Automated Meter Reading and SCADA Application for Wireless Sensor Network

Francisco Javier Molina, Julio Barbancho, and Joaquin Luque

Departamento de Tecnologia Electronica, University of Seville,
C/ Virgen de Africa, 7. Seville 41011, Spain
{fjmolina,jbarbancho,jluque}@us.es
Tel.: (+034) 954 55 28 35, Fax: (+034) 954 55 28 33

Abstract. Currently, there are many technologies available to automate public utilities services (water, gas and electricity). AMR, Automated Meter Reading, and SCADA, Supervisory Control and Data Acquisition, are the main functions that these technologies must support. In this paper, we propose a low cost network with a similar architecture to a static ad-hoc sensor network based on low power and unlicensed radio. Topological parameters for this network are analyzed to obtain optimal performances and to derive a pseudo-range criterion to create an application-specific spanning tree for polling optimization purposes. In application layer services, we analytically study different polling schemes.

Keywords: Automated Meter Reading Application, SCADA, Ad Hoc Networks, Spanning Tree Algorithm, Multihop Routing Protocol.

1 Introduction

Since the 70's, many technologies have been developed for Automatic Meter Reading functions (AMR) and Distribution Automation (DA) for utility applications (water, gas and electricity) [1,2]. Many studies show that solutions based on low power radio networks are viable and that they offer the best cost/performance ratio[3,4,5]. However, it is only in the late 90's that, radio and microcontroller technologies have allowed the development of smart sensor networks. We propose (in this paper) the use of ad-hoc network technologies to support this application because:

- ad-hoc protocols are best suited to low power systems,
- nodes can be located without pre-planning,
- and topology is more flexible, making management simpler.

Public utilities' management has many different aspects closely interrelated that must be coordinated within a corporative network: (e.g. Meter reading from customer meters, Distribution management, Economic dispatch. . .) These applications are often distributed throughout many computers. But, customer data polled from sensor networks, queries, remote control orders and network management messages must be processed by a unique computer named UC -

Utility Controller. UC works as the master that controls many remote units (sensor nodes), like a well known architecture called SCADA, Supervisory Control and Data Acquisition. We have called this application ASCADA, Augmented SCADA, as it has to support typical SCADA functions and additional AMR services. Some of these are listed below:

a) **AMR functions.** Meter reading or checking an individual customer, a group-cluster of meters, and all meters (global reading). These queries could be simultaneous, and the execution time must be as short as possible to reduce the reading period. Nowadays, the on-site reading period, carried out by an operator, is about two months. The goal is to manage global readings daily or weekly.
b) **Telemetry functions.** These services obtain data from sensors, and they control some elements located at selected points of the distribution network (flow, power, state of valves or switches, etc.) Distribution sensing and automation will enhance supply services, reducing failure, alarm and response times. All this data must be polled periodically, and the completion time must be as short as possible to reduce the bandwidth load.
c) **Remote control orders.** Security and reliability are the main characteristics of these services. Minimization of transmission time is a general objective. In this case, it is quite important to reduce multiple hops and provide dynamic routing capabilities to enhance reliability. Further functions would be:
   – encrypted data and sender identification for secure operations,
   – order sequencing to avoid duplication,
   – receipt request for confirmation,
   – message transmission indicating the end of the command.
c) **Alarm transmission.** From distribution elements, nodes detect transmissions caused by an exceptional situation. Nodes from customer meters must not have this service to avoid network overload.

Some AMR services need to use a high percentage of network capacity. This fact will be present throughout the paper. The next section analyses topology characteristics more closely related to individual and overall polling. Section 3 presents different strategies to compute and optimize global polling and simulation results. Finally, we outline future work and alternative solutions that are currently being tested.

2 Topological Model and Properties

The IEEE working group SC-31 has proposed a set of topological models for AMR systems. Figure 1 shows the model based on a fixed radio network. Each element is conceptual and does not necessarily exist in the form shown. Devices in the topology could be combined or reduced to a null element. When they are present, elements A, B, C, D and X are intermediate devices. The topology can support multiple delivery points, as shown, separated by service boundaries.

