A Cluster-Based Quantitative Reliability Model

Eduardo Cañete, Manuel Díaz, Luis Llopis and Bartolomé Rubio
Dpto. Lenguajes y Ciencias de la Computación
University of Málaga
(ecc,mdr,luisll,tolo)@lcc.uma.es

***Abstract*—It is well known that Wireless Sensor and Actor Networks are error-prone as multi-hop communications are carried out. Furthermore, the further the distance between two nodes is, the less the communication reliability is. Despite the fact that this issue has been studied in many publications, there are new publications still appearing due to the importance of this topic. In this paper, we present a tool to help developers to better understand how the distance and the link qualities estimation affect the communication reliability between two nodes. We also present a reliability model to improve the reliability between nodes taking into account their energy consumption. The main feature of our proposal is that developers will be able to specify the desired reliability level quantitatively. Finally, a set of tests are carried out in order to study the performance of the proposed model.**

***Index Terms*—Wireless Sensor and Actor Networks, Reliability, Model, Clusters**

I. Introduction

Wireless Sensor and Actor Networks (WSANs) [1] are a promising technology which allows the monitoring and control of any kind of scenario (indoor environment, whole cities, woods, . . .) [2]. These networks are composed of tiny devices which are resource-constraints enough due to their small size. They are normally characterized by their short-range wireless communications capabilities, short battery-lifes, few memory and limited CPU processing capabilities. In spite of these limitations making WSAN applications difficult to develop because there exists another problem which is even worse. Within a WSAN, the delivery data between nodes (sensor and/or actors) which are N hops away from each other fail quite a lot due to the fact that WSANs are error-prone [3]. And obviously, the probability of fails increases if the distance between the source and destination also increases. Developers should take this into account this issue when they plan to develop and deploy a WSAN, otherwise the network probably does not achieve the goal for which it was thought. Thus, we can conclude that a very important issue in WSANs is to define efficient reliable multi-hop protocols in order to achieve either a high packet delivery probability (PDP) or a high packet reception ratio (PRR) [4].

In this paper, we present a tool developed to study how the delivery data between two nodes is affected as the distance (number of intermediate nodes) between them increases. But, the main contribution of this paper is a reliability model which allows the developers to numerically (0 to 100%) set the desired reliability level between two nodes which are N hops away from each other. Basically, the algorithm is able to know and achieve the needed reliability of the intermediate nodes used to send information between nodes with a reliability previously specified by the user. This way of defining the desired reliability allows us to establish a more direct relationship between PRR and the application layer of our goal application. So, if we want to develop a WSAN application which is able to detect dangerous situations (for example, a high level of radiation), the sensor networks have to be capable of transmitting this kind of information with a reliability close to 100%. Other approaches allow us to set parameters such as high-reliability, medium-reliability, but what exactly is the meaning of these parameters? In other words, what exactly is the reliability reached by using the high-reliability or medium-reliability parameter? What about if we would like to establish a reliability lower than high-reliability and higher than medium-reliability? We think this way of defining the reliability levels among nodes confuses the developers.

The rest of the paper is structured as follows. Section II summarizes the related work. Section III describes the communication model on which the proposed reliability model is based. Section IV presents the proposed quantitative reliability model. Section V describes the reliability tool developed. Section VI discusses the performance evaluation of the proposed model. Finally, section VII concludes the paper.

II. Related Work

There are different kinds of approaches focused on achieving reliable mechanisms to transport the data. Many of these approaches are designed for sensor networks where any scheme to organize the nodes is followed. In [5] is followed a reliability scheme based on the priority queues and work load of the nodes which allows the nodes to estimate the feasibility of delivering a packet on time. However, they impose important restrictions in the test scenario such as each node is assumed to know its position and only sensor-actor interaction is studied. Other reliable approach is established in [6]. In this work, packets are managed depending on the importance of their content, however the authors do not provide any algorithm to obtain the packet importance. In this approach, the packets are transmitted through different paths in order to increase the possibilities of data reception at destination. However, they assume there are not collisions and that packets are not cached in sensor nodes because of memory constraints. This last assumption can lead to a considerable increase of the overhead. In our approach, the reliable communication is reduced to the clusters which make the protocol not only more energy-efficient but also robustness. In addition, our approach takes into account the existence of possible collisions and, on the other hand, the packets are registered in order to avoid an increase of the traffic overhead. It is worth pointing out, most of the current reliable protocols are designed to find the most reliable paths taking into account the energy consumption [7], [8], [9]. Of course, these kinds of approaches achieve good reliability levels and a good trade-off between reliability and energy consumption during the data transmissions but they are not able of quantifying the level of reliability that can be achieved as they are normally not based on a mathematical model. Furthermore, most of them are only designed to achieve reliable paths from the sensor nodes to the sink, cluster-head or base station, but not in an opposite way. In [10], authors propose a reliability model slightly similar to our model. They also allow developers to specify the desired reliability level between two nodes. To achieve that, they exploit the inherent redundancy of dense sensor networks by realizing probabilistic multi-path forwarding. In addition, they assume that all nodes have the same link qualities and the nodes know where are located geographically by using GPS coordinates. In contrast, our protocol is able to achieve the same goal without having to assume the constraints above mention (dense sensor network, same link qualities and geographical position).

