

A Stochastic Approach for Modeling Message Dissemination in Opportunistic Networks

Javier Orozco · Rodrigo Santos · Sergio
F. Ochoa · Roc Meseguer

Received: date / Accepted: date

Abstract The opportunistic networks have shown to be an interesting communication infrastructure for supporting mobile collaborative systems in several application scenarios like disaster relief efforts, hospital work and tourism. The message delivery strategies used in these networks affect in different ways the performance and energy consumption of these systems. Therefore, it is important to understand the message dissemination process in opportunistic networks to envision the impact that it will have on mobile collaborative systems. Most approaches used to study the message dissemination in these networks use simulations or empirical tests. Although useful, the first one requires an important modeling effort, and the second one involves high design and experimentation costs. This limits our capability to evaluate alternatives for message dissemination and choose the most appropriate one according to the features of the application scenario. In order to help address this limitation, we propose a stochastic approach for modeling message dissemination in opportunistic networks. This approach involves a low modeling and usage effort that allows designers not only to compare several dissemination strategies, but also design particular ones using an iterative process. The usability and usefulness of this approach is illustrated through the modeling and analysis of the dissemination process of two well-known message routing strategies.

Javier Orozco · Rodrigo Santos
Electrical Engineering and Computers Department
Universidad Nacional del Sur
E-mail: ieorozco@criba.edu.ar, ierms@criba.edu.ar

Sergio F. Ochoa
Computer Science Department
Universidad de Chile
E-mail: sochoa@dcc.uchile.cl

Roc Meseguer
Department of Computer Architecture
Universitat Politècnica de Catalunya - BarcelonaTech
E-mail: meseguer@ac.upc.edu

Although the results are still preliminary, they show the feasibility of using this proposal, and opens several opportunities for studying these networks and improving the mobile applications that use them.

Keywords Opportunistic networks · Message dissemination · Dissemination cost · Modeling approach · Analytical model

1 Introduction

The widespread adoption of mobile devices with several wireless communication capacities has changed the way in which people interact with other people, the environment and remote resources. Several mobile and ubiquitous applications running on these devices use opportunistic networks (oppnets) to support the activities of mobile users; for instance, firefighters performing search and rescue activities [16,19,21], nurses and physicians doing hospital work [5] or tourists looking for attractions when visiting a new place [20]. The oppnets are mobile peer-to-peer networks that combine the capabilities from both, mobile ad hoc and delay tolerant networks. In these networks the communication opportunities are intermittent, therefore an end-to-end path between the source and the destination may never exist. The modeling of the message dissemination process allows software designers to envision the impact of such a process on the performance and energy consumption of mobile applications, and thus determining the best option to deliver messages before writing code.

Provided these applications are usually intensive in terms of message exchange, the designers have to deal with the trade-off between energy consumption and message dissemination speed. Typically, the higher the dissemination speed the higher the amount of resources required in such a process; particularly energy. Therefore, the analysis for determining the best message dissemination strategy requires to consider these two aspects.

Most approaches used to study the message dissemination in oppnets involve simulations or empirical experiments, which have shown to be useful, but also impose some restrictions that limits our capability to understand this process and envision its impact on the applications that use it. Particularly, the simulations involve an important effort to implement and validate each simulated scenario. Moreover, the cost of tuning the models used in the simulations is also considerable, which jeopardizes the possibility to design particular dissemination strategies in an iterative way; i.e., improving the models based on the previous results.

Concerning the empirical tests, they provide accurate results, but involve important efforts for designing and running the experiments. The feasibility to use this approach decreases when increases the number of participants. Moreover, the empirical tests are difficult to replicate, which limits our capability to perform an iterative design process.

Contrarily, the modeling of message dissemination in oppnets requires that designers can iteratively tune the network and process features to address the

trade-off between dissemination speed and energy consumption. Therefore, these modeling approaches limit the capability to perform this interactive design, and compare the dissemination alternatives.

By trying to contribute addressing this limitation, we present a stochastic approach to model the message dissemination in oppnets. The proposal involves a modeling and usage effort that is considerable minor than performing simulations or empirical experiments, and it allows designers to perform an interactive modeling. This approach uses the Markov Chain theory to represent the dissemination process, and thus it produces an analytical propagation model of an oppnet. Each model represents a particular dissemination strategy that could be tuned or compared with other strategies involving a small effort. Once created and validated the model of a particular dissemination strategy, it can be reused not only by designers of mobile applications, but also by researchers as an instrument that helps them study these networks.

In order to illustrate the usability and usefulness of this approach, two well-known routing strategies were modeled and their performance in various scenarios were calculated using the analytical process, and considering message dissemination speed and energy consumption. This validation process shows not only the feasibility to use this proposal, but also the advantages that using it has in terms of modeling effort and flexibility. Recognizing that the communication scenarios used in this work are not enough to generalize the results and more evaluation is required, the current outcomes are highly encouraging and open several opportunities for studying these networks in a more simple and interactive way.

