A Performance Comparison of Data Dissemination Protocols for Wireless Sensor Networks

Tatiana Bokareva, Nirupama Bulusu and Sanjay Jha

E-mails: tbokareva@cse.unsw.edu.au, nbulusu@cse.ogi.edu, sjha@cse.unsw.edu.au

***Abstract*—In recent years a variety of new data dissemination protocols have been developed specifically for wireless sensor networks (WSN), but no realistic performance comparison between them has been attempted. This paper reports on the results of a simulation comparison made by an independent researcher using the ns-2.26[1] simulator for the WSN protocols: *Directed Diffusion (DD)*[2], *Two-Tier Data Dissemination (TTDD)* [3] and *Gradient Broadcast (GRAB)*[4].**

Our performance study provides useful insights for the network designer - such as which protocols (and design choices) scale control traffic well, improve data delivery or reduce overall energy consumption. We observe that despite the designers intentions to make these protocols self-configuring, they in fact rely on a number of statically configured parameters which are the cause of the reduction in peformance. For example, the static preconfiguration of the cell size in TTDD. is one of the reasons why TTDD exhibits larger routing overhead than DD by 67.6% on average. Although GRAB produces approximately 93.6% smaller overhead than TTDD and 89.27% smaller than DD, because of statically configured amount credit GRAB delivers on average 6 times more of the redundant data packets than TTDD and DD. We suggest that making these protocols truly self-learning can significantly improve their performance, and comment on how some of these parameters can be dynamically derived through measurements of network and event dynamics.

I. Introduction

A WSN (Wireless Sensor Network) can be described as an ad-hoc network formed by a large collection of very simple devices combining sensing, computation and communication abilities[5]. These kinds of devices could dynamically form a network without support from existing infrastructure and human administration[5]. A network of such devices has a variety of applications ranging from high-profile military applications through to civilian applications such as environmental monitoring[5]. Once deployed, nodes in the network in many current applications of wireless sensor networks are stationary.

There is extensive research occurring in the area of protocol design for WSNs. Communication in wireless sensor networks is data-centric and must minimize the energy consumed by unattended battery-powered sensor nodes[2]. As a result of this many different data dissemination protocols have been proposed to solve WSN challenges[2][3][4] [6][7][8].

Each design is based on different assumptions and intuitions regarding the application scenarios of the network and its operational behavior. Although each of the protocols claims to solve some of the challenges identified during the development process, little is known about the *relative* performance of these protocols as there have been no significant attempts to compare these protocols to each other.

This work would appear to be the *first* of its kind to provide realistic analysis of the comparison of the protocols. The main significance of this work lies in its attempt to formalize some of the comparison parameters and procedures for the evaluation of WSN protocols.

This work does not attempt to find the best possible operational scenario for each protocol. Instead it concentrates on the direct comparison of three recently proposed protocols under a set of given scenarios: Directed Diffusion(DD)[2], Two-Tier Data Dissemination(TTDD)[3] and Gradient Broadcast(GRAB)[4].

Contributions:

- Our performance study reveals useful insights for the network designer — such as which protocols (and design choices) scale control traffic well, improve data delivery or reduce overall energy consumption so that they can be used in the future enhancement of protocols. One of the surprising finding was that although GRAB on average produces 93.6% less routing overhead than TTDD and 89.2% less routing overhead than DD, overall DD consumes 3.5% less amount of energy than TTDD and 20.6% less than GRAB . TTDD produced 67.6% larger amount of routing traffic than DD. TTDD and DD have very similar data delivery ratio and they are both very close to the ideal one, whereas GRAB delivers on average 6 times more of redundant data packets across simulations.
- Our key observation is that despite their design intentions to make these protocols self-configuring, they in fact rely on a significant number of statically configured parameters. We suggest which parameters for each protocol should be dynamically configured in response to *measured* network state, using passive measurement techniques such as Bayesian inference to reduce the measurement overhead. Making these protocols truly self-learning techniques could significantly improve their performance.

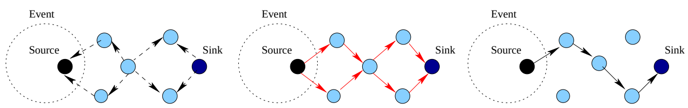
The rest of the paper is organized as follows. A brief overview of the each protocol that is included in this work is given in Section II. Section III is devoted to an explanation of the simulation environment. The results of the direct comparisons are located in Section IV. Section V outlines suggestions for improving individual protocols. Finally Section VI sets out to conclude this work.

