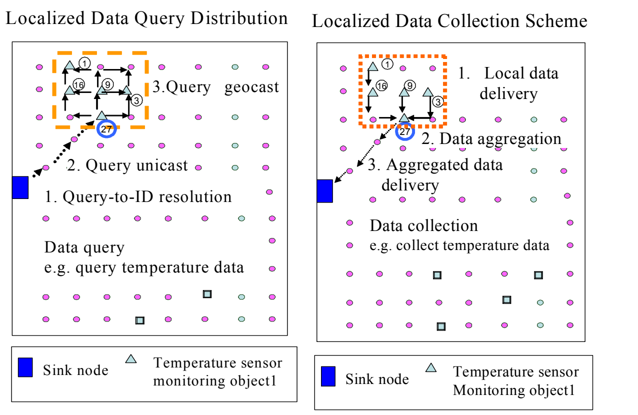
$$Query\_Area=Rect[min(x1,x2,...xn),\ min(y1,y2,...yn),\ max(x1,x2,...xn),\ max(y1,y2,...yn)],$$

where n is the number of corresponding sensor nodes.

2.3 Localized data query and collection

Localized Data Query Distribution (LDQD) consists of two steps: Query unicast, and Query geocast, as shown in Figure 2(a). The unicast distribution is utilized to deliver query messages from a sink node to the query area. The query message is unicasted from the sink node to a sink proxy, which is defined as the node in the query area, e.g. node 27 in the example shown in Figure 2 (a). After the query message arrives the sink proxy in the query area, it is geographically broadcasted to all sensor nodes inside the query area. [9].

Localized Sensing Data Collection (LSDC) consists of three steps: local data delivery, data aggregation, and aggregated data delivery to the sink node, as shown in Figure 2(b). Upon receiving a query message, the corresponding sensor nodes send the sensing data required by the query, back to the sink proxy in a local region. This is achieved by using the reverse path got from the query geocast initiated by the sink proxy [10].



(a) (b)

Figure 2 Localized data query and collection

The sink proxy, such as node 27 shown in Figure 2, collects the sensing data locally before forwarding it to the sink node. The sink proxy aggregates the sensing data by placing multiple sensing data into one packet and delivers the aggregated data to the sink node.

3. SINK-PROXY SELECTION SCHEME

At first, we analyze and describe the energy cost model in the data query and data collection with respect to a certain sink proxy. There are four routing components on which the energy is consumed during the query and data collection, as shown in the Figure 3. The first is the query unicast route (R1) from the sink to the proxy. The second is the broadcast routes (R2) from proxy to the corresponding query area. The third is the data collection routes (R3)from corresponding sensor nodes to the sink proxy. The fourth is the unicast route (R4) from sink proxy to the sink node, for the delivery of aggregated data.

We use E1 and E2 to denote the energy costs in the data query on routes of R1 and R2 respectively, and E3 and E4 to denote the energy cost in data collection on routes of R3 and R4 respectively. We describe the cost at with regard to a sink proxy SPi as follows.

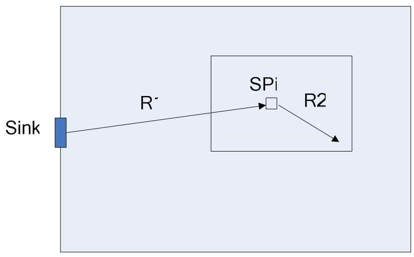
$$E1(SPi) = E(R1, SPi) = E(Query\_Unicast\_from\_Sink\_To\_SPi)$$

$$E2(SPi) = E(R2, SPi) = E(SPi\_Loca\_Broadcast\ )$$

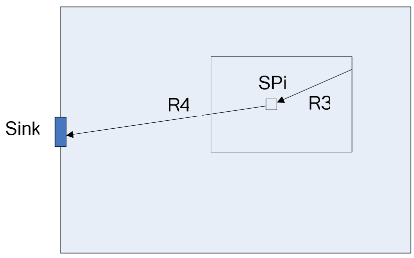
$$E3(SPi) = E(R3, SPi)$$
$$= \sum_{m=1}^{k} E(Unicast\_of\_Corresponding\_Node[m]\_to\_SPi)$$
$$E4(SPi) = E(R4, SPi)$$
$$= L * E(Unicast\_from\_SPi\_to\_Sink)$$