III. Communication Model

Most sensor networks are designed to transmit information from sensor nodes to one powerful node called sink, base station or cluster-head as the topology of the network. For example, we could have a network organized in several clusters within which all collected information is sent to the cluster-head which form another cluster whose cluster-head is the sink or the base station. The proposed reliability model thought, is to be implemented in networks which follow the aforementioned topology taking into account the following assumptions:

1) The communication pattern is many-one and one-many. There are groups of sensor nodes which transmit information to their leader node (cluster-head, sink, or base station). Leader nodes can also send information to the members of their groups. Although this kind of communication is less frequent, but not any less important.
2) Nodes are organized in levels (distance in hops to their leader node). So, each node knows what its level is regarding its leader node. This is known as gradient-based routing.
3) Nodes situated in level L also know their neighbors located in levels $L-1$, $L+1$ and L. Thanks to this information sensor nodes will know what the shortest paths are to reach their leader node as well as knowing how many hops there are to it. On the other hand, leader nodes have to send the information by using broadcast as they do not know where the member nodes are located. However, they know in which level packet retransmission has to be interrupted.
4) Nodes know the link quality estimation between themselves and their neighbors.

IV. Reliable Transmission Model

As mentioned in previous section, it is assumed the nodes within a same cluster know in which level they are located and which are the different and shortest paths to send information to their cluster-head. To know the node level is equivalent to knowing the number of hops between this node and its cluster-head. Let us suppose that we have deployed the following lineal sensor network: 1-2-3-4-5 where 1 is the cluster-head, 5 is a sensor node and 2,3 and 4 are the intermediate nodes. The PDP of sending a packet from node 5 to the node 1 and viceversa comes defined by the product of the intermediates PDP as the following expressions show:

$$PDP_{51} = PDP_{54} * PDP_{43} * PDP_{32} * PDP_{21}$$
$$PDP_{15} = PDP_{12} * PDP_{23} * PDP_{34} * PDP_{45}$$

To make the discussion easier, PG_{ij} will be the PDP between nodes i and j when they are not neighboring and PC_{ij} when they are. Therefore, expressions above can also be expressed in the following way:

$$PG_{51} = PC_{54} * PC_{43} * PC_{32} * PC_{21} \tag{1}$$
$$PG_{15} = PC_{12} * PC_{23} * PC_{34} * PC_{45} \tag{2}$$

It is noteworthy that $PG_{51} \neq PG_{15}$ due to RSSI asymmetry. While RF theory states that the two directions of RF propagation have identical attenuation, in practice this is not the case [11].

These equations 1 and 2 show that to achieve a specific reliability (for example about 90%) during the transmission of packets from node 5 to node 1 (PG_{51}) it is also necessary to know a priori what the reliability is of the intermediate communications (PC_{54}, PC_{43}, PC_{32}, PC_{21}) which is quite hard due to the fact that the quality of each link changes in a dynamic and independent way over time. It implies we have to deal with an equation of $X-1$ variables, where X is the number of nodes that participate in the communication process. Therefore, our first goal is to achieve that the equation used to calculate PG_{51} has only one variable. If all PC_{ij} were equals, we would have just one variable and:

$$PG_{ij} = PC^L$$

where L is the level of the source node i. From this equation we can find the value of PC in the following way:

$$PG_{ij} = PC^L$$
$$\sqrt[L]{PG_{ij}} = \sqrt[L]{PC^L}$$
$$\boxed{PC = {PG_{ij}}^{\frac{1}{L}}} \tag{3}$$

Continuing with our own example, equation 3 means that if we want to establish a reliability level of PG_{51}, the intermediate PC_{54}, PC_{43}, PC_{32} and PC_{21} must be equals to ${PG_{51}}^{\frac{1}{4}}$.