Next section presents and discusses the related work. Section 3 presents the proposed modeling approach for representing the message dissemination and energy consumption in oppnets. Section 4 illustrates the usability of the proposal by instantiating the network model to capture its behavior when an epidemic or a spray-and-wait routing strategy is used for the message dissemination. Section 5 presents and discusses the evaluation results, as a way to show the usefulness of this proposal. Finally, Section 7 presents the conclusions and the future work.

2 Related Work

Some recent works raise different considerations on the way in which the inter-meeting time between nodes (for message passing) can be modeled in opportunistic networks. In particular several stochastic distributions are analyzed and compared with data obtained empirically in real scenarios. The use of random walk distributions with exponential meeting times is considered a suitable model for independent nodes moving in close areas, while power-law distributions are used in open areas [3, 17, 24].

In [12] and [29] the authors analyze different alternatives of epidemic routing to improve the overall performance of a mobile ad hoc network. They use a Markov model to represent the message propagation and a Markov Chain

model to describe the evolution of the system over time. However, this proposal is based on the probability density function, which is specific for the configurations used in their studies.

In [22] the use of an opportunistic network is analyzed as communication support of a mobile collaborative application, and the first concepts of time constraints are introduced. In [23] there is an analysis of real-time message traffic for the case of FIFO scheduling at the gateway without priorities.

In [13] the Opportunistic Network Environment simulator is introduced. This tool was designed for evaluating routing and application protocols on these networks. The simulator provides a framework for implementing routing and application protocols based on different network interfaces, for example Bluetooth or Wi-Fi.

Concerning the strategies reported in the literature to evaluate the behavior of Delay Tolerant Networks (DTNs), they are mainly based on simulations or empirical studies. In [11] various of these strategies are presented and discussed. As mentioned before, they are time consuming and have low flexibility to explore the message dissemination in an iterative way. Therefore, an analytic approach would be recommended to address this challenge.

Several papers report analytic studies of oppnets and DTNs. In [30], a model of epidemic routing is introduced based on ordinary differential equations derived as limits of Markovian models. The proposal calculates the expected delay and the number of message copies (i.e., resources limitations), but it does not consider the energy consumption involved in the dissemination process.

An analytical model, also based on Markov chains, was proposed in [2] for evaluating a single copy forwarding strategy that follows an opportunistic social-aware dissemination. The model considers the number of hops needed for a message to reach the destination, and also the expected transmission delay. The nodes mobility follows a social behavior, i.e., some users may cluster and move together, and others may never get in touch with each other. Although this proposal is interesting, it does not analyze the energy consumption involved in the message transmission.

Similarly, Spyropoulos et al. introduce an analytical model to determine the expected number of hops and the expected delay of the messages when they are delivered in an opportunistic social-aware fashion [25]. As in the previous works, the model is based on human behavior and there is no evaluation of the energy consumption of the network. This proposal is evaluated using synthetic and real mobility traces.

In [1] the authors introduce a Markov model to represent the data-dissemination in stationary regimes. The model is used to determine convergence towards stationary regimes instead of evaluating the network performance.

Whitbeck et al. propose a model for Epidemic propagation on edge-Markovian dynamic graphs, which capture the correlation between successive connectivity graphs [27]. This proposal analyzes the impact of the bundle size and node mobility in the propagation delay. In [10] the authors studied the

performance of message propagation in people-centric opportunistic networks using Random Way Point and Random Direction mobility models. Unfortunately few proposals address the modeling of energy consumption in oppnets and DTNs. In this sense, Wang et al. [26] model the contact and inter-contact time, and validate the model with real traces. Using this information they evaluate the trade-off between the energy consumed in the search for neighbors, the probability of finding them and the frequency with which the process is performed. However, this energy analysis does not consider the whole transmission process.

In [28] the authors analyzed the message response time considering energy consumption restrictions related to the transmission range and sojourn time of the message in the nodes. The message propagation model used in this work is different to the one we are proposing, therefore it is not possible to compare both approaches.

Neglia and Zhang present a first attempt to study analytically the trade-off between delivery delay and resource consumption for epidemic routing in DTNs [18]. The authors computed both, the average number of message copies and the average delay in the transmission. The energy analysis left out the device discovery protocol, which is relevant in oppnets.

These previous works show that the stated problem is still open. In order to help address it, the next section describes the analytical approach proposed for modeling the message dissemination in oppnets. This approach introduces a Continuous Time Markov Chain (CTMC) analysis to determine the process performance, which is obtained by solving a set of differential equations using the tools provided by Markov calculus. Particularly, mathematical software tools, like Octave or Matlab, or even by hand solving the Laplace Transform of the differential equations. Moreover, this proposal also models the energy consumption of the whole transmission process.

3 The Stochastic Modeling Approach

The performance of the oppnet depends on several factors, such as the nodes mobility, the size of the application area, the communication range and the number of network nodes. Modeling the message dissemination in oppnets is complex since it requires to consider these factors and the relationships among them. The network behavior would also be affected by the particular layout of the physical area where the oppnet is deployed. All these factors determine the inter-meeting time probability distribution of nodes. As messages are assumed to be exchanged completely when two nodes meet, the message propagation is tied to the inter-meeting time distribution.