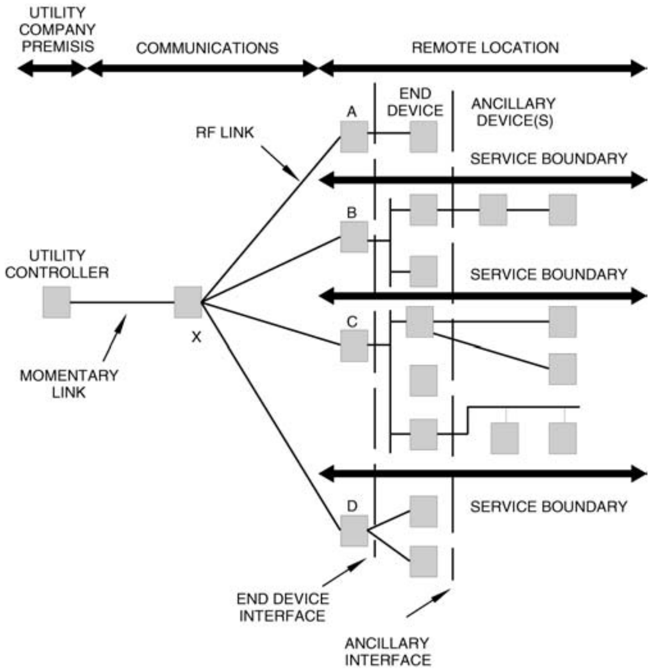

Fig. 1. Radio AMR topology model.

Table 1. Radio ranges on different scenarios and 100mW radiated power.

Band	**Indoors**	**Outdoors (with obstacles)**	**Outdoors (with in line-sight)**
433 MHz	<100 meters	<500 meters	1 - 2 Km
800 MHz	<50 meters	<300 meters	0.8 - 1.5 Km
900 MHz	<30 meters	<200 meters	< 800 meters

Sensors connected to End Device and data stored in memory may vary depending on location or application: customer buildings (meters) or distribution surveillance and control (flow meters, power meters, valves or switchers states).

For this project, only application, network, link and physical layers have been developed. Many of them were inspired by ad-hoc network protocols [6,7,8]. Application layer services include ASCADA functions for local and cluster meter reading, tampering warning, alarm warning and the remote control of actuators. The Network Layer supports a multihop routing protocol and network topology maintenance. Data Link Layer offers a medium access contention (CSMA-CA), sequence and synchronization control. But the Physical Layer is probably one of the most important aspects. License exempt standards fix frequency bands and radiated power output for telemetry applications, so consequently radio range is also limited. Table 1 shows European ETS300-220 limits and bands for telemetry, and radio ranges for different locations.

The End Devices used include radio OEMs which are compliant with ETS 300-220. In European cities, most of them are located indoors (but shouldn't), so the radio range ends up being about 100 m (433 MHz band). The Radio Network consists of many End Devices forming a dense network where each node needs a multihop transmission to reach the Utility Controller node. There is no planning to select node locations, so the network topology has an arbitrary structure like an ad-hoc network, although the amount of nodes will be greater (thousands for medium size cities).

Currently, ad-hoc networks are classified into two categories:

1. Mobile Ad-hoc Networks -MANET.
2. Sensor Ad-hoc Networks.

Although the proposed network has some common aspects with sensor networks, it differs from both:

- Nodes have no mobility.
- Communication is usually between nodes and the UC.
- Power is not a main priority.
- Nodes are prone to failure.
- They are densely deployed within the range area.
- There are few topological changes (on very few occasions a node is added or eliminated and radio range changes rarely occur).

Ad-hoc networks do not have any special nodes. However, for an application layer, the Utility Controller will have true a special node. We will use this property for optimizing network performance. To measure the relationship between these network performances and topology, we define a simple parameter - Medium Number of Hops a node needs to reach the Utility Controller. is related to medium access time from UC to a node, and to global polling time, also. We have estimated in various scenarios. The first one is shown in figure 2, and it assumes the following conditions:

1. Network nodes are uniformly distributed across the city, so we can define a density parameter.
2. All devices are indoors, so the radio range will be short and the same for all of them.
3. UC is located at center of the net.

