A Constrained Multipath Routing Protocol for Wireless Sensor Networks

Peter K. K. Loh and Y. K. Tan

Nanyag Technological University, School of Computer Engineering,
Nanyang Avenue, Singapore 639798
{askkloh@ntu.edu.sg)

Abstract. A deployed wireless sensor network must be managed by an efficient and reliable routing protocol to overcome node degradation and failure, RF communication disruptions and limited energy reserves. RF transmissions are inherently broadcast. Existing research proposals for routing protocol designs typically use either *flooding* (routing same packet over all available paths) or *selective broadcasting* (routing packet over a specified path). Research has shown that routing protocols that use selective broadcasting exhibits better performance with lower communication overheads. Flooding-based protocols typically require lower control overheads and exhibit better fault tolerance. This paper presents two variations of a novel routing protocol, called SOS, which uses *constrained broadcasting* (packet routed over small subset of available paths) to adapt to network failures and disruptions. Simulations show that SOS compares favourably with existing selective broadcasting and flooding-based protocols in terms of performance, reliability and energy efficiency.

Keywords: **r**outing protocol, sensor networks, selective broadcasting, constrained broadcasting, flooding

1 Introduction

A smart environment needs information about its surroundings and its sensory system to operate reliably and efficiently. The importance of a wireless sensor network (WSN) as a sensory system is demonstrated by several recent initiatives [11]. Advances in radio and micro-electro mechanical systems (MEMs) technologies have made sensor nodes more cost effective. Hence, these nodes may be deployed in large numbers in various operating scenarios. Deployment is, however, often under unpredictable and/or inhospitable conditions that may prevent the sensory system from maintaining a stable network configuration to sustain long-term operations [2,8]. To tolerate these conditions, the WSN must be managed by a reliable routing protocol that will maintain its configuration and its operations in the presence of faults [6, 9, 10]. Existing routing protocols typically depend on either *flooding* [13, 15] or *selective broadcasting* [12, 14] to relay data packets. In flooding, each data packet is routed over all available RF links at a node. It is a brute-force approach to increase the chances of packet delivery. In selective broadcasting, only one neighbouring node receives the packet though other neighbours may hear it. The packet is progressively routed along a single path to the hub. It has been demonstrated that selective broadcasting protocols exhibit a performance edge over flooding protocols but the latter can exhibit better fault tolerance in small to moderate-sized networks with lower control overheads [3, 5, 12, 14]. In this paper, we propose a routing protocol that relays data with *constrained broadcasting* (small set of available paths for routing) while meeting the conflicting requirements of low communications and energy costs.

2 Related Work

In this section, we present a concise survey of four routing protocols: the Periodic, Event-Driven and Query-Based protocol (PEQ) [3-4], Gradient-Based Routing protocol (GBR) [14], Efficient and Reliable protocol (EAR) [12], and Gradient Broadcast protocol (GRAB) [13, 15]. These routing protocols are similar in the sense that they make use of path length and/or energy metrics for data dissemination. The first three protocols are based on selective broadcasting while the fourth uses network flooding and serves as a control reference for flooding-type protocols.

The PEQ (Periodic, Event-Driven and Query-Based) routing protocol uses an ACK (acknowledgement)-based scheme to identify node failures or weak signal conditions. To set up route distance information, each network node is initialised with a hop count to the hub. The hub is initialised with a hop count of 0. If a node receives hop information from more than one neighbour, only the lowest value is stored. Each node will only send data to the next node that is of a lower hop value. Hence, a single, shortest path for each node is created to deliver its data to the hub. To tolerate node failures and noisy links, PEQ uses two phases: (i) failure detection at receiver node and (ii) location of a new neighbour. A sender node forwards data to its neighbour and sets a timeout to wait for the ACK. The ACK message will only be received after the neighbour has forwarded the packet. Thus, the sender node on receiving its neighbour's ACK, knows that its neighbour is alive and that the packet has also been forwarded to the next node. However, if no ACK message is being received, the neighbour node is deemed to have failed and the sender node will select another neighbour as its new intermediate destination. To do this, the sender node broadcasts a search message to its neighbours. Neighbours reply with a message containing their hop count and identification. The sender will choose the lowest hop count and the first one to reply becomes the new neighbour node. The sender's routing table will be updated for relaying subsequent packets and the sender node sets its own hop count to be the neighbour node's hop count plus one. This also avoids looping in paths.