Considering several previous works, we assume a random walk distribution for the nodes mobility, which produces an exponential distribution in the inter-meeting times [4, 6, 14]. Based on this assumption, the message transmission can be modeled as a renewal process with Poisson distribution, characterized by a unique parameter λ that measures the nodes encounters rate. In what fol-

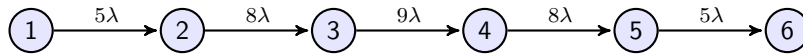


Fig. 1 Markov chain model for an oppnet of six nodes

lows, we model the *message dissemination* and the *energy consumption* based on the previous assumption. These models can be instantiated to address specific communication scenarios reusing thus the knowledge embedded on them.

3.1 Message transmission model

The message transmission follows a *birth process* that can be modeled as a CTMC. This is a mathematical approach widely used to study different communication models where sojourn times, in the different states, have an exponential distribution with the well-known memoryless property (related to the Markov processes). Each state in the CTMC represents the number of message copies present in the network. The Markov chain has a *source* node and also a *destination* one that is represented by an absorbing state. Figure 1 shows a schematic model for a six nodes network that use the classic epidemic routing strategy for message dissemination, i.e., every node holding the message is able to pass it on to other node, whether or not is the *destination*.

The *source* node is represented as the first state in the Markov chain. In this case, the source node may pass the message to any of the other five nodes in the network, one of them being the *destination* one (i.e., the absorbing state). When the message is transmitted to the second node in the chain, there are two copies and still four nodes to reach. In this case the transition probability is doubled. With the third copy, there are three nodes with possibilities of meeting the fourth node, and three nodes left without the message in the network. When the fourth copy has been transmitted, the transition probability is reduced, because even if there are four nodes with probability of copying the message, only two nodes are left without a copy. At the end of this process, there will be five nodes with a message copy, and only one node left without it.

Three metrics are used to understand the message dissemination process: the mean time to absorption (*MTTA*) that reflects the average message delay, the number of message copies present in the network (m_c) at the absorption time, and the average energy consumed (E_m).

The *MTTA* is used to determine how long a message should be alive consuming memory and energy in the nodes while it tries to reach the destination node. Provided that the behavior of the network is stochastic, it is possible to set a certain period of time in which there is a high probability that the message is delivered to the destination node. The (m_c) represents the amount of resources used for the message transmission, which is also related to the energy consumption involved in such a process (E_m). With the first two parameters it is possible to configure the lifetime of the messages allowing them

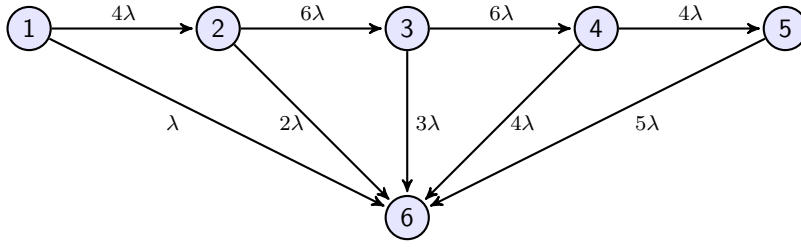


Fig. 2 CTMC with absorbing state

to be eliminated once this time is elapsed. The third parameter is useful for evaluating the energy propagation cost of the transmission process.

The CTMC described before constitutes a stochastic process. In order to compute the $MTTA$ and m_c , it is necessary to calculate the probability density function for each state in the CTMC, by using the differential equations that describe the chain stochastic behavior:

$$\dot{\pi} = \pi Q \quad (3.1)$$

In this equation, π is a vector where each element π_i is the probability density function for state i , and Q is the transition matrix. This matrix is built from the Markov chain and it represents the transitions among all the states in the chain. The variable q_{ij} represents the rate at which the process may move from state i to state j , where q_{ii} is the sum of all the transitions rates.

The set of linear differential equations presented in (3.1) can be solved in several ways. In particular, the analytical solution can be reached using the Laplace Transform (LT). However, the LT solution may have numerical problems for a relatively small number of network nodes. In order to avoid this, the differential equations can be solved using numerical solutions like the one proposed by the ODE45 algorithm [15]. The next equation computes m_c :

$$m_c(t) = \sum_{i=1}^N i\pi_i(t) \quad (3.2)$$

The expected transmission delay from the *source* node to *destination* one is computed by analyzing the behavior of the CTMC. The *destination* node acts as an *absorbing* state. In fact, once the message gets into the *destination* node, such a node will no longer propagate the message.

The previous CTMC is redrawn in Figure 2 to show the transitions to the absorbing state. Even if the message has reached the *destination* node, it may continue propagating copies to other nodes until all of them have a copy of the message.

The $MTTA$ is obtained from the cumulative distribution function (*cdf*) that is calculated by excluding the *absorbing* node from the Markov chain.