We will refer to this topology as SR - Short Range Topology. We have proven that depends mainly on geometric parameters, and it can be computed approximately when $R_{GC} \gg R_{SR}$ as it follows:

$$\overline{NH} \approx \frac{2}{3} H = \frac{2}{3} \frac{R_{GC}}{R_{SR}} \tag{1}$$

Where

- R_{GC} - *Global Radius.* It is the radius of a circle that covers all nodes in the city or a significant number of them.

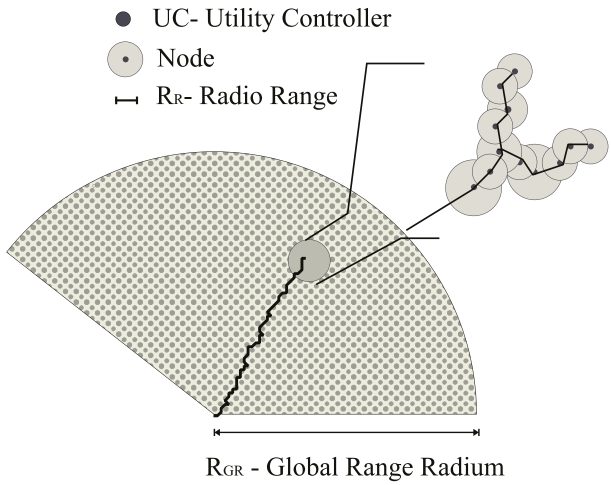

Fig. 2. SR - Short Range Topology.

– R_{SR} - *Short Radio range.* It is the medium value of short radio range of communication equipment.

$\overline{NH}$ is closely related to network topology. It enables computing network performances for several protocols and topologies. Minimizing $\overline{NH}$, will greatly enhance the completion time for the global reading service. This optimization is very important because global reading would probably be the major load service. Other functions like order and alarm transmission will also be enhanced. As $\overline{NH}$ decreases, transmission reliability will grow and service execution time will reduced. One way to enhance previous SR topology is to use those nodes that have much longer range than R_{SR} as a bridge to reach UC reducing intermediate hops. For example, equipment located outdoors may have a range up to ten times greater. Long range nodes could act as a long range subnetwork, able to connect any city area with the UC through fewer intermediate hops. Network protocols must enable message flow between short range nodes and the closest long range nodes in order to continue through the long-range subnetwork. In this way, the network is divided into different clusters within a main node that belongs to the long range subnet (see figure 3). We will refer to this topology as SR-LR (Short Range - Long range) architecture.

In this case, as usual, the number of short range nodes is significantly greater than the number of large range nodes. It can be shown that $\overline{NH}$ will be:

$$\overline{NH}_{SR-LR} \approx \overline{NH}_{SR} + \overline{NH}_{LR} = \frac{2}{3}\frac{R_{LR}}{R_{SR}} + \frac{2}{3}\frac{R_{GC}}{R_{LR}} \qquad (2)$$

Note that $\overline{NH}_{SR}$ does not depend on geometric parameters. It depends on radio transmission characteristics. However, $\overline{NH}$ depends on city geometry. A medium number of hops resulting from SR-LR architecture is significantly less

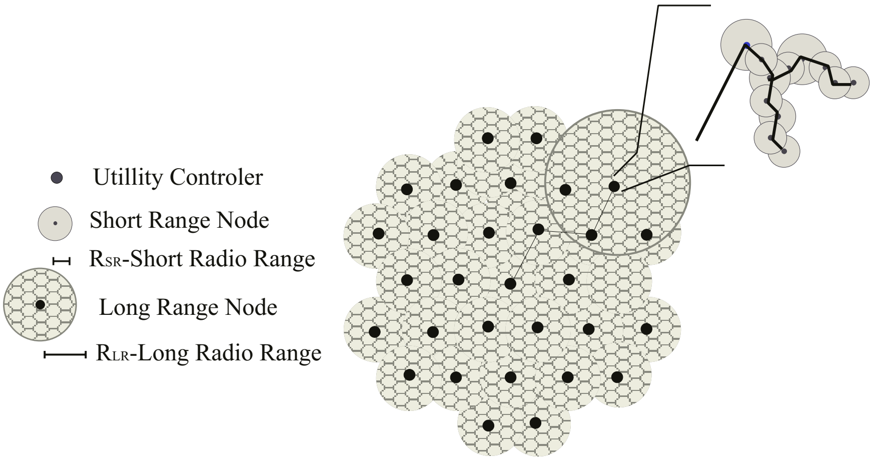

Fig. 3. SR-LR architecture.