At this point, we know what must be the needed reliability level (PC) during the communication of the intermediate nodes to reach a specific reliability (PG_{ij}) between two nodes (i and j) which communication distance is L hops. Once, we know this information, the next step is to find the way of increasing PC_{ij} to $PG_{ij}{}^{\frac{1}{L}}$. It is obvious that if several retransmission are carried out from node i to node j the reliability between them will increase, but how many retransmissions are necessary to increase this reliability level from PC_{ij} to $PG_{ij}{}^{\frac{1}{L}}$? The following expression gives us the solution to this question:

$$
\begin{aligned}
1-(1-PC_{ij})^{N_{ij}} &= PG_{ij}{}^{\frac{1}{L}} \\
-(1-PC_{ij})^{N_{ij}} &= (PG_{ij}{}^{\frac{1}{L}})-1 \\
(1-PC_{ij})^{N_{ij}} &= 1-(PG_{ij}{}^{\frac{1}{L}}) \\
\ln(1-PC_{ij})^{N_{ij}} &= \ln(1-PG_{ij}{}^{\frac{1}{L}}) \\
N_{ij} * \ln(1-PC_{ij}) &= \ln(1-PG_{ij}{}^{\frac{1}{L}}) \\
\boxed{N_{ij} = \left\lceil \frac{\ln(1-PG_{ij}{}^{\frac{1}{L}})}{\ln(1-PC_{ij})} \right\rceil} \qquad (4)
\end{aligned}
$$

In the equation, $(1-PC_{ij})^{N_{ij}}$ is the probability that node j does not receive a packet from node i after it is sent N times. Thus, $1-(1-PC_{ij})^{N_{ij}}$ is the probability that at least one packet sent by the node i arrives to the node j.

Basically, the communication reliable protocol is based on equations 3 and 4. For example, let us suppose that a developer has to create an application where nodes must send an alarm packet to the sink when they detect a high temperature (over a given threshold). If the furthest distance from them to the sink is four hops and the desired reliability is around 92%, equation 3 shows us that the link reliability between intermediate nodes must be $0.92^{\frac{1}{4}}$ which is 0.979 (about a 98%). Now, let us assume that the values of PC_{54}, PC_{43}, PC_{32}, PC_{21} are 78%, 85%, 88% and 81% respectively. Then, according to equation 4, node 5 needs to transmit the same packet to node 4 at least 3 times which comes from $\left\lceil \frac{\ln(1-0.98)}{\ln(1-0.78)} \right\rceil$. N_{43}, N_{32} and N_{21} would be equal to 2, 3 and 3 respectively.

This protocol depends heavily on the estimation of the current reliability between neighboring nodes. Thus, the more accurate the link quality estimation between neighboring nodes is, the better the achieved reliability between nodes which are far away from each others N hops is.

V. Reliability Tool

In order to help the developers understand what is going to be the impact of their established reliability levels. We have developed a tool (see figure 1) to help them analyze how the different reliability levels affect to the sensor networks depending on the number of levels established within a cluster and the desired reliability level in a multi-hop communication. The tool graphic interface can be classified in 4 parts:

1) Multi-hop communication parameters. This part is located in the top left corner of the interface and has the following elements:
   - A numeric field which allows us to introduce the number of levels of the cluster where a reliable communication is going to be carried out. It also means, the maximum number of hops needed to send information from the sensor nodes to its cluster-head or viceversa.
   - A numeric field where the desired level of reliability is indicated.
   - A label which indicates to us the needed reliability during the communication of the intermediate nodes to achieve the global reliability specified.
2) Information Zone. It is located in the top middle of the interface. It just shows us the information mentioned above in a graphical way.
3) Single-hop communication parameters. It is located in the top right corner of the screen and it only has two fields:
   - A numeric field which simulate the possible reliability current levels between two nodes.
   - A numeric field which indicates to us the needed retransmissions number to achieve the reliability level calculated and is showed in the top left corner of the screen knowing that the current reliability between two nodes is the value established in the above field.
4) Graphic information. It is formed by four graphics through which it is easier to analyze how the application will perform depending on the established parameters.
   a) Retransmissions Vs. Current Reliability. This graph is obtained from equation 4. It shows us how many retransmissions are needed to achieve the desired global reliability depending on the possible current reliabilities between neighboring nodes. For example, if we want to achieve a global reliability about 80% between two nodes which are 8 hops away from each other, equation 3 indicates that the required reliability per link is a 98%. The graphic shows us the number of retransmissions each node would need to achieve a global reliability of 80% on the basis of their current reliabilities. If we want to study a concrete data, the graph has a horizontal red line that shows us this kind of information. This line can be moved by modifying the data refereed to the single hop communication (top right corner of screen). In our own example, the red line shows us that if a reliability between two nodes is about 20%, 17 retransmissions are needed to achieve a reliability level of 98% which is necessary to achieve the global reliability of 80%.
   b) Derivative. It shows us the derivative of the previous graphic. Thanks to this graph it is possible to analyze what is the point from which the number of needed retransmissions goes up exponentially. Therefore, this graphic shows developers that when

