This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the ICC 2007 proceedings.

A Node Recovery Scheme for Data Dissemination in Wireless Sensor Networks

F.Bouhafs, M. Merabti, and H. Mokhtar
Computing and Mathematical Sciences School
Liverpool John Moores University
Liverpool, UK
{cmpfbouh, M.Merabti, H.M.Mokhtar}@livjm.ac.uk

Abstract—In wireless sensor networks, sensor nodes are expected to gather specific information about the environment and send it to the user. The data dissemination usually involves many nodes in a multi-hop transmission from the source towards the user base station. Since sensor nodes are energy limited and may fail at any moment, this data delivery is far from secure. Therefore, it is important to design new solution to allow a robust and reliable data dissemination. In this paper we propose a node recovery scheme that exploits the network density and the broadcasting nature of the wireless medium to replace the energy depleted or the failed nodes in the communication. We evaluate this work by simulations and show that our approach improves the communication reliability and extend the routing path time.

Keywords-Wireless sensor networks; data dissemination;

I. INTRODUCTION

Explosive growth in embedded computing and rapid advances in low power wireless networking technologies are fuelling the development of wireless sensor networks (WSNs). These networks can be used in the future in several applications such as healthcare, environmental and habitat monitoring, target tracking, etc.

Wireless sensor networks consist of densely deployed sensor nodes capable of sensing particular physical events in their vicinity and communicating between themselves using wireless transceivers. These tiny sensor nodes are usually deployed in various environments to collectively gather some data required by the user and deliver them to a central monitoring unit called the sink or the base station (BS).

Such sensing delivery operation may include several hops among the resource-limited sensor nodes. If a sensor node on the data dissemination path runs out of energy or fails, the whole delivery operation will be comprised. Therefore, the challenge is to design a solution that makes the data dissemination more reliable and extends the routing path time as much as possible.

In this paper we propose a node recovery scheme for data dissemination in wireless sensor networks. By exploiting the network density and the broadcasting nature of the wireless medium, we propose to replace the energy drained or failed nodes by other neighbouring nodes that can relay the data from the source to the destination, without changing the routing path.

The remainder of the paper is structured as follows. Section II reviews the major routing techniques for WSNs found in the literature, and outlines theirs drawbacks. Section III introduces our new node recovery scheme and describes its different phases. Section IV studies the impact of the WSNs parameters such as the network density and the sensor nodes radio range on the recovery scheme. In section V we evaluate our protocol by simulations. Section VI, outlines open research issues and future work

II. BACKGROUND

Due to several sensor networks constraints, designing efficient communication architecture brings many new challenges. The problem of routing in wireless sensor networks has been the subject of intense study in the last few years.

Several algorithms and protocols have been developed in the last few years, with the goal of achieving more efficient and reliable data dissemination in WSNs. Flooding is obviously the most reliable dissemination scheme to deliver data from source to the sink, as it does not need any extra cost for topology maintenance or route discovery. However, this technique suffers form significant redundancy with too many duplicated messages in the network, which make it an energy costly solution.

Several energy efficient routing protocols have been proposed [1]. Most of them use a single path approach to transmit the data to the user. In single path routing the optimal path is selected according to a predefined metric such as the gradient of information, the distance of the destination or the node residual energy level [2, 3]. Although, single path approach achieves shortest delay and involves the minimum number of nodes, it concentrates the traffic on the same path. In case of continuous data transmission such approach may result in energy exhausted nodes and the loss of the network connectivity.

On the other hand, some other routing protocols that use multipath dissemination choose the network reliability as their design priority. In this approach the data delivery relies mostly on the optimal path. The alternative paths are used only when some nodes on the primary path fail. In [4] and [5] , a multipath extension of Dynamic Source Routing (DSR) and Ad hoc On-demand Source Routing (AODV) were proposed to improve the energy efficiency. In directed diffusion [6] the flooding of interest by the sinks allows the gradients to be set within the network. In [7] a multipath routing approach is proposed for directed diffusion to improve resilience to nodes failure, by exploring the possibility of finding alternate paths connecting the source and sinks when node failures occur. In [8] a probabilistic routing protocol is proposed which use a retransmission probability function to reduce the number of copies of same data. This probability function use the hop distance to the destination and the number of steps that the data packets have travelled as parameters. In [9] the multipath routing is formulated as linear programming problem with an objective to maximize the time until the first sensor node runs out of energy. The sources are assumed to be transmitting data at a constant rate. Although multipath approach achieves reliable data dissemination, using several paths and frequently changes of routing paths results in important energy consumption especially if the source is far away from the sink.

