Les Cahiers du GERAD ISSN: 0711–2440

Analytical performance analysis of a large-scale RF-mesh smart meter communication system

F. Malandra
B. Sansò

G–2015–46

May 2015

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Analytical performance analysis of a large-scale RF-mesh smart meter communication system

Filippo Malandra
Brunilde Sansò

GERAD & Polytechnique Montréal, Montréal (Québec) Canada, H3C 3A7

filippo.malandra@polymtl.ca
brunilde.sanso@polymtl.ca

May 2015

Les Cahiers du GERAD
G–2015–46

Abstract: Advanced meter infrastructures (AMIs) are now widespread and their importance within smart grid systems continues to increase with the advent of new applications. Performance analysis of the infrastructure is key to assess the limits of application deployment. However, due to the large-scale nature of AMI networks that are often composed of tens of thousands of nodes per collector, performance analysis is often carried out in contained experimental trials. To our knowledge, no thorough mathematical performance analysis of real-sized systems has been carried out so far. In this work, we present a model to analyze the performance of a large-scale RF-AMI system and show its application to large-scale real-case scenarios.

Key Words: AMI, smart grid, RF mesh, smart meter, multi-hop.

Acknowledgments: This work was supported by an ECO Energy Innovation Initiative grant from Natural Resources Canada.

1 Introduction

AMI systems are increasingly popular and massive worldwide installation of smart meter devices is currently ongoing. Several studies forecast an acceleration of this phenomenon in the following years, mainly driven by the large revenues expected from the use of smart meter applications within the smart grid market.

Several technologies can be adopted for AMI but the RF-mesh based system seems to be one of the most popular. RF-mesh systems are mainly used for remote reading, advanced metering and for some other applications, such as demand-response or load management, that do not have strong requirements in terms of bandwidth and delay. However, utilities, that have spent millions of dollars installing such a widespread communication infrastructure, may want to exploit it for other types of applications that, in some cases, may require shorter response time.

A fundamental question that arises in this context is how to be able to assess the limits of the installed infrastructure. Two methods can be though of: stochastic simulation and field trials. The difficulty of stochastic simulation lays on the size of the system, that may contain thousands of nodes, thus requiring specialized codes for parallel implementation and still would need a very large amount of resolution time [1]. On the other hand, field trials are limited, because they depend not only on the particular conditions of the households involved in the trial, but also on the time those trials are carried out. In fact, depending on the application, load conditions may influence the amount of communication that can be exchanged and load conditions greatly vary with time of day, month and season, among others. Moreover, the nature of field trials is observation-based, as, most of the time "what ifs" cannot be easily implemented.

There is, therefore, an important need for a flexible analytical model for large scale RF-smart meter performance evaluation that would allow not only to assess the current system, but also to do extended "what ifs" studies, evaluate the suitability of the system for future applications as well as help in the assessment of the features most needed in the evolution of the system. The object of this work is precisely to fill this gap by proposing an analytical characterization of large-scale RF-smart meter performance.

The document is structured as follows: Section 2 is a brief overview of the literature concerning the performance of wireless mesh networks, with a particular attention to smart grid applications; Section 3 presents the modelling of the system as well as some relevant performance indexes; Section 4 reports some results; Section 5 summarizes the conclusions of this work.

2 State of the art

An RF-based AMI system is a largely distributed wireless mesh network for which it is extremely difficult to put in place stochastic simulation results given its large scale. Nevertheless, some authors have attempted this approach (e.g., [1–4]). Others have dealt with real-time trial measurements. For instance, [5] reported some results obtained in a real AMI while [6] combined the simulation approach with real-time measurements. Our methodology is totally analytic and, in that respect, it is based on fundamental findings of wireless networks performance.

Interference among the links is a key issue to assess wireless mesh network performance. [7] defined a *connectivity graph* and associated to it a *conflict graph*, using the information on nodes, links and distances. The problem of maximum throughput was transformed into the search of the maximum independent set (a set of vertices that can transmit simultaneously), which is a NP-hard problem and which does not guarantee a solution. [8] defined a *contention* graph using flows rather than links to account for interference: the objective of the authors was to provide a *fair* access to the medium to different commodities. [9] tackled the problem of performance analysis in a multi-hop wireless network focusing on routing and scheduling. The authors modelled the problem as a graph-coloring one and assumed a simplified interference model in which each node cannot transmit and receive on the same channel at the same time. In our work, we took inspiration from this paper in what concerns the interference analysis, that is one of the steps needed to develop the analytic tool for performance evaluation.