GBR distributes traffic evenly among all nodes and prevents non-uniform traffic overloading. The hub will broadcast an interest message that is flooded throughout the network. Each node upon receiving the interest message will record the number of hops taken by the interest message. Each node then knows the number of hops it needs to reach the hub. The *gradient* between two nodes is the difference in their hop counts. A hop gradient is set up from the nodes to the hub and all messages will flow in that direction towards the hub. A node will forward a message to a neighbour with the largest gradient (nearest) to the hub. When there are multiple neighbours with the same gradient to the hub, one is randomly chosen to spread traffic levels more uniformly. When less than 50% of a node's energy level remains, the node lowers its gradient with respect to its neighbours. This reduces the chance of future packets being routed through it. Any changes in gradient will be progressively updated across the network to maintain routing consistency.

Routing decisions in EAR are based on four metrics: hop-count, distance traveled, energy level of next hop node and link transmission success history. A weighted combination of these metrics is used to relay packets. A "sliding-window" keeps track of the last *N* successful transmissions via a specified RF link to compute success history. An optimal route in EAR is not necessarily the shortest but represents the best combination of distance, energy requirements and link quality. Control packets are minimized by "piggy-backing" routing information onto MAC-layer protocol packets. EAR deals well with communication faults in WSNs with low to moderate traffic volumes. However, with a high-volume of network traffic, the more complex route management incurs an appreciable overhead affecting its performance.

GRAB is designed for reliability by routing redundant copies of messages in a mesh from a source node to the hub. A cost field is set up in the network and the value of a node in the field is the minimum cost to reach the hub from that node. The cost field has a value of 0 starting from the hub and the value at each node increases with the distance from the hub. Messages will flow through the cost field in the direction of decreasing cost, that is, towards the hub. When a node generates a packet, it initialises the header with its cost to the hub and assigns a credit value to that packet before broadcasting it. When neighbouring nodes receive the packet, only those nodes that have a lower cost will enter a decision process of whether to route or drop that packet. A message's credit is consumed at each node on the path to the hub. When a message has enough credit, the node will route the message or else the message will be dropped. By assigning an appropriate amount of credit to each message, duplicate copies of that message will travel in multiple paths from the source node to the hub.

3 Protocol Design Details

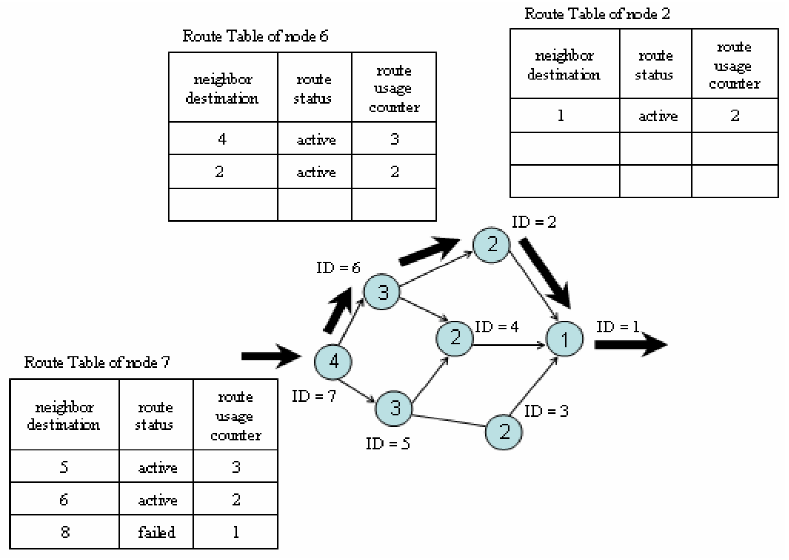
This section details the design of the SOS and $\mathrm{SOS_d}$ routing protocols. $\mathrm{SOS_d}$ is a double path variant of SOS that incorporates limited route redundancy. Both protocols support single or multiple hubs in a multi-hop WSN [16]. SOS and $\mathrm{SOS_d}$ can also be easily modified to support either query- or event-detection based WSNs.