III. A NODE RECOVERY SCHEME FOR DATA DISSEMINATION IN WSN

As sensor nodes are usually densely deployed and due to the broadcasting nature of wireless channel, it is possible that nodes overhear another node's transmission if they are within its transmission scope. Although this redundant reception might result in further energy consumption; it can be useful to recover nodes failure in the routing process, as it is illustrated in fig.1.

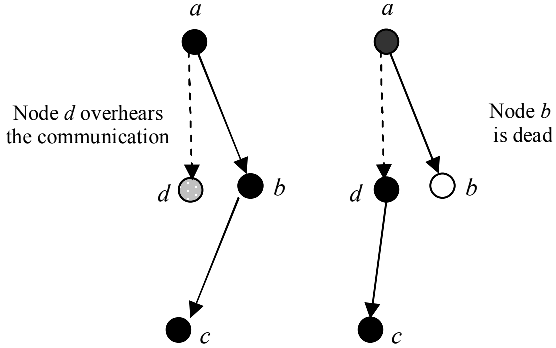

Figure 1: Node *d* could be the recovery node for *b*

However, such recovery needs to be organized in order to allow each node on the data dissemination path to discover its potential recovery node if it exists.

Our recovery scheme can work with any gradient routing protocol like directed diffusion or our semantic clustering protocol described in [10]. Using any of these routing protocols we assume that a source node *S* is sending data to the sink following an already defined path.

We introduce the two entities involved in our recovery scheme: the *upstream neighbour entry*, and the *downstream neighbour entry*. The upstream neighbour entry is the memory space used by each node to save the id of the neighbour from where it receives data packet to send it towards the sink. The downstream neighbour entry is the memory space used by each node to save the id of the neighbour from which it receives the interest and the corresponding number of hops to the sink. In other words the upstream neighbour represents the previous in the routing path and the downstream neighbour represents the next node in the routing path.
Our recovery scheme consists of three phases:

- Neighbour discovery phase.
- Recovery nodes discovery phase.
- Failure node recovery phase.

A. Neighbours Discovery

This operation starts at the deployment time and aims to inform each sensor node about its neighbours. Upon deployed, each sensor node broadcasts a *neighbours_discovery* message, and each node that receives this message replies to the sender by sending a *neighbour_reply* message. At the end of this operation each node will have a complete list of its neighbours.

B. Recovery Nodes Discovery

For this phase we assume that the interest propagation has been successfully finished and a routing path has been established, and each in the path save the previous node id and the next node id in the upstream neighbour and downstream neighbour entries.

When the routing path is defined each node on the path starts looking for a recovery node by broadcasting *recovery_node_discovery* to its neighbours containing: its id, the *upstream neighbour* id, and the *downstream neighbour* id. If a node *m* receives a *recovery_node_discovery* message from a node *n* on the routing path, it starts looking for the *upstream neighbour* and *downstream neighbour* in its neighbours list. If the two nodes' ids are found, which means that it has both the nodes as neighbours, the node *m* declares it self as a potential recovery node for node *n*. The node *m* sets then a timer s for a random but short period of time T_{wait}, after which it sends a *recovery_node_reply* message to the node *n*, and the node *m* is considered as the recovery node for the node *n*. If within the period T_{wait}, the node *m* hears another *recovery_node_reply* message towards the node *n* it cancels the transmission. This mechanism helps to reduce the number of the recovery_*node_reply* transmissions. This operation will be repeated for all nodes in the routing path from the source to the sink, so that each node on the path will have a known recovery node.

Note that all recovery nodes once known turn their radio systems to the overhearing mode.

C. Failure Node Recovery

Suppose that the node *n* is a sensor node on the routing path and *m* is its recovery node. When the node *n* receives a data packet to forward it to the sink, the node *m* overhears the transmission and set a timer T_{transmit} and waits for the node *n* to transmit the data packet to its downstream neighbour. If the timer T_{transmit} expires and the node *m* still did not hear the transmission from *n* to its downstream neighbour, it considers the node *n* as dead, and forwards the data packet instead.

Node *m* informs both the upstream neighbour and downstream neighbour of node *n* that it is the new node on the routing path. Consequently all three nodes start another recovery node discovery sequence as already described in the section B.

IV. NODES DENSITY AND TRANSMISSIONS OVERHEARING

In this section we will discuss how successful is our scheme, the probability of finding a recovery node, and the factors that affect this probability.