The other issue that has a strong impact in the overall network performance is the effect of the MAC layer, that, in the case under study is ALOHA based. Performance of the ALOHA system has been well documented (e.g., [10, 11]). Our system differs from the others in literature because of the large number of nodes ($\approx 10^3$) and of the use of Frequency Hopping Spread Spectrum (FHSS) protocol to tackle interference (see Section 3 for further details).

Even though there are some simulation and field studies on AMI performance as well as a rich literature on mesh network performance, to the best of our knowledge this is the first time that a comprehensive analytical model is put in place to assess several performance parameters of a large-scale AMI system. Our analytical framework is particularly useful because it can be adapted to different types of scenarios and network features and it does not present the computational issues of very large scale systems simulation.

3 RF mesh communication system modelling

3.1 Main features

Many deployed AMI systems present proprietary features that are not easily disseminated to the public. So, in this paper, the case-study was a deployed system for which some data is publicly available. The communication system in object is a multi-hop wireless mesh network made of smart meters, routers and collectors with a current infrastructure of 1.5 million smart meters. The technology to connect the meters to data collectors is RF mesh and routers are used to extend coverage and increase connectivity. The communication channel is the unlicensed ISM band of $902 - 928$ MHz. The system uses FHSS, which is a technique particularly efficient against low spectrum interference coming from other devices transmitting on the same free band. The access to the medium is regulated by a synchronous ALOHA with time slots of duration $\tau = 0.7$ s. Devices are kept synchronized through the Network Time Protocol (NTP). Collectors are the only devices equipped with GPS receivers, whereas routers and smart meters are synchronized to their collector with NTP with an offset in the order of some milliseconds.

3.2 Topology definition

The topology definition in a large scale metering system is a challenging task. For this particular case, we used public information to extract the general areas where smart meters were installed, that happened to present the position of routers and collectors. Next, we calculated the GPS coordinates by means of Google Maps and Bing Maps application programming interfaces. To locate the smart meters, we assumed one device per home and we developed a script in Bing Maps that inputs public data, such as postal codes or list of municipality streets, and outputs the latitude and longitude of the network nodes. At the end of this process, as shown in Figure 1, we obtained the position of the nodes of the network. To define links, two different covering rays were assumed: a value between 0.15 to 0.5 km for smart meters and of up to 2 km for routers and collectors. The variation of the maximum transmission range of smart meters is due to different propagation conditions in urban, rural and suburban areas. Moreover, routers and collectors achieve extended coverage with directive antennas and top-building installation. Routers and collectors have higher capacity links (19.2 kbps) with respect to smart meters (9.6 kbps), as shown in Figure 1, which is a sketch of a simple topology.

3.3 Shortest paths

To start the analysis, a static shortest path routing based on distance as a metric was assumed, but future works will include other types of routing. The problem can be seen as a large multi-commodity flow since different streams of communication coexist in the system. There is one commodity from each smart meter to the collector and vice-versa.