II. Data Dissemination Protocols Overview

This section briefly describes each data dissemination protocol used in the comparison — Directed Diffusion, Two-Tier Data Dissemination, and Gradient Broadcast.

A. Directed Diffusion

Directed Diffusion[2](Figure 1) is the first proposed data-centric communication protocol for wireless sensor scenarios. The data generated by the producer is named using attribute-value pairs. The consumer node requests the data by periodically broadcasting an interest for the named data. Each node in the network will establish a gradient towards its neighboring nodes from whom it receives the interest. The gradient specifies both the data rate and the direction towards which the data should be sent. Once the producer detects an interest it will send exploratory packets towards the consumer, possibly along multiple paths. As soon as the consumer begins receiving exploratory packets from the producer it will reinforce one particular neighbor from whom it chooses to receive the rest of the data. The data will then flow back towards the consumer along the reinforced path. The path reinforcement packets are also used for local path repairs in case of the failure of some nodes during the data delivery phase.



A) Interest propagation B) Initial gradients set up C) Data delivery

Fig. 1. Directed Diffusion

B. Two-Tier Data Dissemination

TTDD[3] (Figure 2) is based on decentralized architecture[9]. It uses a grid structure to divide the topology into cells. Only sensors located at a cell boundary need to forward the data. The consumer actively builds the grid structure through the network and sets up forwarding points in the sensors closest to the grid boundary called dissemination nodes (DN). One tier is the cell at the consumer's current location and the other one is the DN at cells boundaries. The consumer only floods the query within its own cell. When the nearest DN that hears the query, it forwards it to its adjaicent DNs. This process continues until the query reaches the producer or one of the DNs that have the corresponding data. During the query propagation period the network establishes the reverse path towards the consumer, so that it can enable the data path to be the same as that of the query propagation.

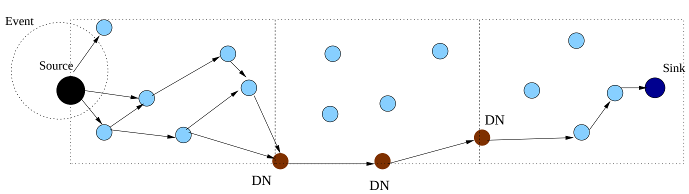


Fig. 2. TTDD

C. Gradient Broadcast

In GRAB[4], a node on deployment sets its cost to reach the consumer at infinity. As soon as the consumer node starts up it broadcasts the advertisement message containing its initial cost. Each intermediate node that hears the advertisement will calculate the receiving cost of the message. At the end of the cost-field setup period each working node will have calculated the minimum cost for it to reach the consumer. Each message carries a "credit" in its header. Depending on the "credit" amount data packets can flow along multiple paths if the "credit " is set to be higher that the minimum cost. Each intermediate node will make its own dicision regarding the forwarding of a packet

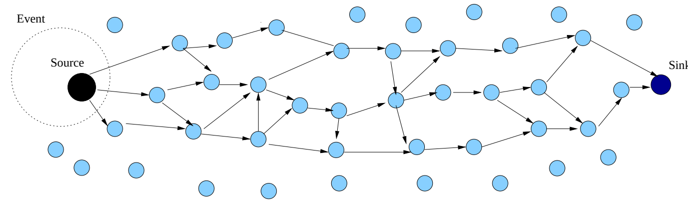


Fig. 3. GRAB

based on the amount of credit in the data message, its own minimum cost value and the remaining ratio.

Each data packet carries in its header the minimum cost of the producer to reach the consumer($Csource$), some constant(α), the amount of current energy used(Pa), and the sender's minimum energy (Ca). The formula for calculating the remaining ratio is:
Let
Cr = remaining ratio
Th = the threshold.
Then

$$Cr = \frac{Csource - (Pa + Ca - Csource)}{\alpha \times Csource} \tag{1}$$

$$Th = (\frac{Ca}{Csource})^2 \tag{2}$$

and if Cr is bigger than Th then a node will rebroadcast a message.

III. Metrics and Methodology

This section describes the simulation methodology and the metrics used for the comparison of protocols.