Where $L = \frac{Collected\_Data\_Size}{Aggregation\_Capability}$; L is the number of transmission packets of collected data in the unicast from sink proxy to the sink node, the aggregation capability is the max number of individual sensor data that can be collected in the payload of a packet at the sink proxy.

$$E(SPi) = E1(SPi) + E2(SPi) + E3(SPi) + E4(SPi)$$



(a) Data query



(b) Data collection

Figure 3 Four routing components

The sink proxy selection procedure is described in Algorithm1, 2,3,4. Algorithm1 described the overall scheme of sink proxy section method among sensor nodes. Algorithms 2 and 3 described the calculation of the cost of data collection (including E3 and E4), and data query (including E1 and E2), respectively.

The candidate sink proxies include the corresponding nodes to the query, and also the sink node itself. This allows the approach includes the opportunity of using the conventional direct sink mechanism of data query and collection.

Algorithm1 Sink proxy selection

Input: Candidate sink proxy; Sink node,
Output: The sink proxy with lowest cost

```
Set  E0= E(SP[0]);
For Sink_proxy_candidate_node [i]=1 to M do // M can be either all node or can be selected nodes
   {
     if (E(SP[I])<E0)
        {
          the proxy sink ID I0=I;
          E0= E(SP[I]) ;
        }
      End if
End for
The lowest cost sink proxy's ID = I0.
End
```

Algorithm2 Data collection cost

Input: Candidate sink proxy SPi; sink resolved query area
Output: The cost of data collection

```
Set E3(SPi)=0; and E4(SPi)=0;
For sensor node I=0 to N do // nodes in the sensor netowrks
   {
    If (node I is in the query area)
         { E3(SPi)=E3(SPi)+E3(I,SPi)}
          // E3(I,SPi) is cost of unicast from node I to SPi
     Endif
   }
End for
 E4(I)=E(Sink,SPi)  //. Described in algorithm 4
 Ec(I)=E3(I)+E4(I)
End
```

Algorithm3 Data query cost

Input: Candidate sink proxy; Sink node, resolved query area
Output: The cost of data query

```
Set E1(I)=0 and  E2(I)=0;
Set Eb(I)=E(TX+cRX) // Eb is the cost of one hop broadcast at node I , where c is the the average number of neighbor nodes of a sensor node.
For sensor node I=0 to N do  //nodes in the sensor netowrks
   {
    If (node I is in the query area)
         {E2(SPi)=E2(SPi)+Eb(I)}
    Endif
   }
End for
 E1(I)=E(Sink,SPi)
 E(SPi)=E1(SPi)+E2(SPi)
End
```

Algorithm4 Basic unicast cost

Input: Souce node (SRC) ID; Destination node (DST) ID, radio distance d0
Output: The cost of routing from SRC to DST cost(S,D)

```
set s=src, E(s,d)= Eb, and NodeX=src; dm=d0;
set D=(Distance between NodeX to DST)
While (distance to the destination (D>0) do
    {
      For node ID i=1 to I<N do
        {
          if (distance(s,i)<d0);    // the neighbor nodes
            if (distance(s,i)<dm)
              {
                dm=d(s,i)   // update the shortest distance
                hop=hop+1;
                N(hop)=i;
                NodeX=N(N(hop))
              }
            endif
          endif
        }
      end for
    s=N(hop);
      // the node on the "hop"th hop to the source node
    cost(S,D)=cost+B_Tx+B_RX;
      // Unicast cost
    }
end while
End
```

4. SIMULATION EVALUATION

4.1 Evaluation metric, objects, and simulation setup

We evaluate our proposed protocols by using NS-2 simulator integrated with IEEE 802.15.4 based sensor node module [11], [12].