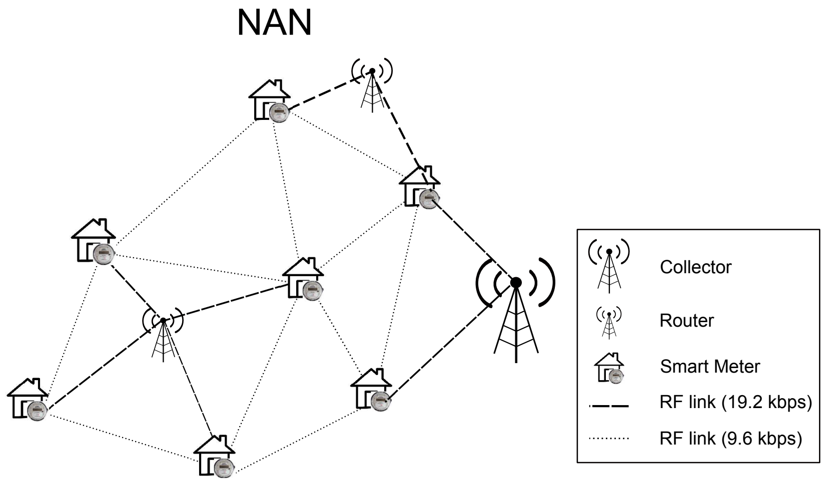

Figure 1: Topology definition

3.4 Traffic characterization

Let $G(V, E)$ characterize the graph representing the AMI system, where V and E are the sets of nodes and links of the RF network. The set of nodes V is composed by the union of the set of smart meters M, the set of routers R and the set of collectors C. Let λ_{up} be the mean traffic from each smart meter to the collector (uplink) and λ_{down} the mean traffic from the collector to a single smart meter (downlink). In both cases, Poisson arrivals are assumed so that the aggregation of traffic streams at each node is also Poisson with a mean value equal to the sum of all sub-streams mean values.

Let λ_i be the transmission rate of generic device i. In order to characterize λ_i, the routing behaviour of each device must be analyzed. Intermediate nodes are in charge of transmitting packets from the origin to the destination of the shortest paths they belong to. For this purpose, we introduce ξ_i, the number of shortest paths that contain node i as an intermediate node. Then, a given smart meter $i \in M$ transmits its own packets to the collector at a rate λ_{up}, packets from ξ_i smart meters to the collector at a rate $\xi_i\lambda_{up}$ and finally packets from the collector to the ξ_i nodes at a rate $\xi_i\lambda_{down}$. Routers do not generate traffic so a given device $i \in R$ is in charge of transmitting packets of ξ_i different streams in both uplink and downlink directions. On the other hand, collectors transmit to each of the $|M|$ smart meters with a rate λ_{down}. These considerations can be summarized as follows:

$$\lambda_i = \begin{cases} \xi_i(\lambda_{up} + \lambda_{down}) + \lambda_{up}, & \text{if } i \text{ is a smart meter} \\ \xi_i(\lambda_{up} + \lambda_{down}), & \text{if } i \text{ is a router} \\ |M|\lambda_{down}, & \text{if } i \text{ is a collector} \end{cases} \tag{1}$$

The actual transmission rate also depends on the number of packet retransmissions caused by collisions. In particular, if N_i represents the average number of retransmissions at node i, the actual transmission rate can be defined as $\tilde{\lambda_i} = N_i\lambda_i$. The average number of retransmissions and the link with the collision probability is deeply analyzed in Section 3.5.

3.5 Probability of collision

One of the main issues to a wireless communication system is interference. If we assume that the interfering ray equals the covering ray, every node's neighbours are its possible interferers. If we neglect for the moment FHSS, we can state that there is a collision at node i when at least one of its neighbours attempts to transmit during the same time slot.

When a collision is experienced, the involved packets have to be retransmitted. The average number of times a packet at node i is retransmitted, N_i, is related to the average collision probability by:

$$N_i = \frac{1}{1-p_i} \tag{2}$$

As stated in the previous section, the transmission rate of a set of nodes has a Poisson distribution with the sum of all λ_i as mean value. Given I_i the set of interferers of node i, the probability that none of the nodes $j \in I_i$ transmits in a time-slot of duration τ is:

$$P(X_{I_i} = 0) = e^{-\tau \sum\limits_{j \in I_i} \tilde{\lambda_j}} = e^{-\tau \sum\limits_{j \in I_i} \lambda_j N_j} = e^{-\tau \sum\limits_{j \in I_i} \frac{\lambda_j}{1-p_j}} \tag{3}$$

where X_{I_i} is the number of nodes in I_i that transmit during that time-slot. Then, the probability that collisions occur is:

$$p_i = P(X_{I_i} > 0) = 1 - P(X_{I_i} = 0) = 1 - e^{-\tau \sum\limits_{j \in I_i} \frac{\lambda_j}{1-p_j}} \tag{4}$$