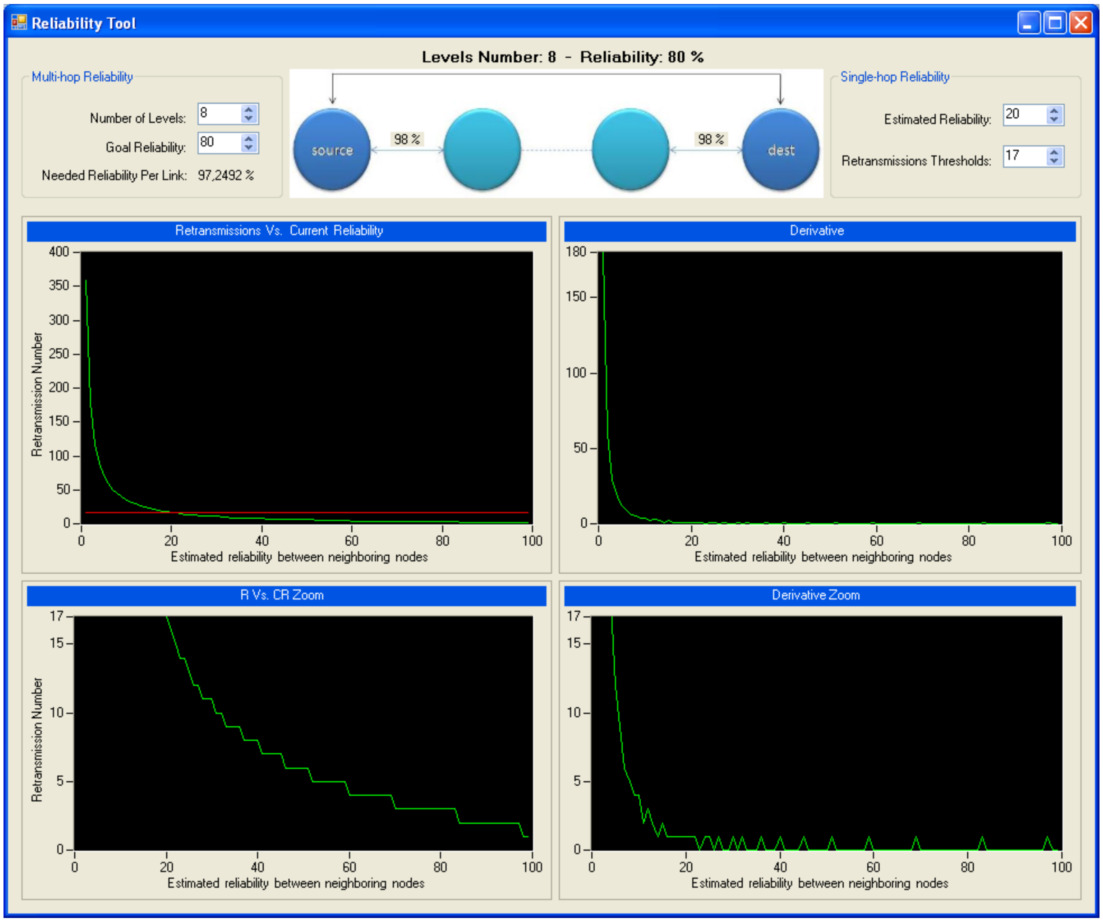

Fig. 1. Reliability Tool

the estimated reliability between two nodes is less than 20%, it is not advisable to retransmit a packet N times since N could be so big that the energy consumption of the nodes would be too costly.