The modified transition matrix is denoted as \widehat{Q} . The *cumulative distribution function* for each state is computed by solving the following set of equations:

$$\dot{\mathbf{L}}(t) = \mathbf{L}(t)\widehat{Q} + \pi(0) \quad (3.3)$$

The time spent before absorption can be calculated by taking the limit $\lim_{t \rightarrow \infty} \mathbf{L}(t)$. As the equations are restricted to the non-absorbing states, the limit can be applied on both sides of (3.3) to obtain the following set of linear equations:

$$\mathbf{L}(\infty)\widehat{Q} = -\pi(0) \quad (3.4)$$

$$MTTA = \sum_{i=1}^N L_i(\infty) \quad (3.5)$$

By replacing t with the solution of (3.5) in (3.2) it is possible to compute m_c . Once determined the number of message copies present in the network (m_c) and the average message delay ($MTTA$), we can calculate the energy consumption in this message dissemination process.

3.2 Energy consumption model

The participation in an oppnet implies the users must spend part of their device energy to transport messages of other people. How much energy a certain routing strategy will demand in the network is a key question for estimating the cost of the message dissemination process. Based on several previous works [8,9,26] we introduce a model to represent the energy consumption of the network, during an end-to-end message transmission. The model has four terms (3.6):

$$E_m = E_{mt} + E_{mr} + E_{dd} + E_{idle} \quad (3.6)$$

Where, E_{mt} is the energy consumed by the oppnet during the message transmission and E_{mr} represents the consumption due message reception. E_{dd} is the consumption during the device discovery process and E_{idle} is the energy consumption when the devices are idle. Equation 3.7 expresses the energy consumption during the message transmission. This consumption depends on both, the message length [9] and the kind of devices involved in the process [8,9]. For simplicity, in this analysis it is assumed that the message length is constant and the energy consumption is assumed to be the mean value among all devices (Δ_t). MT is the number of messages that are transmitted during the period that is being evaluated.

$$E_{mt} = \Delta_t MT \quad (3.7)$$

Equation 3.8 expresses the energy consumption during the reception of a message. It is assumed that a message is not broadcasted, but sent from one

node to another one. However, it is impossible to avoid that other neighbor nodes listen the message; therefore, those nodes will discard the message after reading the header (i.e., after verifying the target node). For this reason, γ represents this extra consumption, which is calculated for each particular case. In this proposal it is assumed that there is a 15% extra consumption due the header processing during the message reception.

$$E_{mr} = \Delta_r(1 + \gamma)MT \quad (3.8)$$

As in the previous case, Δ_r represents the energy consumption during reception and MT is the number of messages that are transmitted.

Equation 3.9 represents the energy consumed during the device discovery process. This operation must be done periodically and discover as much neighbors as possible. If a node fails in the detection of a neighbor and that neighbor is the destination node, then the message transmission will be unnecessary delayed.

On the other hand, a node cannot be continuously scanning for other nodes as the battery would be exhausted in a short time. The period T_{dd} represents the trade-off between the energy consumption and the probability of detecting new neighbors. As a rule of thumb, it can be set to be five times the rate of the inter-meeting times. β is a parameter that represents the energy consumption in the node during the device discovery process.

$$E_{dd} = \beta N \sum_n \frac{(t_n - t_{n-1})}{T_{dd}} \quad (3.9)$$

Finally, equation 3.10 computes the energy consumed by the nodes during the idle intervals. In that equation, α is the mean power demand while the device is idle, N is the number of nodes in the network and $(t_n - t_{n-1})$ is the time elapsed between two consecutive states (i.e., since the last successful transmission).

$$E_{idle} = \alpha N \sum_n (t_n - t_{n-1}) \quad (3.10)$$

The complete expression for the energy consumption is then obtained from equations 3.7, 3.8, 3.9, and 3.10. The equation 3.11 computes the energy consumption during a message propagation.

$$E_m = \Delta_t MT + \Delta_r(1 + \gamma)MT + N(\alpha \sum_n (t_n - t_{n-1}) + \beta \sum_n (t_n - t_{n-1})/T_{dd}) \quad (3.11)$$

For comparing the energy consumption among different message propagation strategies, it is interesting to consider the consumption in a relative way. Basically, it is assumed that the network has, before starting the transmission, a certain amount of energy that is computed as the energy stored in each node. For the sake of simplicity, let us say that the initial energy $E_{init} = NE_c$, that is the number of network nodes by the initial energy in each node. Equation

3.12 computes the percentage of residual energy left in the network after a message has been transmitted.

$$E_{res} = 100 \frac{E_{init} - E_m}{E_{init}} \quad (3.12)$$

The evaluation of the energy consumption in terms of Joules is not significant as there are many different devices and batteries. The literature reports that even for different devices and communication protocols, the consumption of an idle device is between 20% and 30% of the node consumption while transmitting/receiving a message [7,26]. In this work we assume such a consumption as 25%. The consumption of the device discovery process is almost identical to a message transfer; therefore, it is assumed as 90%.

This model considers the use of a single communication protocol in the message transmission process, however, it can be extended to support several alternatives. In order to do that, the discovery process energy consumption requires to perform the scanning of several radio antennas. While the energy consumption during transmission and reception is almost the same, the device discovery process consumes proportionally more than that, since the scanning process is performed for each protocol.

Another important aspect to consider is the time period required to transfer a message from source to destination node. Using a Markov chain it is possible to determine not only the Mean Time To Absorption (MTTA), but also the time necessary to achieve a 90% probability of successful delivery. The energy evaluation is done over that time interval.