3.1 Route Discovery Phase

This phase determines the WSN's configuration and builds appropriate routing tables at each node. A hop-level distribution from the hub is established in a similar manner to PEQ. A node will use its hop-level to identify the next possible neighbour(s) to relay the data. If the hop level received is lower than the receiving node's hop level by more than one, the receiving node updates its own hop level by adding one to the received hop level and retransmitting the 'hop' message with its own hop level to its neighbour(s). Otherwise, there will no updating and the ‘hop’ message will be discarded. Each node may store up to three possible routes to enhance routing reliability. To minimise the probability that an important node may not be discovered or a link that could have failed due to interference between two neighbouring nodes, each node transmits the ‘hop’ message twice, with each message separated by a random delay to prevent collision of messages. The hub, however, only broadcasts the ‘hop’ message once to avoid energy wastage as in periodic flooding.

3.2 Data Relay Phase

Here, data is routed from each node to the hub and vice versa (Figure 1).



Route Table of node 6

neighbor destination	route status	route usage counter
4	active	3
2	active	2

Route Table of node 2

neighbor destination	route status	route usage counter
1	active	2

Route Table of node 7

neighbor destination	route status	route usage counter
5	active	3
6	active	2
8	failed	1

Fig. 1. Illustration of data packet relaying in SOS

Each node maintains a routing table with up to three different routes to the hub. Each table entry is a 5-tuple: <*hub address*, *neighbour node ID*, *route status*, *times route failed*, *route usage counter*>. *Hub address* is the hub ID in the WSN, *neighbour node ID* is the ID of the intermediate node to which a data packet is to be sent en-route to the hub, *route status* indicates route availability, *times route failed* tracks the number of times the route has failed. The route will be removed from the routing table once a specified maximum number of retries is reached, *route usage counter* tracks the number of times a route has been used. A node will choose the route with the lowest usage. This helps distribute network traffic more uniformly. As a consequence, network lifetime can be increased with probability of node failure due to insufficient energy reduced. Another advantage is that message data packets forwarded along different routes will reduce the chance that a fault in one route causes loss of the entire message. Figure 1 shows node 7 receiving a data packet. In ‘Route Table of node 7’, although the route to node 8 has the smallest route usage of 1, the route status indicates it has failed. The next suitable route is to node 6, an active route with lowest route usage. After relaying the packet to node 6, the route usage value of this route will increase to 3, same as the route to node 5. The route to node 5 will be chosen the next time round as, with the same usage value, the first node in the routing table will be selected. Thus, the data packet will be routed from nodes 7-6-2-1 to the hub.

3.3 Failure Recovery Phase

Node x relays a data packet to neighbour y by first sending RTS. After a specified number of retries, if no CTS is received from node y, node x updates its routing table by deleting node y's ID, decrements its 'number of routes to hub' value and broadcasts node y's ID in an error message to its neighbours. Neighbours receiving this error message remove the problematic node ID from their routing tables. If at least one route to the hub is left, it will be selected as the alternate route to forward the data packet. Presence of extra routes to the hub will increase the chance of delivery. Node x will then broadcast 'search' messages to neighbouring nodes to replace the problematic node. Node x could also be isolated if all its neighbours have failed. After the search, if no replacement is found, the process will terminate. Neighbours of higher hop level now know that node x has no more routes to the hub and will remove it's ID from their routing tables. A node will respond to a 'search' message when it has at least one route to the hub. Reply message with its hop level and ID is sent after a random delay to minimize collisions with other replies.