than SR architecture. For example, in a medium-size, European city like Seville (Spain) 95% of the population is located inside of a circle of about $R_{GC} \approx 5Km$. Then:

$$\overline{NH}_{SR} \approx \frac{2}{3}\frac{5000}{1000} = 33.33 \qquad \overline{NH}_{SR} \approx \frac{2}{3}(\frac{1000}{100} + \frac{5000}{1000}) = 10 \tag{3}$$

$\overline{NH}$ is generally a geometric parameter that we have to relate to network or protocol characteristics. As is shown in figure 4, messages flow between the UC, and nodes. They can be described as a token moving through a spanning tree. Each node represents a hop, and its branches show some of the equipment inside its radio range. A first approximation of the number of transmissions may be derived from in a simple way, in terms of tokens moving in a spanning tree, under the following conditions:

1. There is only a token in the tree, so there are no collisions.
2. The token always jumps between parent and children without retransmission.
3. Tokens flow following minimum paths.

Then:

$$\overline{NT} = \overline{NH} \tag{4}$$

These conditions draw an ideal scenario:

– where a tree grows,
– radially without loops,
– all the nodes are in service,
– and there are no collisions or interference.

A first approch for real scenarios can be derived, supposing that p is the error rate of transmissions between nodes. In this way retransmissions will be necessary. We assume also, that p is uniform within the radio range and the same for all nodes in the network. Then, we have proved that the maximum value for medium number of transmissions is:

$$\overline{NT}_{MAX} = \frac{1}{1-p} \cdot \overline{NH} \tag{5}$$

Again $\overline{NT}$ depends on $\overline{NH}$ directly. So, finding the short paths to the UC sold be the basis of topology management and performance optimization. We have developed a custom algorithm to find these paths, but any other would be possible [9,10,11,12,13,14] (most of these are designed for point to point communication). However, ASCADA defines the UC at the application layer as the main node. It is often present as a transmitter or receiver. Topologies described previously make use of these properties to minimize multiple hops and create preference paths to the UC (based on a minimum hops criteria). Moreover, sensor networks usually select paths based either on power or quality link criteria [15,16]. Because a minimum number of transmissions is the objective, our selected criteria quantify these transmissions as close as possible to reality. Let i, j be two neighbor nodes. We define as *pseudo-range from node i to the UC through j*:

$$\rho_{ij} = d_j + \overline{nt_{ij}} \qquad d_j = \min\big(\rho_{jx}\big) \qquad \forall x \tag{6}$$

Where d_{ij} is the minimum pseudo-range between j and UC, and $\overline{nt_{ij}}$ is the medium number of transmissions between i and j, then, equation 6 assures the location of a short path following a reasonable link quality along it, and loop free.

3 Augmented Scada Application Optimization

Nodes always know the path to root through the parent node, but any selected algorithm must allow the root node (UC) to have an approximate image of current tree topology. This way, packets from nodes to the root do not contain information about routing. Conversely, messages from root to nodes generally use an explicit routing scheme, so packets must contain information about the path. Packet header size must be optimized because radio frames should be as short as possible to reduce the transmission error ratio, effective bandwidth, etc.

We propose in the paper to optimize application services in such a way so that message traffic and routing header size are minimal. In the following, we summarize all these application services:
AMR services:

- *AMR_Read*: Root sends a read message to a node to read data or check a meter.
- *AMR_Poll*: Root sends a message to read data from all meters in a node.
- *AMR_Collect*: Root sends a message again, to read data from all nodes of a subtree.

Telemetry services:

- *TLM_Polling.* This service initiates a periodic polling to a subset of nodes, located at distribution points. It creates and updates a table called Image Table with data from sensors which include a timestamp. This table may be accessed by custom primitives (*TLM_Read*).
- *TLM_Read.* Return data from image table with integrity information (timestamp).