If we want to take into account FHSS, the definition of collision changes: a collision occurs at node i when at least one of its neighbours is transmitting on the same channel as i. In our first FHSS modelling, we consider that each node randomly decides the transmission channel among the Q available. Let us consider the event that at least one of the k neighbours of i chooses the same channel as i. This is the complement of the event in which all the k nodes choose different channels than i; therefore, the probability that at least one of the $k = \mid I_i \mid$ nodes chooses the same transmission channel as i is given by:

$$p^{(k)} = 1 - \left(1 - \frac{1}{Q}\right)^k \tag{5}$$

As a consequence, a collision occurs when there is at least one transmitting neighbour of node i using the same channel as i is calculated as follows:

$$p_i = \sum_{g=1}^{+\infty} P(X_{I_i} = g) p^{(g)} = \sum_{g=1}^{+\infty} \left(1 - \left(1 - \frac{1}{Q}\right)^g\right) \frac{\left(\tau \sum\limits_{j \in I_i} \frac{\lambda_j}{1-p_j}\right)^g}{g!} e^{-\tau \sum\limits_{j \in I_i} \frac{\lambda_j}{1-p_j}} \tag{6}$$

This is a so-called *fixed point equation*. The problem of finding the values of p_i (for $i = 1 \ldots |V|$) that solve the non-linear system (6) is equivalent to solving the following optimization model:

$$\min_{\boldsymbol{p}} \|f(\boldsymbol{p})\|_2^2 \tag{7}$$

$$s.t. : \begin{cases} p_i \geq 0 & \forall p_i \in \boldsymbol{p} \\ p_i < 1 & \forall p_i \in \boldsymbol{p} \end{cases} \tag{8}$$

where:

$$\boldsymbol{p} = \begin{bmatrix} p_1 \\ \vdots \\ p_{|V|} \end{bmatrix} \quad f(\boldsymbol{p}) = \begin{bmatrix} f_1(\boldsymbol{p}) \\ \vdots \\ f_{|V|}(\boldsymbol{p}) \end{bmatrix} f_i(\boldsymbol{p}) = \sum_{g=1}^{+\infty} \left(1 - \left(1 - \frac{1}{Q}\right)^g\right) \frac{\left(\tau \sum\limits_{j \in I_i} \frac{\lambda_j}{1-p_j}\right)^g}{g!} e^{-\tau \sum\limits_{j \in I_i} \frac{\lambda_j}{1-p_j}} - p_i$$

One can note that, when the number of channels is 1, (6) is equivalent to (4) since $\sum\limits_{n=1}^{+\infty} \frac{a^n}{n!} = e^a - 1$. A quantitative solution for the fixed point system of equations (4) was obtained using the least square minimization model (7) and numerical results are presented in Section 4. The performance analysis carried out in the rest of this work is based on the vector $\boldsymbol{p}$, solution of the optimization model (7). At this time, numerical results are obtained using the simplified equation (4), neglecting in the first phase the presence of FHSS.

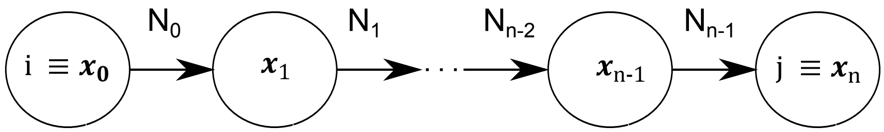

Figure 2: Scheme of the n-hop path from node i to node j

3.6 Delay

A basic parameter in the analysis of the performance of a telecommunication network is delay. In a multi-hop random access system with small packets, such as Time-Slotted ALOHA, it is a standard practice to let the time slot duration include all other types of delay a packet can encounter. In the rest of the discussion, propagation, transmission and processing delays are included in the $0.7s$ time slot. In this first analysis, the queueing delay is considered negligible since we are dealing with very low traffic rates.

In the light of these considerations, we calculate the average delay based on the number of hops a packet makes in order to reach its destination only. The delay of a generic n-hop path such as the one displayed in Figure 2 can be calculated as follows:

$$d_{ij} = \tau \sum_{k=0}^{n-1} N_k \tag{9}$$

Note that in (9), delay depends exclusively on the length of the shortest path as well as the probability of retransmission of the system. If queueing and processing delays were not negligible, (9) should be modified by adding the appropriate M/M/1 delay modelling terms.