3.4 SOS Double Path - SOS$_d$

SOS$_d$ exploits *constrained broadcasting* to relay data packets. It utilizes an additional message path to increase the probability of bypassing node/link failures and avoiding packet loss while relaying data packets. Additional paths to forward data packets will have a higher chance of ensuring delivery to the hub. However, the energy usage will be higher especially in larger networks. The limitation to a double instead of multiple paths is to avoid excessive redundancy, which can cause significant increases in packet collisions and network congestion, and incur unacceptable energy overheads. Only source nodes send packets to two different neighbours. However, if two data packets from a source reaches the same node z, node z also relays the packet to two different nodes due to routing control imposed by the route usage counter (Figure 2).

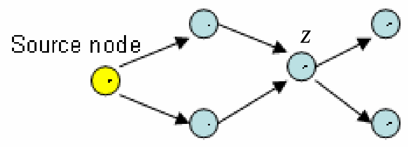


Fig. 2. Illustration of a double path

There are up to three routes to a hub in its routing table.

4 Simulation

UCLA's Global Mobile Simulation Systems Library (GloMoSim) [1] was employed with network nodes modelled after the Xbow MICA2 mote [7]. The MAC protocol was implemented using the distributed coordination function (DCF) of IEEE 802.11. Radio transmission range for each node was 56 meters. Radio model was signal-to-noise (SNR) bounded. Propagation model was ground reflection (two-ray). Frequency used was 433 MHz and the bandwidth used was 76800 kbps. These settings are fixed throughout each simulation experiment conducted. Each simulation was executed for 60 minutes and the size of data payload was 24 bytes.

4.1 Performance Metrics

Performance of each protocol was analyzed using the following performance metrics.

(a) Packet Delivery Ratio (PDR): proportion of packets successfully delivered.

$$\frac{\text{Total number of data packets received}}{\text{Total number of data packets sent}} \quad \text{X} \quad 100\%$$

(b) Packet Latency: average time for hub to receive data packet from source node.

$$\frac{\sum \text{Individual data packet latency}}{\text{Total number of data packets delivered}}$$

It includes processing times at various layers of each node, route recovery and repair, retransmission, queuing, propagation, transfer times etc.

(c) Energy Consumption: energy consumed per delivered packet.

$$\frac{\text{Total energy consumed} = \sum \text{Individual node's expended energy}}{\text{Total number of packets delivered}}$$

Node energy consumption model in GloMoSim is used, where energy used by a node for each packet transmitted at the radio layer/received at the MAC layer is:

Energy expended (J) = Voltage (V) X Current (A) X Transmission/Receive Duration
= 3.0V X 0.0052A X ((packet size X 8) / bandwidth)

Energy consumed will increase proportionally with the number of packets sent. Hence, energy consumption also contributes to the total overhead of a protocol.

(d) Packet Collisions: the total number of packet collisions that occurred during the simulation. It includes collisions by both control and data packets.

4.2 Non-Ideal Conditions

Simulations were carried out for networks of 50 to 500 nodes with 10% and 50% active source nodes. This enables protocol scalability with network size to be evaluated for increasing traffic volume. Active source nodes were randomly chosen and data packet rate at each source node was 1 per 30 seconds. A noise model where every node except the hub takes on a random noise factor between 10% and 50% was used. The noise factor is the probability that a packet received by the node is corrupted or lost in transmission. A fault model where 50% of the nodes, except the source nodes, were randomly selected to fail at a random time within the simulation duration was used. For each network size, results were averaged over 30 runs.