Remote control services:

- *RC_Send.* Root sends a message with orders to control remote device (valves, breakers…). Message must be sequenced to avoid duplication and encrypted for secure operation. It may be necessary to notify a receipt message, and an order completion message.
- *AMR_Collect* service may be the most complex because time minimization and network overhead reduction are quite difficult to optimize simultaneously. There are many studies for polling optimization [17,18], but most of them use a well defined topology (rectangular, hexagonal…). The proposed network is a random network, we only suppose that nodes are uniformly deployed either as short-range nodes or long-range nodes. Execution time and network overheads can be evaluated approximately from topology parameters (see previous section). These values may be used to compare how different polling schemes may optimize network performance.

From the medium number of transmissions, we can compute execution time in an ideal context (equation 4) or in a more real one (equation 5). We can consider the collect message like a token moving along the branches from root to nodes, and the answer as another token returning back from nodes. The first approach is a simple collecting schedule consisting of polling each node from the root and, individually, waiting for the answers. Nodes do not have an application layer in that case, and there is only a single token moving in the tree. Moreover, let us consider an ideal scenario with the following conditions:

1. There is no interference or collisions.
2. Waiting time to medium access is zero.
3. Protocol stack computing time is negligible.
4. Only one frame is necessary to transmit all data from a node.
5. All tokens have exactly the same size.

An approach value of execution time, for an overall collect order, may be derived from a medium number of transmissions (equations 2,3,4), as:

$$CT_1 = \left(\mathrm{PTS} \cdot \mathrm{N} \cdot \overline{NT} + \mathrm{ATS} \cdot \mathrm{N} \cdot \overline{NT}\right) \cdot \frac{\mathrm{CharSize}}{\mathrm{BaudRate}} \tag{7}$$

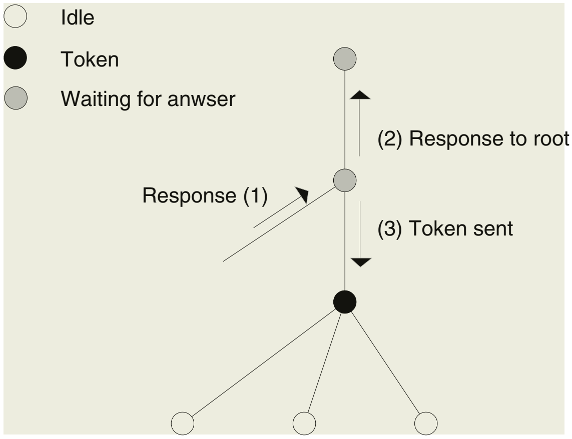

Fig. 4. Polling scheme 2.

Where

- CT_1 - Collect Time for schedule 1.
- N - Network nodes.
- PTS - Polling Token Size.
- ATS- Answer Token Size.

Second collecting algorithm is shown in figure 4. An application layer in all the nodes receives the order. This node passes the token to one of its children, and waits for a response. When it is received, (figure 4(1)) the node sends data to the root (figure 4(2)), and it passes the collect order to the next child (figure 4(3)). Only when there are no more children to be polled, will the node answer with its own data. Only one token is being passed between parent and children, so no path header information is necessary, and the answer token is simultaneously flowing to the root. To prevent a lost token, a timeout period guarantees the token passes to the next child.

Using the previous scenario, we can compute an approximate collecting time value for that schedule. Neglecting initial transmissions and second order effects, the majority of transmissions are caused by answer token passing. Polling token passing occurs simultaneously and so does not have any significant effect. In this way:

$$CT_2 = N \cdot \mathrm{ATS} \cdot \overline{NT} \cdot \frac{\mathrm{CharSize}}{\mathrm{BaudRate}} \tag{8}$$

The ratio between CT_2 and CT_1:

$$\frac{CT_2}{CT_1} = \frac{N \cdot \mathrm{ATS} \cdot \overline{NT}}{N \cdot \mathrm{PTS} \cdot \overline{NT} + N \cdot \mathrm{ATS} \cdot \overline{NT}} = \frac{1}{1 + \dfrac{PTS}{ATS}} \tag{9}$$