c) R Vs. CR Zoom. It is a zoom of the "Retransmissions Vs. Current Reliability" graphic. The zoom is established on the basis of the red line mentioned above.

d) Derivative Zoom. This graphic shows us an interesting piece of information. Let us focus on the range 60-80, we can observe two plain signals in the subranges 60-70 and 72-80. This means that when estimated reliability between two nodes fall into one of these ranges, the number of required retransmissions is equal and therefore, the energy consumption is also the same. Developers may think that is is more costly (in terms of energy consumption) to achieve a reliability of 98% when the estimated reliability is 60% than to achieve the same when the estimated reliability is 70%. This graph reveals the ranges where energy consumption is the same independently wether the value of the estimated reliability is higher or lower.

On the other hand, this information allows the protocol to be more efficient at distributing the energy consumption over the whole network.

Figure 2 shows another different way of analyzing and understanding the relation between link quality among neighbors, the number of retransmissions to increase these link qualities and the desired goal reliability between two nodes which are L hops away from each other. The data represented in the figure have been generated by using the equation 4 and taking into account that L is equal to 9. In order to understand the data, let us focus on the gray area of the figure. For example, let us assume that a developer wants to achieve a reliability about a 90%. The figure indicates that if the estimated link qualities of the nodes are about a 89%, 77% or 67%, the protocol will need 2, 3 or 4 retransmissions respectively to achieve the desired reliability goal. Now, let us imagine that the estimated link quality of two nodes is greater than 67% and lower than 77%. In this case, the number of necessary retransmissions will have decimals (3.3, 3.4, …). Thus, in order to ensure that the reliability goal (90%) is achieved, protocol will use the next integer. In this particular case, it would be the number 4. It could cause the final reliability goal is greater than 90%. We have considered that it is better