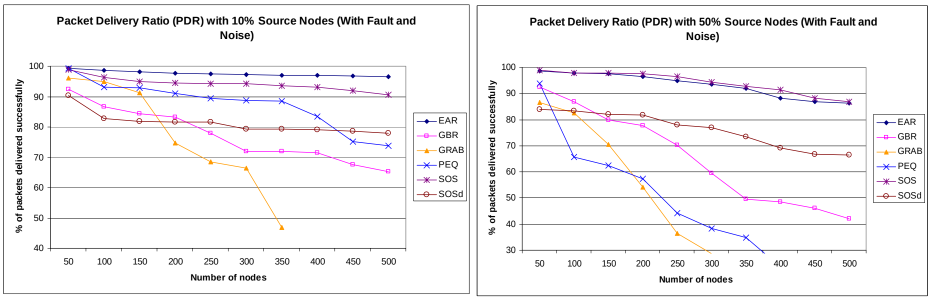


Fig. 3. Packet Delivery Ratio in Non-Ideal Conditions

With 10% source nodes, PDR of PEQ averages above 80% for networks of 400 nodes or less. PEQ is able to maintain a reasonable PDR even at high node failures of 50%. However, PDR drops significantly when active sources increase to 50%. With increased number of active sources, there are more data as well as control packets. PEQ employs an ACK-based scheme that results in an increase in packet collisions - more packets are dropped thus reducing the PDR. A less drastic trend is observed for SOS_d due to the additional path used by each source node. Duplicate data packets transmitted from each source node increases collision rates as in PEQ but performance is increasingly better with more active source nodes as path redundancy in SOS_d is restricted to only source nodes. Figure 3 shows that both SOS and EAR maintain a reasonable PDR above 85%. PDR for GRAB is the lowest for both 10% and 50% active sources. GRAB uses flooding to forward the data to the hub; this generates a significantly large number of data packets in the network and results in over-utilization of bandwidth. Average PDR with 10% is thus better than for 50% active sources where there are considerably more packet collisions and packet loss.

Expectedly, PDR for 50% active sources decreases for all protocols as network size increases. Neighbours within transmission range of a node increases leading to increased collisions. GBR, EAR and SOS protocols make use of a single route for each data packet (selective broadcasting). Reliability of packet delivery is not necessarily directly proportional to the number of multiple paths from source to hub. GRAB and SOS_d achieve fault tolerance by sending duplicate copies of data packets along different routes. However, the PDR of SOS_d for both 10% and 50% source nodes scale better with network size, even exceeding GBR for large networks in excess of 400 nodes. Reconfiguration in GBR does not consider link quality.

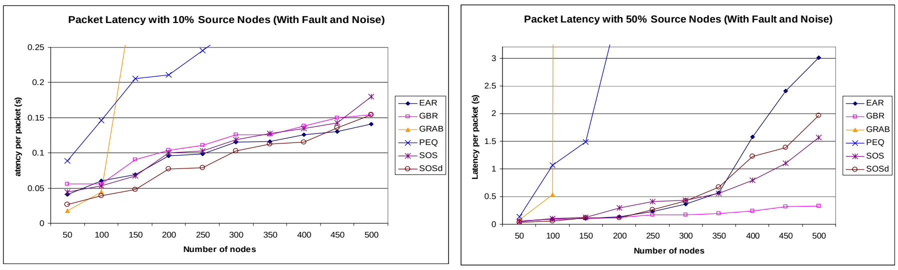


Fig. 4. Packet Latency in Non-Ideal Conditions

GBR maintains relatively low packet latency for 10% active source nodes while it has the lowest packet latency for 50% active sources. In larger networks, GBR's fast and simple random routing has greater chance of locating an active link. However, it does not guarantee that a complete route is available. Also, GBR only updates nodes with network configuration changes when nodes' energy threshold < 50%. Performances of SOS and SOS_d are competitive with EAR. In EAR, the need to continuously process piggybacked control information increases the processing time of data packets resulting in increased latency. SOS informs its neighbours that a change has occurred only when it detects a problem with its intended receiver node. This is more evident with networks in excess of 350 nodes and 50% active nodes.

For networks < 450 nodes and 10% active sources, SOS_d has lower packet latency than SOS and EAR. A double path allows one data packet to reach the hub first while recovery is in progress in the other failed path. Any network reconfiguration process will have a reduced effect on latency. With 50% active sources, EAR's piggybacking advantage is no longer significant with network sizes > 350 nodes. Here, control overhead becomes considerable and compromises latency. With larger networks, SOS reverses its trend against SOS_d as the chance of avoiding failed paths increases with more neighbours available. Now, double transmissions in SOS_d become overheads.