Usual PTS and ATS values allow a relative decrement of about 70 per cent. A greater reduction can be reached by using multiple tokens. If collisions or

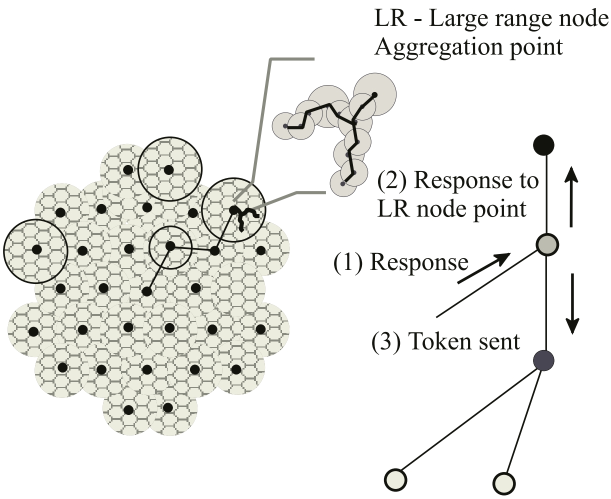

Fig. 5. Simultaneously collecting tokens: polling scheme 3.

interference are not considered, then collecting time decreases to a factor equal to the number of tokens. Obviously, collisions increase as the number of tokens increase. Many response tokens travel to the root node, creating an implosion problem [19] (which is greater near the root node). To avoid this, we propose a polling scheme based on Short range - Long Range topology. The root node must send polling tokens to each node in the LR subnet. Each LR node collects data from all their sub-trees, aggregating and storing data, without passing them to root. LR nodes only send data when the UC requests them. The node application protocol does not change significantly with respect to the second schedule, so this variant is fundamentally the same one for nodes. As it is shown in figure 5, polling tokens are passing simultaneously over different areas, so collision probability is low. Supposing that all the clusters have approximately the same number of nodes (long-range and short range densities are homogeneous), neglecting collisions and second order effects, then:

$$CT_3 = \Big(\frac{N_{SR}}{N_{LR}} \cdot \mathrm{ATS} \cdot \overline{NT}_{SR} + N \cdot \mathrm{ATS} \cdot \overline{NT}_{LR}\Big) \cdot \frac{\mathrm{CharSize}}{\mathrm{BaudRate}} \qquad (10)$$

And supposing that $N_{SR} \approx$ N, and applying (equation 2)

$$\frac{CT_3}{CT_1} = \frac{1}{1 + \dfrac{PTS}{ATS}} \cdot \frac{1}{1 + \dfrac{\overline{NT}_{SR}}{\overline{NT}_{LR}}} \qquad (11)$$

Using topology parameters for Seville, this scheme may reach $CT_3 = CT_1/10$.

4 Future Works

For the proposed network, we are currently working on network simulation, self-configuration and optimization of ASCADA services performance. For the latter, we are studying two complementary polling schedules: Avalanche token passing and data catching/pre-collecting. To reduce interference and to enhance polling order diffusion, each node sends two or more tokens to the children and responses are aggregated to send a single response to the root. Nodes must be selected in such a way that interference probability would be the least possible. Data catching and pre-collecting schemes use data validation to update stored data. Nodes save data from their children with a timestamp, so when a polling token arrives, the node may use this data without passing it to them. Catching update strategies, as periodically or predictive pre-collecting, must be designed to minimize data age and optimize network performance.

Execution times and network overload optimization are the main objectives of algorithms presented in this paper. Reliability however, must also be an important characteristic for a SCADA system.

Currently, we are working on reliability enhancement. Services such as order messaging and alarm transmissions can be critically affected by local failures, interference or collisions. Reliability must be present in all protocol layers, but especially in application and network layers. Some characteristics previously outlined may raise reliability in application layers: order sequencing, encrypted messages, etc. At the network layer, we are researching a routing algorithm able to find an alternate path when an error has occurred. We are testing a modification of the Fish Eye Routing and other algorithms [20,21].