4 Model Instantiation

The analytical model presented in the previous section should be instantiated according to the message dissemination strategy to be used. In order to illustrate the model instantiation process, we have considered the *Epidemic* and *Spray and Wait* routing strategies.

When an Epidemic dissemination is used, the Markov chain always has a symmetric construction. Therefore, general rules can be applied to compute the different transition rates between states based on the number of network nodes.

In case of Spray and Wait, the construction of the model is particular for each pair (C, H) . In many cases, the number of possible states to be considered could increase, as several combinations of messages copies can be present in the network. This feature limits the possibility of computing, in a generic way, the model for this routing strategy. Next we explain how to perform the instantiation of the analytical model for these routing strategies.

4.1 Modeling Epidemic Routing

Epidemic routing regularly uses the maximum amount of resources available in the network. As mentioned before, each node receiving the message becomes a “vector” capable of propagating it to other network nodes. In this way, a copy of the message may be present in every node using an important amount of memory and bandwidth. Provided that the Markov chain is symmetric, its transition matrix can be expressed through the following equation:

$$\forall i, j = 1, 2, \dots, N \quad Q_{ij} = \begin{cases} j(N-j) & j = i + 1 \\ -j(N-j) & j = i \\ 0 & \text{otherwise} \end{cases} \quad (4.1)$$

In particular, the equation 4.2 shows the general form (for the Laplace Transform - \mathbf{LT}) of each state k , for a network of N nodes that uses Epidemic routing. The transient probability function can be obtained from the Inverse Laplace Transform (\mathbf{LT}^{-1}).

$$\begin{aligned} \pi_1(s) &= \frac{1}{s + N\lambda} \\ \pi_k(s) &= \frac{\prod_{j=1}^{k-1} j(N-j)\lambda}{\prod_{j=1}^k (s + (j(N-j) + j)\lambda)} \quad k \geq 2 \end{aligned} \quad (4.2)$$

As mentioned before, the solution for each particular state of the Markov chain can be found with the help of a solver like `Matlab`, or by hand using the regular Inverse Laplace Transform tables. These analytical methods are not suitable for addressing medium-size to large networks (i.e., oppnets with more than 20 nodes); therefore, in these cases we recommend the use of a numerical approximation, e.g., based on ODE45 [15].

For computing the *MTTA* in an oppnet that uses Epidemic routing, it is necessary to reformulate the Markov chain as shown in Figure 2. Then, the \widehat{Q} is obtained from:

$$\forall i, j = 1, 2, \dots, N-1 \quad \widehat{Q}_{ij} = \begin{cases} j(N-1-j) & j = i + 1 \\ -j(N-j) & j = i \\ 0 & \text{otherwise} \end{cases} \quad (4.3)$$

The *cfid* for each state is obtained from equation (3.4) and it is given by the following:

$$\begin{aligned}
L_1(\infty) &= \frac{1}{N\lambda} \\
L_k(\infty) &= \frac{\prod_{j=1}^{k-1} j(N-j)\lambda}{\prod_{j=1}^k (j(N-j) + j)\lambda} \quad k \geq 2
\end{aligned} \tag{4.4}$$

From equation (4.4) it is possible to compute the *MTTA*.

$$MTTA = \frac{1}{N\lambda} \sum_{i=1}^N \frac{1}{i} \tag{4.5}$$

The expected number of copies can be computed from (3.2) for this routing strategy. As there is no general expression for the $\pi_i(t)$, the solution to (3.2) depends on the number of nodes in the network. Then, the energy consumed in the network is a function of the amount of copies present at the moment of absorption.

4.2 Modeling Spray and Wait Routing

This message dissemination strategy limits the number of copies in the network and involves two phases. In the first one, the message is delivered from the *source* node to a limited number of intermediate nodes. In the second phase, these nodes are in charge of transmitting the message to the *destination* one. In this strategy, two parameters define how the messages are propagated in the network. The first one is the number of allowed copies (C) that determines the bandwidth required for the transmissions. The second one is the number of *hops* (H) allowed for the message to reach the *destination* node. It defines how many nodes can propagate the message.

The parameters C and H can be used to set several dissemination conditions in the network. At the moment the message is ready to be transmitted by the *source* node, it has the capacity of delivering C copies of the message. The way in which these copies are distributed depends on the number of hops allowed. Each time the *source* or *intermediate* nodes can pass as much as $H - j - 1$ messages to the next node, where j is the amount of hops already taken. This approach is completely different from the Epidemic strategy, in which transmitting the message to another node does not reduce the capacity of transferring it to another node later.