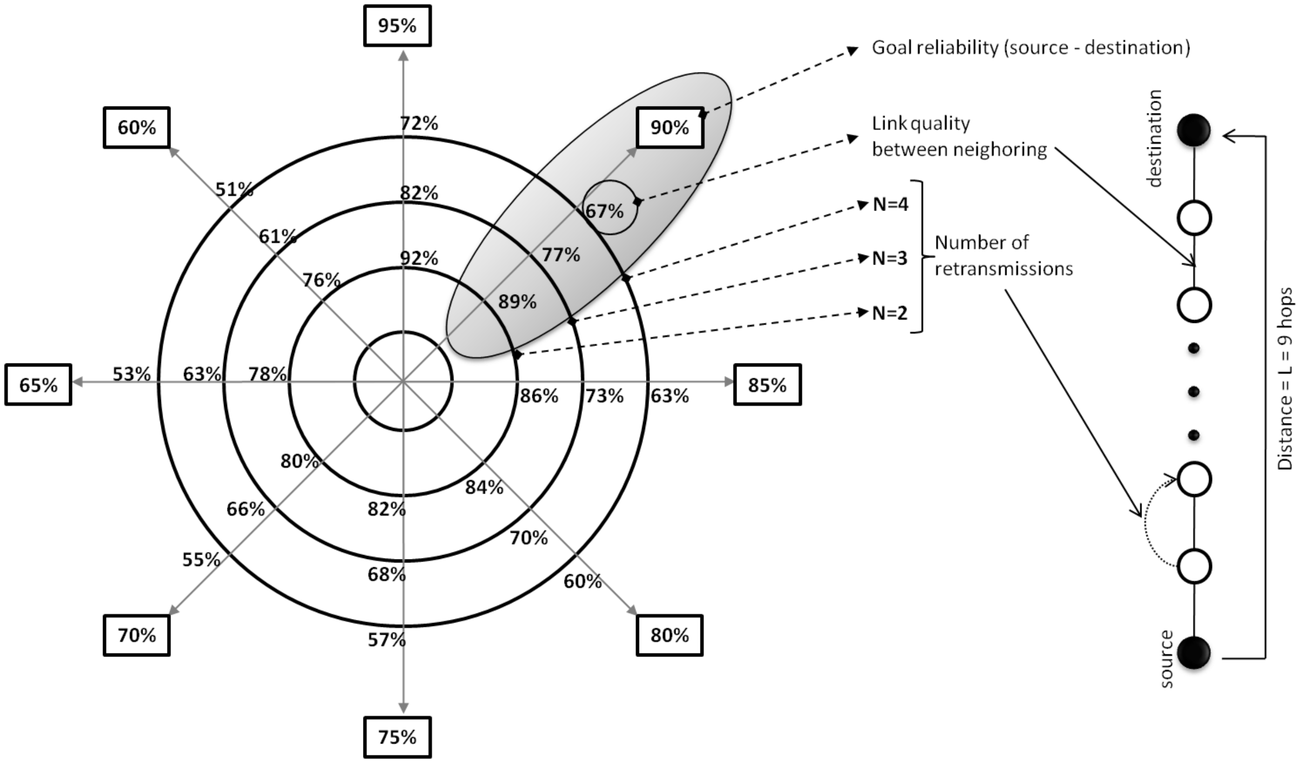

Fig. 2. Reliability analysis between two nodes which are 9 hops away from each other

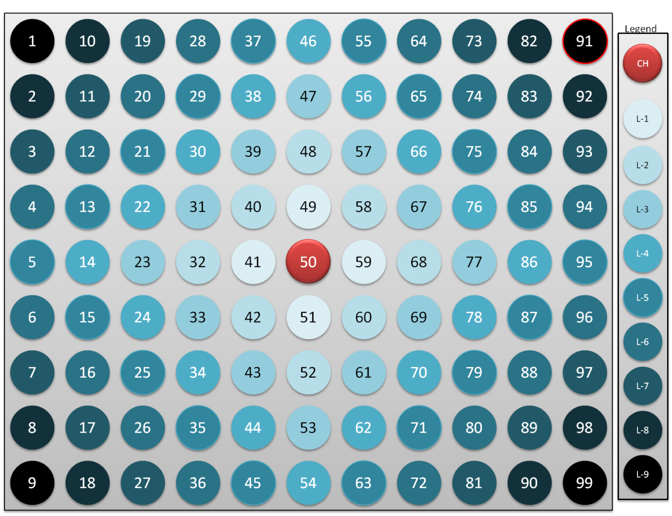

Fig. 3. Evaluation Topology

to achieve a greater reliability in these kinds of situations.

VI. Evaluation

In order to analyze and study the performance of the reliability model presented in this work, several experiments has been carried out. Figure 3 shows the cluster topology used in the simulations to carry out the experiments. A square grid topology with 99 nodes has been used, as it is quite a standard configuration and in addition, it can also represent very well a cluster of nodes. Basically, the experiments have consisted in sending 100 packets (events and commands) between the cluster-head (node 50) and a sensor node (node 91) which are 9 hops away from each other, in order to measure how accurate the reliability achieved by the model is.

A. Environment set-up

COOJA sensor network simulator [12] has been used to carry out all the experiments. COOJA is a power profiling tool that enables accurate network-scale energy measurements in a simulated environment. COOJA simulator offers the possibility of carrying out the simulation in different platforms. We selected the TelosB motes since they are one of the most used by the sensornet community. In order to carry out the simulations, we used Contiki [13] which is an open source, highly portable, multi-tasking operating system for memory-efficient networked embedded systems and wireless sensor networks.

B. Results

In order to evaluate our reliable protocol several experiments have been carried out by using the simulator Cooja. The goal of the experiments was to analyze the packet delivery ratio (PDR) of the protocol after sending 100 packets in both directions, from a sensor node to the cluster-head and viceversa. Concretely, the protocol was configured to achieve a reliability level of 80% between two nodes which were 9 hops away from each other. This scenario was studied by using different link quality estimations. Table I shows the results obtained from the simulations:

Packet Delivery Ratio					
Nodes Link Quality	Theoretical Reliability	Reliable Events	No Reliable Events	Reliable Commands	No Reliable Commands
84%	20%	82%	14%	94%	58%
70%	4%	84%	2%	97%	11%
60%	1%	87%	0%	92%	2%
52%	0%	66%	0%	76%	0%

TABLE I
RELIABILITY RESULTS

- Column 1 shows the four different link qualities set in the simulator.
- Column 2 shows the the theoretical reliability between two nodes which are 9 hops away from each other. For example, if the link quality of the nodes is 84%, the reliability between the two nodes mentioned above will be 0.84^9 which is equal to $0.2082 = 20\%$.
- Columns 3 and 4 show the achieved reliability level after sending 100 events (communication from a sensor node to the cluster-head) by using the reliable protocol and the basic protocol.
- Columns 5 and 6 show the achieved reliability level after sending 100 commands (communication from the cluster-head to a sensor node) by using the reliable protocol and the basic protocol.