Acknowledgment

All of this work has been made possible thanks to the project being jointly financed by the Spanish Government (*MYCT-Ministerio de Ciencia y Tecnología*) and private companies, in particular EMASESA and ISOTROL.

References

1. A. Bond. The water industry (automatic meter reading). *IEE Colloquium on 'Low Power Radio and Metering'. IEE, London, UK*, (Digest No.1994/060), 1994.
2. Philips M. Adams, B. Trends towards standard communications for metering. *Ninth International Conference on Metering and Tariffs for Energy Supply. IEE, London, UK*, (Conf. Publ No.462), 1999.
3. A.M. Fox. The business case for radio based amr. *IEE Colloquium on 'Low Power Radio and Metering'. IEE, London, UK*, (Digest No.1994/060), 1999.
4. Radford D. Mak, S. Design considerations for implementation of large scale automatic meter reading systems. *IEEE Transactions on Power Delivery*, 10(Digest No.1994/060), Jan 1995.
5. F.J. Molina, M.G. Gordillo, J. Luque, and J. Barros. Radio network architecture for automatic meter reading. *Conférence Internationale des Grandes Réseaux Électiques, Krakow POLAND*, 1999.

6. J. Broch, D. A. Maltz, D. B. Johnson, Y.C. Hu, and J. Jetcheva. A performance comparison of multi-hop wireless ad hoc network routing protocols. *MOBICOM, Dallas, TX*, Aug 1998.
7. E. Royer and C.K. Toh. A review of current routing protocols for ad hoc mobile wireless networks. *IEEE Personal Communications*, 6, Apr 1999.
8. S. Das, C. Perkins, and E. Royer. Performance comparison of two on-demand routing protocols for ad hoc networks. *IEEE, INFOCOM 2000*, 2000.
9. Z. J. Wang and J. Crowcroft. Analysis of shortest-path routing algorithms in a dynamic network environment. *ACM SIGCOMM*, 1992.
10. J. Behrens and J. J. Garcia-Luna-Aceves. Hierarchical routing using link vectors. *IEEE, INFOCOM 1998*, 1998.
11. J. J. Garcia-Luna-Aceves and M. Spohn. Source-tree routing in wireless networks. *Proc. IEEE ICNP 99, 7th ntl. Conference on Network Protocols, Toronto, Canada*, Oct 1999.
12. D. B. Johnson and D. A. Maltz. *Dynamic Source Routing in Ad Hoc Wireless Networks. In Mobile Computing.* Kluwer Academic Publishers, 1996.
13. C. E. Perkins and E. M. Royer. Ad-hoc on demand distance vector routing. *WMCSA'99, New Orleans, LA*, Feb 1999.
14. Couto and B Aguayo. Performance of multihop wireless networks: Shortest path is not enough. *Proceedings of the HotNets, Princeton, New Jersey*, Oct 2000.
15. S. Singh and C. S. Raghavendra. Power aware routing in mobile ad hoc networks. *Proceedings of MOBICOM*, 1998.
16. Chandrakasan A. Heinzelman, R. W. and H. Balakrishnan. Energy-eficient routingprotocols for wireless microsensor networks. *In Hawaii International Conference on System Sciences (HICSS '00)*, Jan 2000.
17. G.A. Cheston and S. T. Hedetniemi. Polling in tree networks. *Proc. Second West Coast Conf. Computing in Graph Theory*, 1983.
18. A. A. Rescigno. Optimal polling in communication networks. *Parallel and Distributed Systems, IEEE Transactions on*, 8, May 1997.
19. T. Imielinski and S. Goel. Querying and monitoring deeply networked collections of physical objects. *Proceedings of MobiDE'99, (Seattle, Washington)*, 8, Aug 1999.
20. Gerla M. Pei, G. and T. W. Chen. Fisheye state routing: A routing scheme for ad hoc wireless networks. *Proceedings of ICC 2000, New Orleans, LA*, 8, Jun 2000.
21. M. Spohn and J.J. Garcia-Luna-Aceves. Neighbourhood aware source routing. *Proc. of ACM Symposium on Mobile Ad Hoc Networking and Computing (MobiHOC '01)*, 2001.