Figure 3 shows an example of a Binary Spray and Wait message propagation. In this case, when the *source* node meets an *intermediate* node, it passes half of the copies it has, and each *intermediate* node does the same. For this to occur, it is required that $H = \lfloor C/2 \rfloor$ and $H > 2$. Figure 3 shows one possible path for delivering the message to the destination node. The complete Markov

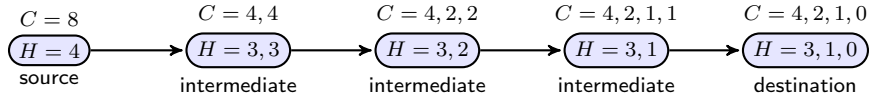


Fig. 3 Spray and Wait example for $C=8$ and $H=4$

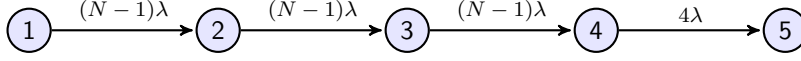


Fig. 4 Spray and Wait model for $C=4$ and $H=2$

Chain for this case contains many more states, as there are different possible combinations with the same number of copies in the network.

Next subsections show the message transmission process using several combinations of C and H as a way to illustrate how several dissemination conditions can be set in this model. Although these parameters are interdependent, by increasing one of them does not ensures the improvement of the network performance. For instance, given a certain value for H , incrementing C improves the network performance since more nodes will be able to transmit the message once they receive a copy. However, given a value of C , incrementing H does not improve the performance in every case. As it will be shown in the next subsections, the network performance for $H = 2$ and $H = 4$ is identical when $C = 4$, but it is marginally better when $H = 3$.

4.2.1 Two hops allowed

Using this configuration ($H = 2$), the message can go through only one *intermediate* node that will eventually pass it to the *destination* node. The *source* node distributes C copies to an identical amount of *intermediate* nodes. Figure 4 shows the way in which the message is distributed for the case of four copies ($C = 4$). The size of the network (N) defines the transition rates, but the amount of states in the CTMC is independent of the network size. The chain will have $C + 1$ states in every case. When the message has been copied to C *intermediate* nodes, it will have a constant ratio of $C\lambda$ to reach the *destination* node. This particular case is regular and the transition rates can be expressed in general terms as a function of N and C .

The Q matrix is built in the following way:

$$\forall i, j \in \{1, C+1\} \quad Q_{ij} = \begin{cases} -(N-1) & j = i \text{ \& } i < C \\ (N-1) & j = i+1 \text{ \& } i < C \\ -C & j = i \text{ \& } i = C \\ C & j = i+1 \text{ \& } i = C \\ 0 & \text{otherwise} \end{cases} \quad (4.6)$$

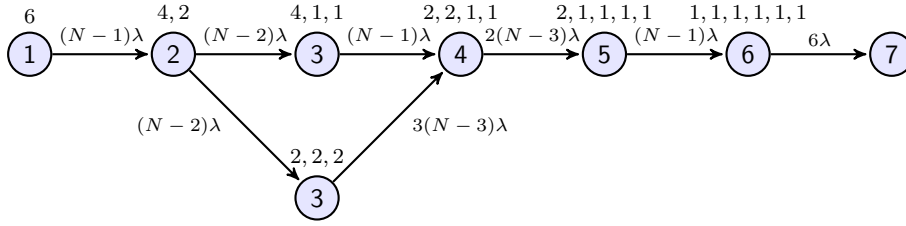


Fig. 5 Spray and Wait model for $C=6$ and $H=3$

The reduced \hat{Q} matrix can be obtained redrawing the CTMC and eliminating the absorbing state.

$$\forall i, j \in \{1, C\} \quad \hat{Q}_{ij} = \begin{cases} -(N-1) & j = i \text{ \& } i < C \\ (N-i-1) & j = i+1 \text{ \& } i < C \\ -C & j = i \text{ \& } i = C \\ 0 & \text{otherwise} \end{cases} \quad (4.7)$$

4.2.2 Three hops allowed

In this case ($H = 3$), the *source* node transfers to the *intermediate* ones two copies of the message. These nodes can transfer one of these copies to other nodes, and eventually the message arrives to the *destination* node. The number of copies allowed in the system should be at least three. With three or four copies, the propagation model is similar to the case of $H = 2$, as there is only one possible network state for each distribution of copies in the nodes. With $C \geq 4$, there is more than one state having the same amount of copies in the system. Figure 5 shows a transmission example for $C = 6$, where two possible states can be assumed involving the three message copies in the system.

The matrix Q has no general form since the amount of additional states, representing the same amount of copies in the network, depends on the maximum allowed number of copies.

The second state in the CTMC has the following distribution of copies. One node has four copies while the other has only two copies. The chain can progress in two different directions. The node holding four copies may find another node and pass it two copies of the message. In this case, there will be three nodes, each one holding two copies. However, if the nodes holding two copies meet another node, then the chain evolves towards other combination with one node holding four copies and two nodes holding only one. It is clear that both paths have exactly the same probability. In the second state both nodes holding copies of the message have the same probability of finding another node. This is the reason why we divide the output rate from the second state, for the states representing three different nodes holding copies of the message in equal parts $(N-2)$. In these cases, there are more states in the CTMC than copies of the message in the network. Thus, for the computation of m_c , the amount of copies related to each state should be considered. For instance, when $C = 6$ and $H = 3$ the following instantiation should be done:

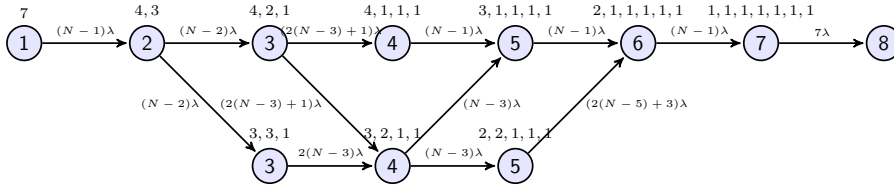


Fig. 6 Spray and Wait for $C=7$ and $H=4$

$$m_c(t) = \pi_1(MTTA) + 2\pi_2(MTTA) + 3\pi_3(MTTA) + 3\pi_4(MTTA) + 4\pi_5(MTTA) + 5\pi_6(MTTA) + 6\pi_7(MTTA) \quad (4.8)$$

4.2.3 Four hops allowed

In this case ($H = 4$), the minimum C is four. The *source* node propagates two copies to *intermediate* nodes and keeps a copy for the case of meeting the *destination* one. *Intermediate* nodes propagate one copy to another *intermediate* one and keep one for the case of reaching the *destination* node.

In Figure 6, the CTMC for seven copies is shown. The results in the second state indicate that the *source* node has four copies while there is only one *intermediate* node with three copies. At this point, like in the case of $H = 3$, there are two possible paths. In the first one, the *intermediate* node meets another node and transfers two copies, resulting in a distribution of four, two and one copies for the *source*, *first intermediate* and *second intermediate* nodes respectively. In the second path, the *source* node meets another node and transfers three copies of the message, keeping just one for itself. The distribution in this case is three and one copy for the *intermediate* and source nodes respectively.

In the third state the situation is repeated. There are three nodes with message copies, but only two of them can propagate it to other *intermediate* nodes. Again, it may happen that the *source* node meets another one and in that situation it transfers three copies, keeping one for itself. After this, there will be two nodes with one copy, one node with two copies and one node with three copies. The other path is followed if the *intermediate* node holding two copies meets another one. In that case, it transfers one copy and keeps the other one for itself, resulting in a distribution of one node with four copies and three nodes with one copy each.

The fourth state can only progress to the sixth one with a distribution of two nodes with one copy, one node with two copies and one node with three copies. In the sixth state there are again two possible paths. In the first one, the node with three copies meets another one and transfers two copies, keeping one for itself. After this, the distribution will be two nodes with two copies, and three nodes with one copy. The other path is followed when the node with two copies meets another one and transfers one copy. In that case the distribution is one node with three copies, and four nodes with one copy each.

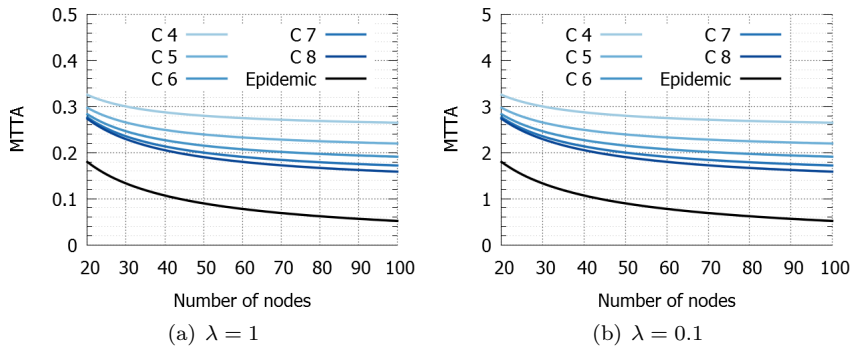


Fig. 7 $MTTA$ for $H=2$

As it can be seen in Figure 6, the states representing three, four and five nodes with at least one copy of the message are duplicated. This has to be considered when building the transition matrix Q and \hat{Q} , so the $MTTA$, m_c and E_c can be properly computed.

It is important to remark that once built the model, it can be reused, avoiding thus that software or communication protocol designers have to create it every time that they need to compare the performance of various dissemination strategies or design a new one. The model setting process is also simple. This shows that this modeling proposal reduces the effort required to count on and use the message dissemination model, by comparing it with effort required when using simulations or empirical tests.

5 Message Delivery Performance Evaluation

In order to illustrate how to use the instantiated model, we evaluate the performance of the Epidemic and Spray and Wait routing strategies by comparing their $MTTA$. This was done by computing the CTMC transition matrix Q and \hat{Q} for several combinations of the number of network nodes (N), message copies (C) and hops (H). The differential equations for each combination were solved using the ODE45 in Octave.

Figure 7 shows the behavior of the $MTTA$ for the message delivery when up to two hops are used. The $MTTA$ remains almost identical if we use three or four hops. As it can be seen, reducing the inter-meeting ratio ten times increases the $MTTA$ in the same proportion. This can be explained by looking at equation 4.5. As shown, the inter-meeting rate is just a scaling factor. For simplicity, in the rest of the performance evaluation we assume $\lambda = 1$.