With PEQ and 50% active sources in larger networks, increased transmissions are accompanied by increased acknowledgements. PEQ has to wait for an ACK message timeout to detect a problem before path repair begins. The next data packet has to wait for PEQ to find a new path before it can be relayed. Thus, PEQ's packet latency is higher due to no immediately available alternative routes when transmission fails. GRAB's use of flooding generates large quantities of redundant packets, resulting in unacceptable levels of collisions (also see Figure 6), severely degrading latencies.

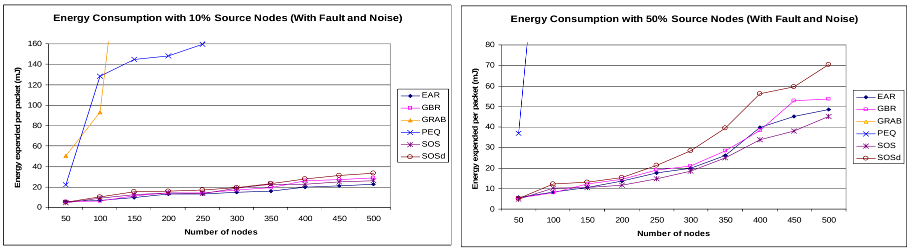


Fig. 5. Energy Consumption in Non-Ideal Conditions

GRAB has the highest energy consumption for both 10% and 50% active sources. This is one of the drawbacks of flooding-based strategies to achieve reliability. During routing, a large amount of redundant data packets are generated (Figure 5).

PEQ's ACKs raise packet transmissions two-fold and reduce its energy efficiency to below that of SOS, SOS_d, EAR and GBR, which are not acknowledgement-based. This is more evident with 50% active sources. Performance of SOS_d may be compromised in a high traffic and noisy environment. Sending duplicate copies of the same packet over a second path consumes 100% extra energy and increases traffic loads in a congested network resulting in more collisions and packet loss.

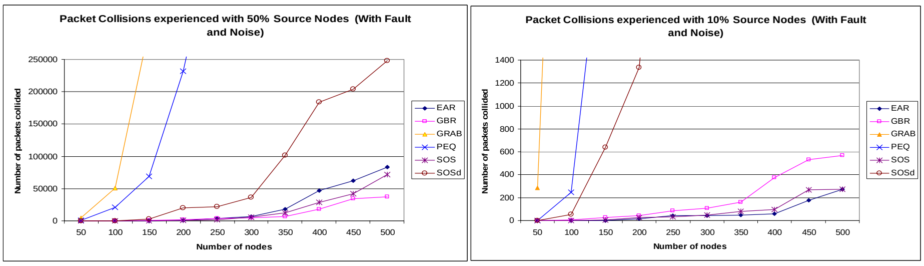


Fig. 6. Packet Collisions in Non-Ideal Conditions

Figure 6 shows that as the number of active sources increase, collisions become a significant performance factor affecting all routing protocols. Extra overhead will be involved in dealing with collisions resulting in less packets delivered successfully. Packet collisions cause unnecessary retransmissions of packets and will impact the latency and energy efficiency of the protocols. PEQ and GRAB, which experience high levels of packet collisions, experience large latencies and high energy consumption figures. With PEQ, increased packet collisions result as a consequence of its ACK-based scheme. With 50% active sources in larger networks, this is more evident allowing SOS to better it.

Both EAR and SOS have lower packet collisions compared to other protocols. EAR uses piggybacked control information to detect node failures. For lower traffic levels, EAR experiences the least number of packet collisions. However, as the traffic levels increase, more failed transmissions are experienced. EAR will send out recovery messages whenever a node detects a problematic neighbour. This results in more control packets exacerbating the collisions. SOS, however, sends out recovery messages only after error detection during transmission. Hence, the situation reverses with SOS experiencing lower packet collisions than EAR with 50% active source nodes. SOS_d experiences higher packet collisions than SOS, EAR and GBR as it exploits an additional path to transmit data packets at each source node. This is moderated to an extent as the path redundancy is restricted to only source nodes.

With 10% active source nodes, GBR experiences more packet collisions than SOS and EAR. The situation reverses with 50% active source nodes. With more active sources transmitting, levels of packet collisions are higher for every protocol. However, this congestion of packet traffic has slightly less effect on GBR due to its stochastic random route selection, enabling it to avoid more collisions with "well-spaced" transmissions. However, a randomly selected path may not lead to successful delivery due to failed nodes as seen in GBR's comparatively low PDR (Figure 3).