We account for two different types of delay:

d_i^u *uplink delay*, the time necessary for a packet generated by a smart meter to get to the collector;
d_i^d *downlink delay*, the elapsed time for a packet to travel from the collector to a smart meter.

4 Results

This section presents some numerical results obtained with the model introduced in the previous sections. We chose a typical rural area with an extension of $240\,km^2$ with 3415 nodes (1 collector, 114 routers and 3300 smart meters). Moreover we defined 32 different traffic scenarios, based on message exchange rate on the uplink, λ_{up}, and the downlink, λ_{down}, assuming that all the meters had the same message frequency (see Table 1). The reader should be aware that such an assumption was adopted to ease the interpretation of results, but it could be easily relaxed within our modelling framework.

Results were obtained using MATLAB, with an Intel(R) quad Core(TM) i7-3770 CPU @ 3.40GHz processor. The average computational time was 1684.86 s with the maximum value of 7250 s obtained in the 3^{rd} scenario and the minimum of 342.52 s in the 32^{nd}.

Table 1: Traffic scenario ID according λ_{up} and λ_{down}

$\frac{1}{\lambda_{down}}$[h]	$\frac{1}{\lambda_{up}}$[h]							
	0.5	**1**	**1.5**	**2**	**2.5**	**3**	**3.5**	**4**
1	1	5	9	13	17	21	25	29
2	2	6	10	14	18	22	26	30
3	3	7	11	15	19	23	27	31
4	4	8	12	16	20	24	28	32

4.1 Collision probability

The solution of the fixed point set of equations (4) provided the probabilities of collision that each node is subject to when it attempts to transmit a packet. Figure 3 represents some statistics of the collision probability in different scenarios: the maximum and the average values related to each kind of device are reported. The black curves with triangles are related to routers. The maximum is, in many scenarios, close to 1 while the average goes from 0.35 in the first scenario to less than 0.1 in the last scenario. The grey curves with crosses refer to the maximum (continuous line) and the average (dotted line) values related to the smart meters. In this case, both maximum and average collision probabilities are considerably lower than those of routers. We can also remark that the average collision probability for smart meters is in all scenarios below 10%. When it comes to the collector, we can see that there is only one curve since we have a unique device of this type in each scenario. The dashed-dotted black line with circles shows that the performance of the collector is in between that of smart meters and routers: its collision probability is not excessively high (it is above 7% only in the five first scenarios).

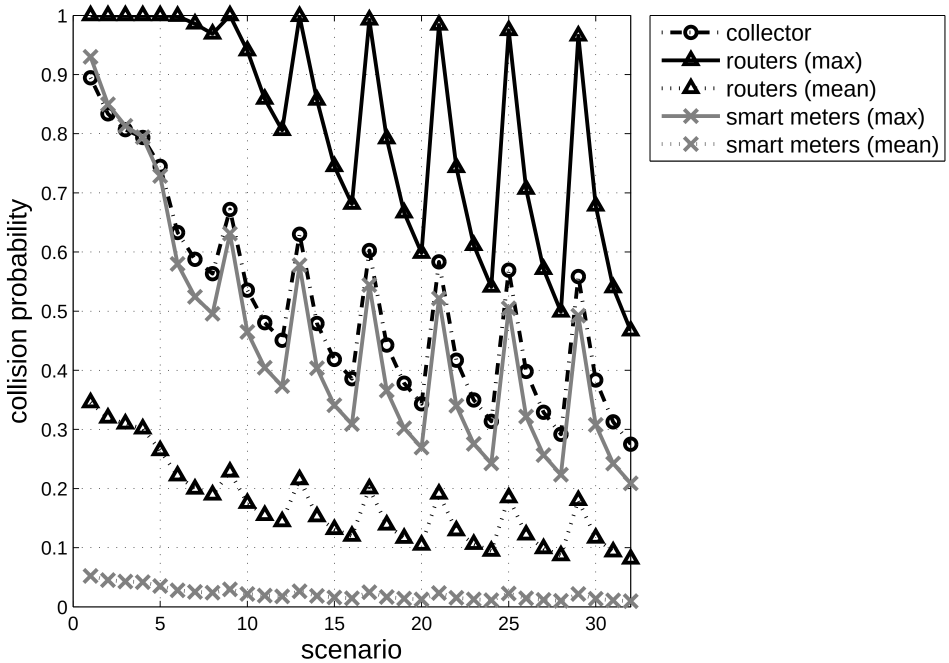

Figure 3: Statistics of the probability of collision of different devices in different scenarios

4.2 Delay profile

Provided that we calculated the delay as the sum of the number of retransmissions in each hop (3.6) and that the average number of retransmissions is inversely proportional to the probability of collision (2), we do not show the trend of delays in the system since it is very similar to that for the probability of collision presented in the previous section.