A. Methodology

Ns-2.26[1] was used for the simulation of protocols. Each of the data dissemination protocols studied has the same underlying IEEE 802.11 MAC layer, the same radio propagation model based on the 914Mhz frequency of the Lucent WaveLan DSSS radio with omni-directional antenna placed 1.5 meters above the node and the same data load. 2 different topologies with uniformly distributed nodes have been generated. The size of the topology, the number of nodes that are deployed and the NINRA (the Number of nodes In Nominal Range Area) can have significant impact on protocol behavior. The same topology scenarios are used across different protocol simulations. Given the radio range of a node, the topology size and the number of nodes deployed NINRA represent the largest possible number of neighbors that a node can hear from and it is calculated according to the following formula:

$$NINRA = \frac{N}{A} \times (\pi \times R^2) \tag{3}$$

Where:
N= the number of nodes in a topology
A = the area of the topology
R = the radius range of the radio.
Table I shows the parameters used for generating the various simulation topologies.

TABLE I
TOPOLOGY PARAMETERS USED IN SIMULATION.

Number of nodes	*Dimensions*	*NINRA*
20	360x360	30
40	511x511	30
60	626x626	30
80	723x723	30
100	809x809	30
120	886x886	30

To represent the worst case scenario only one producer and one consumer used for each simulation. The producer and consumer are located at opposite sides of the topology so that a large number of sensor nodes in the topology, participate in the protocol. Six different topology scenarios are used for the simulation. The first one consists of 20 nodes in the topology. The number of nodes deployed is progressively increased by 20 until there are 120 nodes in the topology. Data packets are generated at intervals of 1 second. The simulation is run for 500 seconds therefore each protocol has enough time to discover the route from the consumer to the producer and produce substantial amount of data traffic.

B. Metrics

For the evaluation of protocols the following three metrics have been chosen[7]. Each metric is evaluated as a function of the topology size, the number of nodes deployed, the NINRA and the data load of the network.

B.1 Average Energy Consumption (Ea)

The average energy consumption is calculated across the entire topology[10]. It measures the average difference between the initial level of energy and the final level of energy that is left in each node. Let
Ei = the initial energy level of a node
Ef = the final energy level of a node
n = number of nodes in the simulation
Then

$$Ea = \frac{\sum_{k=1}^{k=n}(Ei_k - Ef_k)}{n} \tag{4}$$

This metric is important because the energy level that a network uses is proportional to the network's lifetime. The lower the energy consumption the longer is the network's lifespan.

B.2 Routing Overhead(R)

This metric represents the total amount of routing packets transmitted during the simulation time. Let
Tr = the total amount of routing packets that a node transmits during the simulation
n = the number of nodes deployed
Then

$$R = \sum_{k=1}^{k=n}(Tr_k) \tag{5}$$

This metric is important for the comparison of these protocols as it indicates the scalability of a protocol. Each protocol has to function in low bandwidth and congested environments, so this metric is a good indication of the degree of functionality for a protocol and its efficiency in terms of resources consumption. Also it operates as a very good indication of how much effort is needed to construct and maintain a route between the producer and the consumer.

B.3 Packet Delivery Ratio(r)

This metric represents the ratio between the number of data packets that are sent by the producer and the number of data packets that are received by the consumer. Let
Ps = the number of data packets sent by the producer
Pc = the number of data packets received by the consumer including duplicates.
Then

$$r = \frac{Pc}{Ps} \tag{6}$$

This metric indicates both the loss ratio of the routing protocol and the effort required to receive data. In the ideal scenario the ratio should be equal to 1. If the ratio falls significantly below the ideal ratio, then it could be an indication of some faults in the protocol design. However, if the ratio is higher than the ideal ratio, then it is an indication that the consumer receives a data packet more than once. It is not desirable because reception of duplicate packets consumes the network's valuable resources. The relative number of duplicates received by the consumer also important because based on that number the consumer, can possibly take an appropriate action to reduce the redundancy

IV. COMPARISON RESULTS

A. Routing Overhead

Figure 4 shows the relative routing overhead for all three protocols. As can be seen, TTDD exhibits the largest routing overhead. This is an indication that the grid construction and maintenance operation is very expensive for TTDD in terms of the routing overhead. Additionally, the size of the cells plays a major role in the way TTDD behaves. For the current version of TTDD the cell size has to be set up before the simulation and there is no way for the protocol to change it in order to respond to changes in its environment.