The following protocols are evaluated in the simulation:

Flooding: Conventional on-demand data query and collection, which adopts flooding based query and collection of data by forwarding data back toward the sink node [13].

LBM: Location based multicast for query and data collection with reverse paths of multicast [14].

STSP: Spatial-Temporal Sink Proxy based data query and collection with the proposed sink proxy selection scheme.

RANSP: Spatial-Temporal Sink Proxy based data query and collection scheme with RANdom selected sink proxy.

The protocols are evaluated by using the following metrics:

Energy Consumption Distributed at Each Node: The distribution of average energy consumption in a round of data query and collection at each sensor node during the simulation.

Energy Consumption Improvement Ratio: The ratio of performance improvement with regards to energy consumption for STSP comparing with RANSP.

Unless otherwise stated, the simulation is setup as follows. There are 100 sensor nodes, and 1 sink node in the network. The network topology is static and all nodes are set to be connected to each other by either single-hop or multi-hop links. The sensor nodes are distributed over a square of 100m x 100m, and the detail topology of the sensor network is shown in Figure 4. Each sensor node is equipped with a IEEE 802.15.4 radio module, with a radio range of 15m. The maximum number of sensing data that can be aggregated in a packet is set to 20.

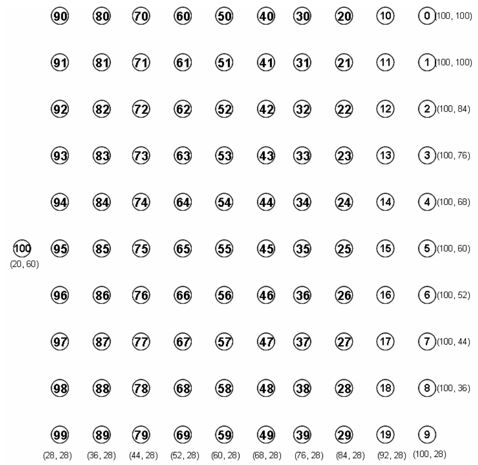


Figure 4 Simulation topology.

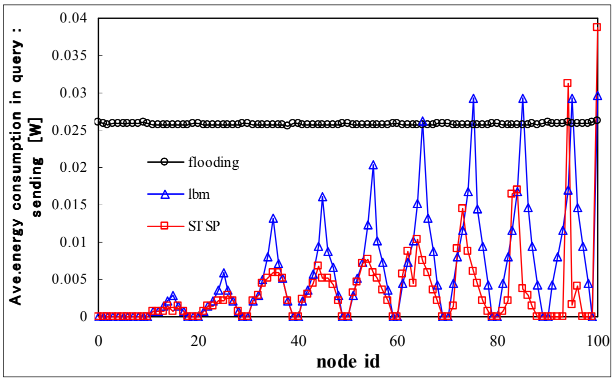
4.2 Numerical results

At first, we evaluate the distribution of average energy consumption at each sensor node in a round of data query and collection where query is dynamically setup with a area size of 20m*20m square. The energy consumption is the average of 40 cycles of data query and collection.

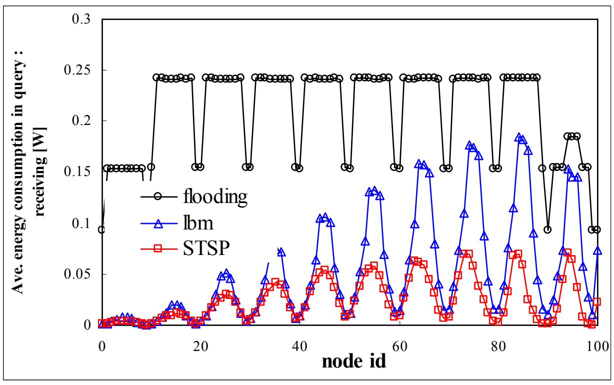
Figure 5 (a) and (b) show the energy consumption of data queries at each sensor node. Flooding-based queries have high energy consumption at most sensor nodes. LBM based query generally has a higher energy consumption than STSP. In the simulation, energy consumption on receiving data packets is much larger than that on sending packet. This is because that each sensor node sends less packets than it receives, since when a node receive a same packet as it has already relayed, it will discard the packet without relaying it again [15].