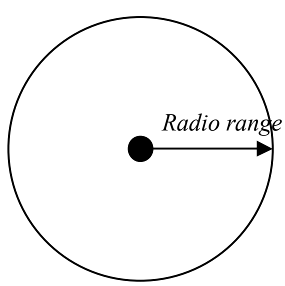

Figure 2: Radio coverage model

Since sensor nodes are equipped with wireless radio transceivers, their scope of transmission is limited. As the propagated signal strength decays exponentially with respect to distance [11], the radio coverage area can be simply modelled as a disk where the transceiver is at the centre. The diameter of this disk is considered as the radio range of the transceiver, as it is illustrated in fig.2.

Fig.3 illustrates an example of a one hop transmission between node *a* and node *c* through node *b*. To make this communication possible it is necessary to have node *b* within the radio coverage area of both nodes *a* and *c*. Consequently, increasing the number of nodes within this area increases the chances of recovery and maintaining the network connectivity much longer.

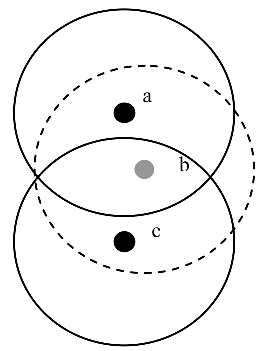

Figure 3: Wireless transmission model

As sensor nodes are usually randomly and densely deployed it is important to study what factors can help to increase the chances of nodes recovery and maintain network connectivity.

We consider a network consisting of a large number of sensor nodes deployed randomly in a two-dimensional geographical region. Under this assumption, the sensor locations can be modelled by a stationary two-dimensional Poisson point process [12, 13]. Denote the density of the underlying point process as λ, and it represents the number of nodes per m^2. The number of sensor nodes located in a region *R*, *N(R)*, follows a Poisson distribution of parameter $\lambda\|R\|$, where $\|R\|$ represents the area of the region, as following:

$$P(N(R)=k)=\frac{e^{-\lambda\|R\|}(\lambda\|R\|)^k}{k!}$$

Fig.4 shows the maximum area of interconnection of two nodes radio coverage areas without possibility to communicate with each other, denoted R_{Max}. This area can be calculated using only the radio range *r* as following:

$$R_{Max}=Ct(r)=\frac{r^2(4\pi-3\sqrt{3})}{6}$$

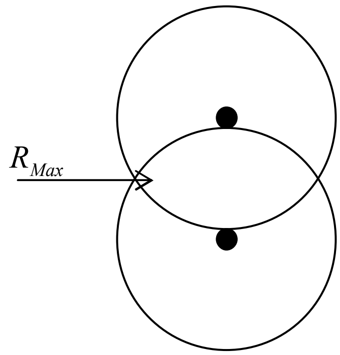

Figure 4: Maximum area of interconnection without direct communication possibility

Consequently, knowing the area R_{Max} and by using the equation (1), it is possible calculate the probability to have at least one recovery node for each node on the transmission path, by calculating the probability of having at least two nodes within the area R_{Max} as following:

$$P(N(R_{Max}) \geq 2) = 1 - P(N(R_{Max}) \leq 2)$$

Where:

$$P(N(R_{Max}) \leq k) = \sum_{i=0}^{k} \frac{e^{-\lambda \|R_{Max}\|} (\lambda \|R_{Max}\|)^i}{i!}$$

Fig.5 shows the probability of having at least two nodes within the area R_{Max}, for different network densities (number of nodes/m^2) and different radio ranges. From this figure we can observe that the probability of having at least two nodes within the area R_{Max} increases as the network density and the radio range increase.

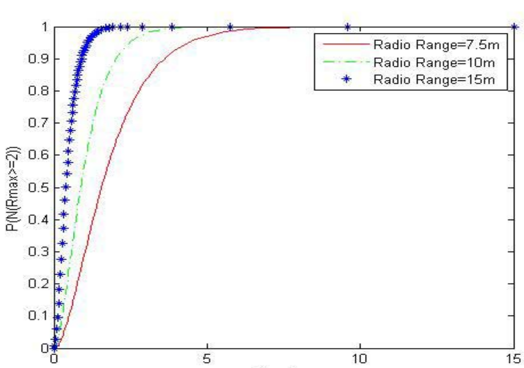

Figure 5: probability of having at least two nodes within the area R_{Max}

V. EVALUATION

To evaluate our protocol we have used the Georgia-Tech Network Simulator (GTNetS)[14] and we used the same radio model discussed in [1]. We firstly simulate a WSN in a 100mx100m area, where sensor nodes are densely deployed following a Poisson distribution, and where the base station (BS) is at position (x=0, y=0). We simulate source node in the middle of the area at location (x=50, y=50), and we establish a path between the source node and the BS, using a simple gradient scheme.