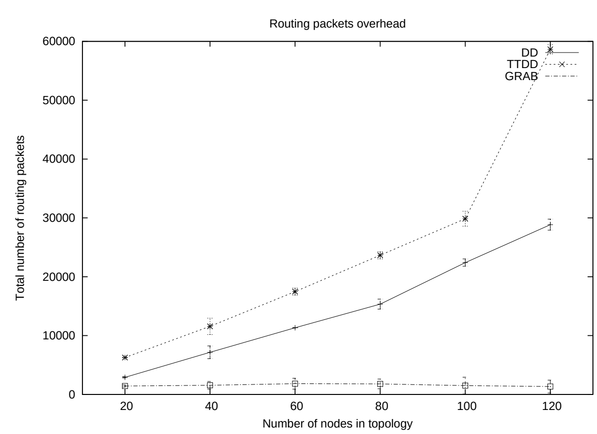


Fig. 4. Routing overhead

Although as shown in Figure 5 the number of routing overhead packets produced by GRAB fluctuates significantly across simulations and therefore it has most unpredictable behavior in terms of routing overhead. Overall it has generated the smallest routing overhead. The refreshment of the cost field in response to major changes appears to be a very positive feature of GRAB.

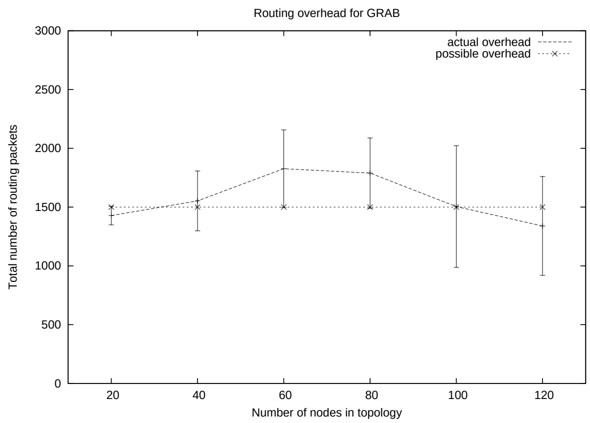


Fig. 5. GRAB overhead

B. Delivery Ratio

Figure 6 shows the relative delivery ratio of data packets for all the protocols. TTDD and DD have very similar delivery ratios and very close to the ideal one. DD, however has slightly more fluctuations. GRAB on the other hand has a larger delivery ratio than the other two protocols with a very large error bars. Therefore even for the constant amount of the credit and the stable topology of nodes we can not predict the exact delivery for GRAB at the beginning. It is also much higher than the ideal one. This feature of GRAB may increase the robustness of data delivery in the case of noisy channels. However, this feature is not particularly desirable while operating on clear channels, as it leads to high energy consumption.

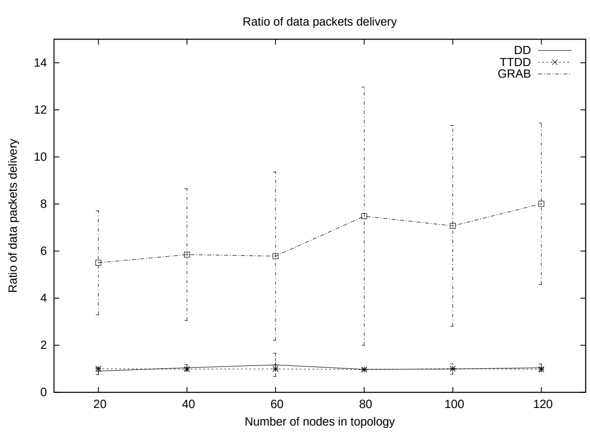


Fig. 6. Delivery ratio

C. Average Energy Consumption

Figure 7 shows the relative energy consumption of all three protocols. As expected GRAB has a much higher average energy consumption compared to TTDD and DD. TTDD and DD have very similar energy consumption with TTDD being slightly higher form 20 to 80 and at 120 nodes. TTDD also performed marginally better for the 100 nodes scenario. The reason for this is that TTDD limits its flooding of packets to one cell. Therefore the choice of the cell size is a very important parameter. Table II shows the approximate number of cells for TTDD based on the number of nodes deployed and the default size of the topology.