Figure 5(c) and Figure 5(d) shows the distribution of energy consumption in the data collection operation at each sensor node. Flooding-query based data collection has high and non-uniform energy consumption at sensor nodes. Nodes near the sink node have the highest energy consumption. STSP based data collection has much lower energy consumption at sensor nodes. The highest energy consumption of STSP is much less than that of Flooding-query based data collection and LBM based data collection. Further, STSP has more uniform energy distribution among sensor nodes.

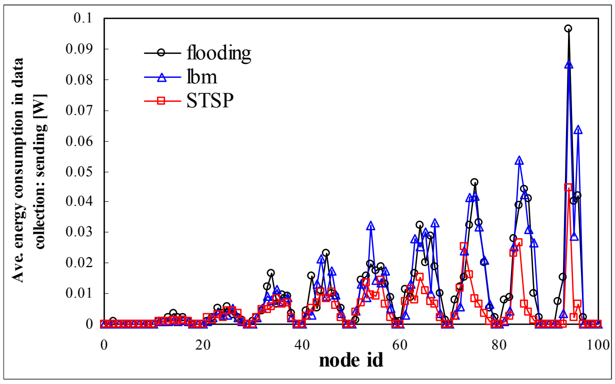
Figure 6 illustrates the energy consumption improvement ratio of STSP to RANSP. Figure 6(a) shows that, STSP saves about 3 and 7 percents of energy for the sending and receiving of query message, comparing with the RANSP. And that STSP saves about 7.6 and 8.1 percents of energy for the sending and receiving of data collection message, comparing with RANSP. Figure 6(b) shows that STSP saves about 1 percents of energy for both sending and receiving of query message, comparing with the RANSP. And that STSP saves more than 20 percents of energy for both sending and receiving of data collection message, comparing with RANSP.



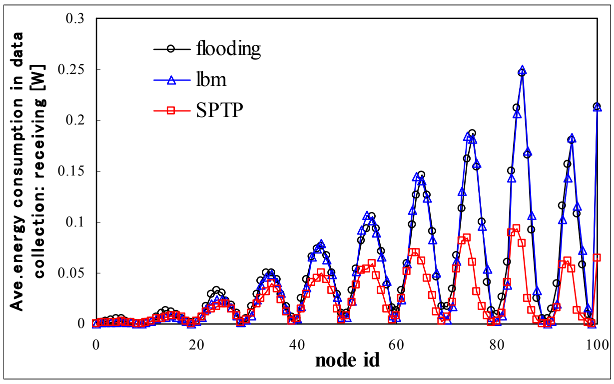
(a)



(b)

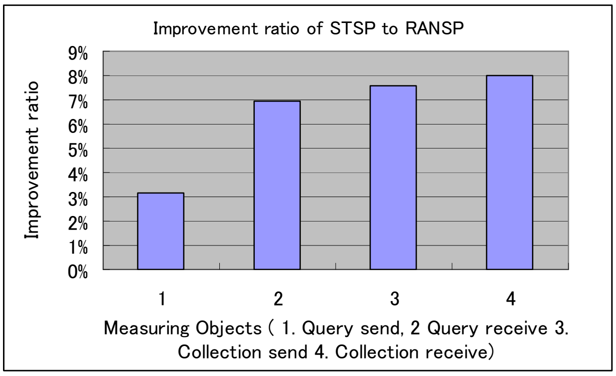


(c)

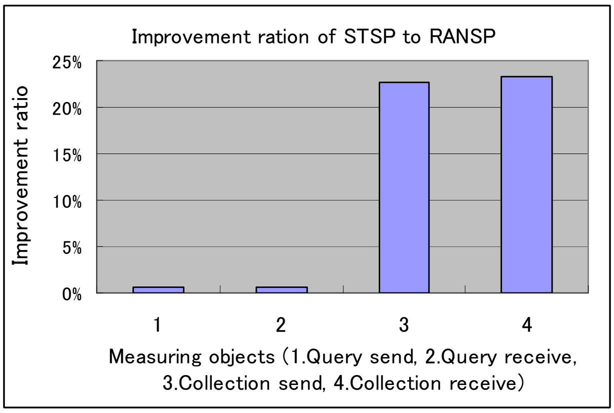


(d)