Let us analyze the results shown in table I. On the one hand, it can be appreciated that the PDR achieved transmitting events is about 80% despite that the theoretical reliability is much more lower. Actually, when the events are sent without using the reliable protocol the results obtained (column four) are similar to the reliability level indicated by the theoretical reliability. On the other hand, the result obtained from the commands are quite different. When they are transmitted the reliability level is higher than 80%. It is due to the fact that when the packets are sent from the cluster-head to the sensor nodes, they can take several paths. Finally, it can also be appreciated that when link qualities of the nodes is set to the 52% the achieved reliability level is lower than 80% (in both cases, event and commands). We believe that it could be due to the collisions produced in the sensor networks since the number of retransmissions is quite high.

VII. CONCLUSIONS

In this paper, we have presented a tool to study and understand how both link quality estimations and distances between source and destination nodes affect the communication reliability. As a novel contribution we have presented a reliability model which allows developers to quantitatively set the desired reliability between a sensor node and its leader node whatever the distance between them is. Finally, a set of experimenters have been carried out to prove the suitability of the proposed model. The results obtained show that when the link quality estimations are greater or equal to 60% the performance of the reliability model is quite accurate.

ACKNOWLEDGMENT

This work was partially supported by Spanish Projects TIN2008-03107 and TIC-03085.

REFERENCES

[1] I. Akyildiz and M. Vuran, *Wireless sensor networks*. LibreDigital, 2010.
[2] J. Gehrke and L. Liu, "Sensor-network applications," *IEEE Internet Computing*, vol. 10, no. 2, 2006.
[3] A. Arora, R. Ramnath, E. Ertin, P. Sinha, S. Bapat, V. Naik, V. Kulathumani, H. Zhang, H. Cao, M. Sridharan *et al.*, "Exscal: Elements of an extreme scale wireless sensor network," 2005.
[4] A. Willig and H. Karl, "Data transport reliability in wireless sensor networks–a survey of issues and solutions," *Praxis der Informationsverarbeitung und Kommunikation*, vol. 28, no. 2, pp. 86–92, 2005.
[5] E. C. H. Ngai, Y. Zhou, M. R. Lyu, and J. Liu, "Reliable reporting of delay-sensitive events in wireless sensor-actuator networks," in *Mobile Adhoc and Sensor Systems (MASS), 2006 IEEE International Conference on*, Oct. 2006, pp. 101–108.
[6] B. Deb, S. Bhatnagar, and B. Nath, "Reinform: reliable information forwarding using multiple paths in sensor networks," in *Local Computer Networks, 2003. LCN '03. Proceedings. 28th Annual IEEE International Conference on*, Oct. 2003, pp. 406–415.
[7] K. Sharma, H. Singh, and R. Patel, "A Reliable And Energy Efficient Transport Protocol for Wireless Sensor Networks," *Global Journal of Computer Science and Technology*, vol. 10, no. 9, 2010.
[8] H. Zhou, Y. Wu, Y. Hu, and G. Xie, "A novel stable selection and reliable transmission protocol for clustered heterogeneous wireless sensor networks," *Computer Communications*, 2010.
[9] J. Paek and R. Govindan, "RCRT: Rate-controlled reliable transport protocol for wireless sensor networks," *ACM Transactions on Sensor Networks (TOSN)*, vol. 7, no. 3, pp. 1–45, 2010.
[10] E. Felemban, C. Lee, and E. Ekici, "MMSPEED: Multipath multi-SPEED protocol for QoS guarantee of reliability and timeliness in wireless sensor networks," *IEEE Transactions on Mobile Computing*, pp. 738–754, 2006.
[11] P. Misra, N. Ahmed, D. Ostry, and S. Jha, "Characterization of Asymmetry in Low-Power Wireless Links: An Empirical Study," *Distributed Computing and Networking*, pp. 340–351, 2011.
[12] J. Eriksson, F. Österlind, N. Finne, A. Dunkels, N. Tsiftes, and T. Voigt, "Accurate Network-Scale Power Profiling for Sensor Network Simulators," *Wireless Sensor Networks*, pp. 312–326, 2009.
[13] A. Dunkels *et al.*, "Contiki-a lightweight and flexible operating system for tiny networked sensors," 2004.