TABLE II

Number of nodes	*Dimensions*	*Number of cells*
20	360x360	3
40	511x511	7
60	626x626	10
80	723x723	13
100	809x809	16
120	886x886	19

If the number of cells is small then TTDD will flood its data very similarly to DD. Also the routing overhead for TTDD is higher than that for DD. This two main factors lead to higher energy consumption by TTDD compared to DD. As the topology grows the number of cells grows and the flooding is therefore constrained to an area of the network. This is the explanation for the slowing in the increase of the slope of the function from 60 to 100 nodes, except for the 120 nodes topology. The average energy consumption for TTDD is higher at that point due to the large routing overhead. The ideal cell size for a given topology is not investigated in this work. The original paper assumes that the size of cells is supplied through some external mechanism.

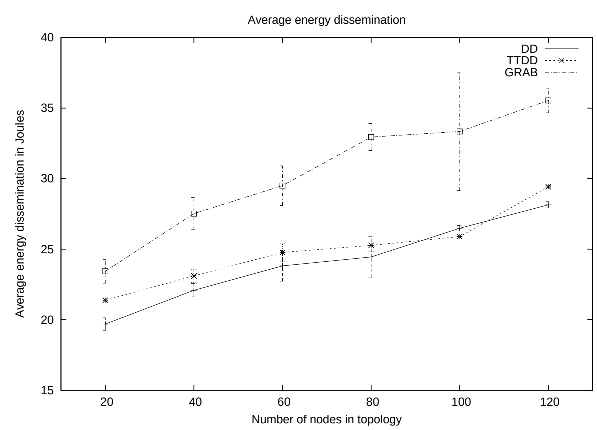


Fig. 7. Average energy

V. Suggestions for Protocol Improvement

A key finding from the analysis relates to the configuration of a protocol's parameters. Comparisons revealed that the performance of a protocol was enhanced where its parameters was not inflexibly predetermined but rather, could be varied by adapting to its environment. To boost their performance, we suggest that making these protocols truly self-learning by configuring protocol parameters in response to *measured* network state, using passive measurement techniques such as Bayesian inference. Below, we comment on which of these parameters needs to be

TABLE III

	GRAB	TTDD	DD
Scaling Control Traffic	Lowest (Best)	Highest	Close to TTDD
Data Delivery Ratio	6 (High redundancy)	1 (Ideal)	0.8 - 1
Mean Energy Consumption	Highest	Close to DD	Lowest
Static Parameters	Credit	Cell size	Refresh rate
Suggestions	Reduce redundancy by adapting credit according to measured path loss	Reduce routing overhead by adapting cell size according to measured network density diameter	Improve data delivery by adapting refresh rates according to measured path latencies

dynamically derived in the case of each protocol, and based on which measured variables.

TTDD has a static cell size. Performance would be improved if it could learn its own network topology variables such as the network density and diameter, and adjust its cell size according to the environment in order to limit the amount of flooding that occurs. One possible enhancement to TTDD can be that it takes the advantage of the knowledge of geographical positions of its nodes. Each node in the topology knows its own coordinates and the number of immediate neighbors, therefore it can estimate the relative density and NINRA to it self. During the construction of the grid and the discovery of producers location the consumer can aggregate this information and negotiate an appropriate cells size.

GRAB could be improved by adding the ability of its consumer to adjust the credit that the data packet carries in order to reduce the redundancy. In the real live deployment there is no way to determine in advance what is the optimal ”credit” should be. This credit could be a function of the applications reliability requirements, and dynamically configured as a function of the mean percentage packet loss, along a given path, which can be dynamically derived based on statistical inference. GRAB can begin with the maximum amount of the ”credit” namely 2.5 and during the cost field refreshment period the consumer can indicate how much credit should be given to a data packet based on the number of duplicates received. If its delivery ratio is to high then the amount of credit and subsequently the mesh width should be reduced. However if the delivery ratio is below 1 it should widen the width of the mesh by indicating the enlargement in the amount of the ”credit”.

DD could reduce the routing overhead by reinitiating its refreshment of interest only when the major changes in the topology are detected. Or it could calibrate its refresh period based on measurements of mean path latency, to improve its path stability, and consequently the stability of its data delivery rate. Table III summarizes the results and suggestions of our performance comparison.

VI. Conclusion

This paper presented the *first* comparative analysis made by independent researcher of three available data dissemination protocols for wireless sensor networks, using ns-2.26[1] simulations — Directed Diffusion (DD)[3], Two-Tier Data Dissemination (TTDD)[8] and Gradient Broadcast (GRAB)[9]. These protocols cover a large number of design choices including the construction of the grid, credit-based adjustable mesh forwarding and the establishment of gradients for neighboring nodes.