Figure 5 Energy consumption distributed at each node



(a) 20m*20m size of query area



(b) 40m*40m size of query area

Figure 6 Energy consumption improvement Ratio of STSP to RANSP

5. CONCLUSION

This paper proposed a spatial-temporal sink proxy mechanism, which localizes information retrieval for sensor networks. A sink proxy is selected by the sink node with the awareness of user query's contents. At the sink node, a query's name is resolved into the IDs and locations of corresponding sensor nodes before being distributed to the network. By utilizing the sink proxy according to the location of sensor nodes, query distribution and data collection are performed in a corresponding local area. Data query is delivered to the sink proxy and is locally disseminated. Sensing data are collected at the sink proxy, at which data are aggregated and sent to the sink node. The analysis and simulation results show that the proposed scheme highly reduces the energy consumption of data query and data collection in the network and the energy consumption is more uniformly distributed among sensor nodes since sensor data are collected at sink proxies that are distributed in the network.

6. REFERENCES

[1] D. Estrin, L. Girod, M. Srivastava, "Instrumenting the World with Wireless Sensor Networks" International Conference on Acoustics, Speech, and Signal Processing, May 2001.

[2] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. A survey on sensor networks. IEEE Communications Magazine, Aug. 2002.

[3] A. Mainwaring, J. Polastre, Robert Szewczyk, David Culler, John Anderson. "Wireless Sensor Networks for Habitat Monitoring." ACM International Workshop on Wireless Sensor Networks and Applications, Sept.2002.

[4] C.Intanagonwiwat, R.Govindan, D.Estrin, J. Heidemann, and F.Silva. "Directed Diffusion for Wireless Sensor Networking" ACM/IEEE Transactions on Networking, Feb. 2002.

[5] "Habitat Monitoring on Great Duck Island," http://www.greatduckisland.net/, 2006.

[6] P. B. Karp, Y.Ke, S.Nath, S.Seshan, "IrisNet: An Architecture for a. Worldwide Sensor Web" IEEE Pervasive Computing, Oct.2003.

[7] W. Liang, Y. Liu, "Online Data Gathering for Maximizing Network Lifetime in Sensor Networks", IEEE Transactions on Mobile Computing, Jan 2007.

[8] J. Heidemann, F. Silva, C. Intanagonwiwat, R. Govindan, D. Estrin, D. Ganesan, Building Efficient Wireless Sensor Networks with Low-Level Naming, Proceedings of SOSP, Oct 2001.

[9] Rui Teng; Bing Zhang; Yasuo Tan, "A Study of Localized On-Demand Data Collection in Sensor Networks" 15th IEEE International Conference, Nov.2007.

[10] D. Johnson, Y. Hu, D. Maltz, " The Dynamic Source Routing Protocol (DSR) for Mobile Ad Hoc Networks for IPv4", RFC 4728, Feb. 2007.

[11] Ns2-simulator http://www.isi.edu/nsnam/ns/ns-build.html#allinone

[12] IEEE 802.15.4 modeling and simulation platform http://ees2cy.engr.ccny.cuny.edu/zheng/pub/

[13] C.Intanagonwiwat, R.Govindan, D.Estrin, J. Heidemann, and F.Silva. "Directed Diffusion for Wireless Sensor Networking" ACM/IEEE Transactions on Networking, Feb.2002.

[14] Y. Ko and N.H. Vaidya. Geocasting in mobile ad hoc networks: Location-based multicast algorithms. In Proceedings of the 2nd IEEE Workshop on Mobile Computing Systems and Applications, Feb.1999.

[15] J. N. Al-Karaki, A. E. Kamal "Routing techniques in wireless sensor networks: a survey." IEEE Wireless Communications, Dec. 2004.