5 Conclusions

Several self-reconfiguring routing protocols were evaluated against two variations of SOS. Results show the effectiveness of both protocols' performance. SOS is preferred over SOS_d in WSNs with moderate to heavy data traffic while SOS_d, with constrained broadcasting, performs better in lightly loaded networks. Results showed that SOS and SOS_d are able to achieve competitive levels of PDR, energy efficiency and packet latency, with comparatively less packet collisions in the former. Limitations identified include poor performance in having unrestricted message path redundancy and the use of acknowledgement-based control. Finally, SOS and SOS_d do not require any changes to the MAC layer to achieve comparable performance to EAR. Practical realization will therefore be easier as only the network layer needs to be implemented.

References

1. Ahuja R., Bagrodia R., Bajaj L., Gerla M., Takai M., "GloMoSim: A Scalable Network Simulation Environment", Technical report 990027, UCLA, Computer Science Dept, 1999.
2. Akyildiz I.F., Cayirci E., Sankarasubramaniam Y., Su W., "A Survey on Sensor Networks", IEEE Communications, Pages 102-114, 2002.
3. Bourkerche A., "Performance Evaluation of Routing Protocols for Ad Hoc Wireless Networks", Mobile Networks and Applications, Volume 9, Issue 4, pp 333-342, 2004.
4. Boukerche A., Werner R., Pazzi N., de Araujo R.B., "A fast and reliable protocol for wireless sensor n/ws in critical conditions monitoring apps", pp 157-164, MSWiM 2004.
5. Broch J., Maltz D. A., Johnson D. B., Hu Y-C., Jetcheva J., "A Performance Comparison of Multi-Hop Wireless Ad Hoc Network Routing Protocols" Proceedings of ACM/IEEE International, Conference on Mobile Computing and Networking, pp 85-97, October 1998.
6. Bulusu N., Estrin D., Girod L., Heidemann J., "Scalable Coordination for Wireless Sensor Networks: Self-Configuring Localization Systems", ISCTA, July 2001.
7. CrossBow MICA 2 motes specification http://www.xbow.com
8. Estrin D., Girod L., Pottie G., Srivastava M., "Instrumenting The World With Wireless Sensor Networks", Proceedings of International Conference on Acoustics, Speech and Signal Processing (ICASSP 2001), Salt Lake City, Utah, USA, May 2001.
9. Estrin D., Govindan R., Heidemann J., Kumar S., "Next Century Challenges: Scalable Coordination in Sensor Networks", MobiCOM August 1999.
10. Kahn J. M., Katz R. H., Pister K. S. J., "Next Century Challenges: Mobile Networking for "Smart Dust", Proc 5th ACM/IEEE Intl Conf Mobile Comp & N/wking, pp 271-278, 1999.
11. Lewis F.L., "Wireless Sensor Networks", Smart Environments: Technologies, Protocols, and Applications, John Wiley, New York, 2004.
12. Loh P. K.K., "A Scalable, Efficient and Reliable Routing Protocol for Wireless Sensor Networks", Ubiquitous Intelligence and Computing (UIC), LNCS 4159, 2006, pp 409-418.
13. Lu S., Ye F., Zhang L., Zhong G., "GRAdient Broadcast: A Reliable Data Delivery Protocol for Large Scale Sensor Networks", ACM Wireless Networks, Vol 11, No.2, March 2005.
14. Schurgers C., Srivastava M. B., "Energy Efficient Routing In Wireless Sensor Networks", MILCOM 2001.
15. Ye F., Zhong G., S. Lu, and Zhang L., "GRAdient Broadcast: A robust data delivery protocol for large scale sensor networks," ACM Wireless Networks, Vol. 11, March 2005.
16. Ye W., Heidemann J., Estrin., D., "An Energy-Efficient MAC Protocol for Wireless Sensor Networks", INFOCOM 2002.