Typically, when these protocols are studied in isolation, the emphasis is on studying only the scaling behavior of the protocol (for example, the impact of network density on scaling behavior). Such an approach can mask the design weaknesses of a particular protocol. Being a relative performance comparison, this study is the first to provide useful insights to what kind of design choices are the most desirable in order to improve the performance of proposed protocols. Each of the protocols studied performed well in some cases, but displayed certain drawbacks in others. The performance of TTDD and Directed Diffusion was quite close, with GRAB’s performance being most distinctive.

TTDD has 67.6% large routing overhead but consumes only 3.6% more energy than DD. This is due to the nature of data forwarding in TTDD. It constrains the flooding of data packets to one cell. However, for large cell sizes relative to the topology size it floods the data in a very similar way to the flooding of interests and exploratory packets used by DD.

GRAB has 89.3% smaller routing overhead than DD and 93.6% smaller than TTDD because of the way it refreshes its minimum cost at each node. The cost is refreshed only when there are major changes in the network topology are detected or the delivery of the data has been delayed. However, because of the way it forwards its data to the consumer it consumes redundantly 26% large amount of energy compared to DD and 21.7% larger compared to TTDD. Overall DD consumes 3.5% less amount of energy than TTDD.

Finally, TTDD has a slightly closer delivery ratio to the ideal ratio than does DD, although the delivery ratios are very similar in both of these protocols. DD appears to have larger fluctuations for the delivery ratio of data packets. The smallest ratio of data packets delivery was approximately 0.8 whereas the TTDD delivery ratio did not fall below 0.9 during the simulation period.

We suggested that parameters for each protocol such as credit (GRAB), cell size(TTDD), and refresh rate (DD) should be dynamically configured in reponse to *measured* network state, such as path loss, latency, network density and diameter, using passive measurement techniques such as Bayesian inference. In summary, making these protocols truly self-learning could significantly improve their performance.

References

[1] USC/ISI UC Berkley, LBL and Xerox PARC. *http://www.isi.edu/nsnam/ns/ns-documentation.html*.

[2] Chalermek Intanagonwiwat, Ramesh Govindan, and Deborah Estrin. Directed diffusion: A scalable and robust communication paradigm for sensor networks. In *Proceedings of the Sixth Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom 2000)*, pages 56–67, Boston, MA, USA, August 2000. ACM Press.

[3] Jerry Cheng Songwu Lu Fan Ye, Haiyun Luo and Lixia Zhang. Two-tier data dissemination model for large-scale wireless sensors networks. In *Proceedings ofACM/IEEE International Conference on Mobile Computing and Networking (MobiCom 2002)*, Atlanta, Georgia, USA, September 2002.

[4] Songwu Lu Fan Ye, Gary Zhong and Lixia Zhang. Gradient broadcast: A robust data delivery protocol for large scale sensor networks. In *IPSN*, Palo Alto, CA, USA, April 2003.

[5] Crossbow Technologies Incorporated. *http://www.xbow.com*.

[6] Wendi Heinzelman, Anantha Chandrakasan, and Hari Balakrishnan. Energy-efficient communication protocols for wireless sensor networks. In *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences (HICSS)*, pages 3005–3014, Big Island, Hawaii, USA, January 4-7 2000.

[7] Lakshman Krishnamurthy Chie-Yih Wan, Andrew Cambell. A reliable transport protocol for wireless sensors networks. In *First Workshop on Sensors Networks and Applications(WSNA)*, Atlanta, CA, USA, September 2002.

[8] Adrian Perrig, Robert Szewczyck, Victor Wen, David Culler, and Doug Tygar. Spins: Security protocols for sensor networks. In *Proceedings of the Seventh Annual International Conference on Mobile Computing and Networking (ACM MOBICOM '01)*, pages 189–199, Rome, Italy, July 2001. ACM.

[9] I-Jeng Wang and Steven D.Jones. The scalability of a class of wireless sensors networks. In *Modeling and Design of Wireless Networks*, page Vol.4531, 2001.

[10] David B.Johnson Yih-Chun Hu Jorjeta Jetcheva Josh Broch, David A.Maltz. A performance comparison of milti-hop wireless ad hoc network routing protocols. In *Proceedings of the Fourth Annual ACM/IEEE International*. ACM Press, August 1998.