Fig.6 shows the number of potential recovery nodes using the recovery node discovery scheme for each node within the dissemination path, for different network densities and different radio ranges. From this figure we can observe that the number of recovery nodes as the network density increases as the network density and the radio range increase. These results correspond to the analytical results obtained in section IV.

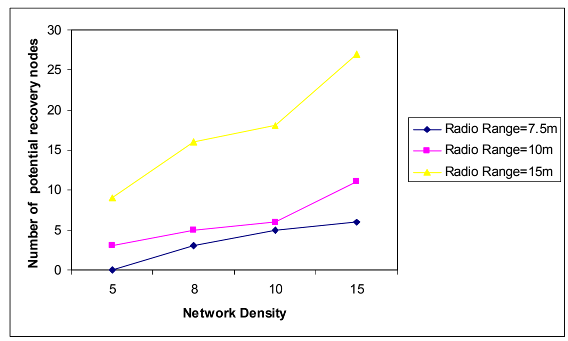

Figure 6: Number of discovered recovery nodes

At a second stage, we fix the network density $\lambda = 10$ (number of nodes per m^2) and the radio range to 7.5m and we evaluate the performances of our recovery scheme on the single path routing operation.

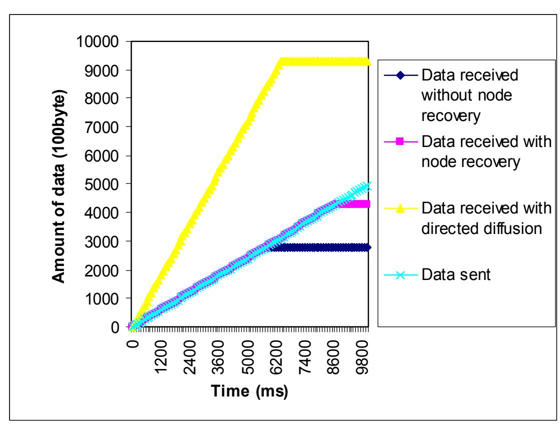

Figure 7: Amount of data sent and received over time

Fig.7 shows the amount of data sent by the source and the amount of data received by the user base station (BS) over time, with node recovery scheme, without node recovery scheme, and using directed diffusion. Form fig.7 we observe that that the single path approach associated with our recovery scheme extends the routing path lifetime and out perform directed diffusion. This figure shows also that directed diffusion deliver more data than the amount sent by the source. This is because directed diffusion sends the data to the base station through several paths to ensure reliability, which consume more energy than single path routing.

Fig.8 shows the number of nodes alive over time, with and without node recovery scheme. From fig.8 we can observe that more nodes are used in the communication with the node recovery scheme than without it. The additional dead nodes in the single path routing with recovery are the recovery nodes used to extend the communication time.

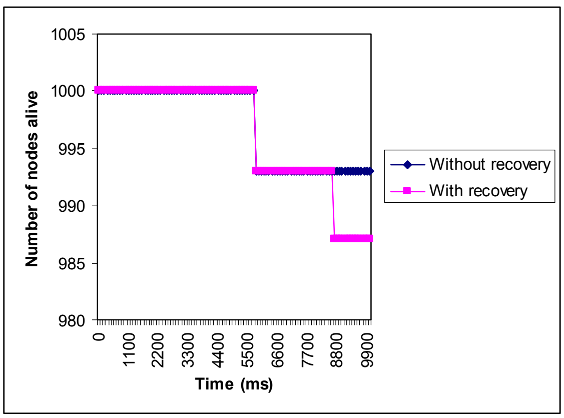

Figure 8: The number of nodes alive over time with and without recovery nodes

We perform the same comparison between single path routing with recovery and directed diffusion as shown in fig.9.