4.3 Survival function: application feasibility

An interesting insight on the performance of our system can be brought by the notion of the percentage of smart-meter nodes presenting an uplink or downlink delay above a certain threshold. In the smart grid context, it is important that the received information be up-to-date, and this is true only if it is received within a short delay. In order to carry on this analysis, we use the concept of survival function. Let Y be a random variable, its survival function is $f(y) = P(Y > y)$ with $y \in [0, \infty]$.

Let D_u and D_d be the discrete variables related to uplink and downlink delay respectively, we define two survival functions as follows:

$$f_u(\gamma) = P(D_u > \gamma) \tag{10}$$

$$f_d(\gamma) = P(D_d > \gamma) \tag{11}$$

The notion of survival function allows us to assess the feasibility of smart grid applications in RF-mesh systems.

Let α and β be two generic load management applications with the following requirements: the collector transmits 1 pkt/4 h in average, each smart meter transmits 1 pkt/2 h in average; a packet is considered old by the collector if received after a delay (in uplink) of 20 s. α can tolerate a maximum of 20% of packets to be old, whereas β can tolerate only 10%. The traffic transmission rates suggest that our example corresponds to scenario 16 ($\frac{1}{\lambda_{down}} = 4$ h, $\frac{1}{\lambda_{up}} = 2$ h). We are analyzing the delay at the collector side, that is to say the uplink delay. Figure 4 shows the survival function of delay in uplink in scenario number 16. In the same graph we plot the points (20 s, 0.2) and (20 s, 0.1), relative to the requirements for α and β. The first coordinate of these points represents the delay threshold (γ), while the second represents the maximum tolerated probability $P(D_u > \gamma)$. If the point is above the curve, as for α, it means that the requirements of its associated application are within the possibilities of the system: the application is considered feasible. On the other hand, if the point is below the curve, as for β, it means that the associated application requirements cannot be met by our system: the application is not feasible.

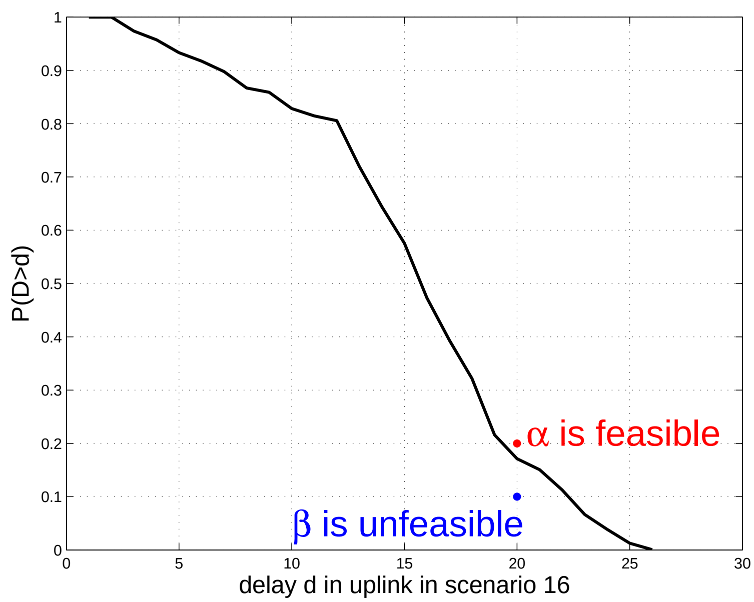

Figure 4: Survival function of uplink delay in scenario 16

5 Conclusions

This work presents, to our knowledge, a first analytical model for the overall performance evaluation of an AMI RF-mesh based system. The model becomes an important tool since it bypasses the difficulty of having to perform costly stochastic simulations for a network containing thousands of nodes, while enhancing the quality of information that can be extracted from field trials. The tool allows different types of what-if studies to assess the applicability of forthcoming applications, which is quite important for utilities.