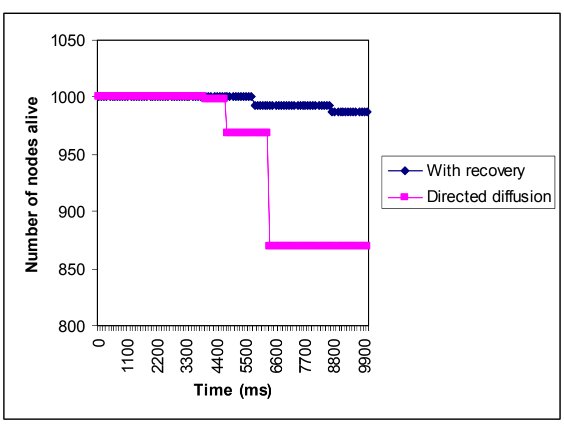

Figure 9: A comparison of the number of nodes alive between directed diffusion and our recovery scheme

This figure shows that our recovery scheme associated with a single routing path solution exploit less nodes than directed diffusions and thus provides better network connectivity and more reliable data dissemination.

VI. Conclusion

Reliable data dissemination is a major problem in WSN. The approaches proposed so far to solve this problem are either not too reliable or consume too much energy. As WSN are usually densely and randomly deployed it is possible to exploit this property to recover energy drained nodes in the routing path. In this paper we presented a new node recovery scheme which exploits the WSN density and the broadcasting nature of the wireless medium to recover energy exhausted or failed nodes that are involved in the communication with the sink. We show through simulations and mathematical analysis that it is possible to find at least one recovery node for each node in the network, if some conditions related to the network density and the radio range are satisfied. We show also through simulations that when our scheme is used in conjunction with a single path routing protocol it results in an extension of the routing path and better network connectivity. We show also that our scheme is more reliable than the single path routing and achieves much more energy saving than multipath routing. Future work will consider the extension this work to hostile environments where sensor nodes are more exposed to failure.

References

[1] K. Akkaya and M. Younis, "*A Survey of Routing Protocols in Wireless Sensor Networks"*, elsevier Ad Hoc Network Journal, 2005, **3**(3): p. 325-349.

[2] H.-J. Huang, et al. "*Magnetic Diffusion: Disseminating Mission-Critical Data for Dynamic Sensor Networks"*, in *International Workshop on Modeling Analysis and Simulation of Wireless and Mobile Systems,* Montréal, Quebec, Canada, October 10-13, 2005.

[3] Y. Xu, J. Heidemann, and D. Estrin. "*Geography-informed Energy Conservation for Ad Hoc Routing"*, in *ACM/IEEE International Conference on Mobile Computing and Networking,* Rome, Italy, July, 2001.

[4] A. Nasipuri and S. R. Das. "*On-Demand Multipath Routing for Mobile Ad Hoc Networks "*, in *International Conference on Computer Communications and Networks,* Boston, MA, USA, October 1999.

[5] M. K. Marina and S. R. Das. "*On-demand Multipath Distance Vector Routing in Ad Hoc Networks"*, in *International Confrence for Network Protocols,* Riverside, CA, USA, November 2001.

[6] C. Intanagonwiwat, R. Govindan, and D. Estrin. "*Directed Diffusion: A Scalable and Robust Communication Paradigm for Sensor Networks"*, in *MOBICOM 2000,* Boston, MA, USA, August 2000.

[7] D. Ganesan, et al., "*Highly Resilient, Energy Efficient Multipath Routing in Wireless Sensor Networks"*, Mobile Computing and Communications Review, 2002, **1**(2): p. 295-298.

[8] C. Barrett, et al. "*Parametric probabilistic sensor network routing"*, in *International Workshop on Wireless Sensor Networks and Applications,* San Diego, CA, USA, September 2003.

[9] J.-H. Chang and L. Tassiulas. "*Energy Conserving Routing in Wireless Ad-hoc Networks"*, in *IEEE INFOCOM 2000,* Tel-Aviv, Israel, March 2000.

[10] F. Bouhafs, M. Merabti, and H. Mokhtar. "*A Semantic Clustering Routing Protocol for Wireless Sensor Networks"*, in *IEEE Consumer Communications and Networking Conference,* Las Vegas, NV, USA, 8-10 January, 2006.

[11] P. M. Shankar, "*Introduction To Wireless Systems"*. 2002, New York: John Wiley & Sons, Inc. 317.

[12] V. P. Mhatre, et al., "*A Minimum Cost Heterogenous Sensor Network with a Liftime Constraint"*, IEEE Transactions on Mobile Computing, 2005, **4**(1): p. 4-14.

[13] J. L. Devore, "*Probability and Statistics for Engineering and the Sciences*". 2004, Belomont, CA, USA: International Student Edition. 785.

[14] G. Riley. "*The Georgia Tech Network Simulator*", in *ACM SIGCOMM Workshop on Models, Methods and Tools for Reproducible Network Research,* Karlsruhe, Germany, August 2003