We have not just developed a model, but also defined a framework to assess AMI RF-mesh performance. This framework is based on the type of routing, the interference model, the traffic analysis and the type of measures that must be taken into account to assess network performance, such as delay, collision probability, critical nodes and a survival function particularly useful to assess the feasibility of the deployment of new applications. This model could also be valuable in the design phase of a real RF-mesh system, to size the network and choose the number and location of routers and collectors to be installed.

We constructed a case study extracting publicly available information. For such a case, even though the conclusions were expected, it was possible to quantify them with the tool. The traffic rate (λ_{up} and λ_{down}) considerably affected the performance of the network and the shortest path routing was not very efficient. The model, however, can easily incorporate any other type of routing. Also expected was the fact that routers are the critical elements of the system, as the percentage of routers in critical conditions is much higher than for the other devices.

It is clear, however, that this model represents a first step towards a more comprehensive engine that would portray in greater detail some of the protocols and RF issues. In particular, a better representation of FHSS features is currently ongoing. Also, dynamic routing algorithms (i.e. RPL and AODV) are being integrated into the model. Despite the lack of those modelling details, this first analytic approach can give different points of view in the performance analysis of RF-mesh systems and it can be very useful to overcome the limits of the current approaches.

Finally, it is clear that in order to carry out real-time or quasi-real-time applications, the current infrastructure needs to be upgraded or must co-exist with some off-loading network such as the Internet.

References

[1] B. Lichtensteiger, B. Bjelajac, C. Muller, and C. Wietfeld, Rf mesh systems for smart metering: System architecture and performance, in Smart Grid Communications, 2010 First IEEE International Conference, 2010, pp. 379–384.

[2] G. Iyer, P. Agrawal, E. Monnerie, and R. Cardozo, Performance analysis of wireless mesh routing protocols for smart utility networks, in Smart Grid Communications, 2011 IEEE International Conference, 2011, pp. 114–119.

[3] A. Patel, J. Aparicio, N. Tas, M. Loiacono, and J. Rosca, Assessing communications technology options for smart grid applications, in Smart Grid Communications, 2011 IEEE International Conference, 2011, pp. 126–131.

[4] A. AlMajali, A. Viswanathan, and C. Neuman, Analyzing resiliency of the smart grid communication architectures under cyber attack, in 5th Workshop on Cyber Security Experimentation and Test, 2012.

[5] S. Céspedes, A. Cárdenas, and T. Iwao, Comparison of data forwarding mechanisms for ami networks, in Innovative Smart Grid Technologies (ISGT), 2012 IEEE PES, 2012, pp. 1–8.

[6] P. Kulkarni, S. Gormus, Z. Fan, and F. Ramos, Ami mesh networks: A practical solution and its performance evaluation, Smart Grid, IEEE Transactions, 3(3), 1469–1481, 2012.

[7] K. Jain, J. Padhye, V. Padmanabhan, and L. Qiu, Impact of interference on multi-hop wireless network performance, Wireless Networks, 11(4), 471–487, 2005.

[8] T. Nandagopal, T.E. Kim, X. Gao, and V. Bharghavan, Achieving MAC layer fairness in wireless packet networks, in Proceedings of the 6th Annual International Conference on Mobile Computing and Networking, ACM, 2000, pp. 87–98.

[9] M. Kodialam and T. Nandagopal, Characterizing achievable rates in multi-hop wireless networks: the joint routing and scheduling problem, in Proceedings of the 9th Annual International Conference on Mobile Computing and Networking, ACM, 2003, pp. 42–54.

[10] M. Médard and A.J. Goldsmith, Capacity of time-slotted ALOHA systems, in Wireless Communications and Networking Conference, 1999. WCNC. 1999 IEEE, pp. 490–494.

[11] M. Kaynia and N. Jindal, Performance of ALOHA and CSMA in spatially distributed wireless networks, in Communications, 2008. ICC'08. IEEE International Conference, pp. 1108–1112.