The figure also shows that the Epidemic routing is always the one with the best throughput for messages; i.e., the one with the shortest delay for delivering messages. However, this performance requires an important consumption of resources as it will be shown in Section 6. This consumption of resources

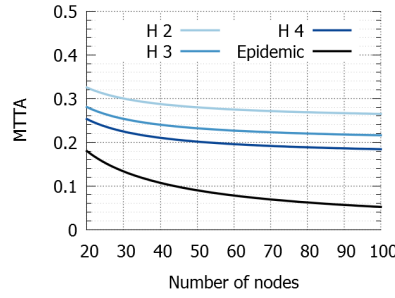


Fig. 8 $MTTA$ for $C=6$

affects the network during a short time period, due the message delivery in Epidemic is the faster than in other routing strategies.

The performance analysis shows that the improvement in the transmission delay, when using a Spray and Wait strategy with a number of allowed copies, tends to saturate the network. There is a remarkable improvement of performance between the use of four and eight copies for the message delivery. However, there is a small difference in performance when we use seven or eight message copies. Actually, the improvement between using seven and eight copies is smaller than the one obtained from four to five. This behavior is similar if we consider two, three or four hops.

The results of the $MTTA$ shown in Figure 8 considers six copies, but using different numbers of hops for the message delivery. The cases involving three and four hops ($H = 3$ and $H = 4$ respectively) have the best performance. This is because with that combination the Spray and Wait is binary, i.e., each node transmits to the next one half of its copies.

This shows that the results obtained from the instantiated model are rich enough to analyze the performance of a certain message delivery strategy, or compare the performance of various of them. Moreover, tuning and running the instantiated model is also simple. This opens several opportunities to perform iterative modeling (or comparisons) of message delivery strategies.

6 Determining the number of copies and the energy consumption

Once analyzed the performance of a message delivery strategy, we have to determine the resources consumed by each routing strategy, particularly memory and energy, according to the probability of delivering the message to the destination node. The first variable (i.e., memory) is determined by the number of nodes with a message copy (m_c) in the network, and the second one represents the energy left in the network (E_{res}).

The probability of a successful transmission can be determined by computing the probability of reaching the absorbing state after a period of time. Figure 9 shows the probability of reaching the destination node using different routing strategies, and according to several $MTTA$ values. These results show

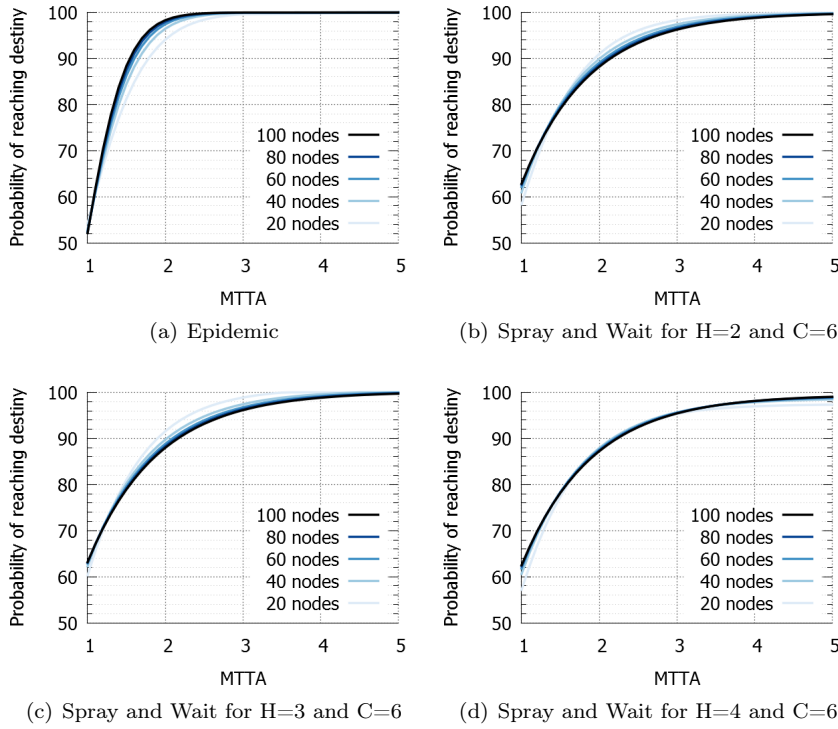


Fig. 9 Probability of reaching destination node at different times (expressed as multiples of $MTTA$)

that the probability of reaching the destination node at $2MTTA$ is over 85% for all routing strategies. Considering this fact, it is possible to determine the demand of resources, setting the message lifetime to two times the $MTTA$.

In all combinations of C and H for Spray and Wait, the expected number of copies in the network at the moment of absorption tends to C . Instead, in the Epidemic routing, the expected number of copies in the system is close to the network size. These results indicate that when using Epidemic routing, almost all nodes will have a copy of the message before it reaches the *destination* node.

Figure 10 shows the residual energy present in the network. In the first one the device discovery process is done at two times the meeting rate, while in the second one it is done at five times the meeting rate. The results show that the device discovery process consumes an important amount of energy; therefore, if we use it frequently, we will degrade the performance of the Spray and Wait (in term of energy consumption). Apparently, the Epidemic strategy demands more energy since a more important number of parallel transmissions are allowed. However, this is partially true, because the delivery process is the shortest, therefore the network is disseminating the messages during a time

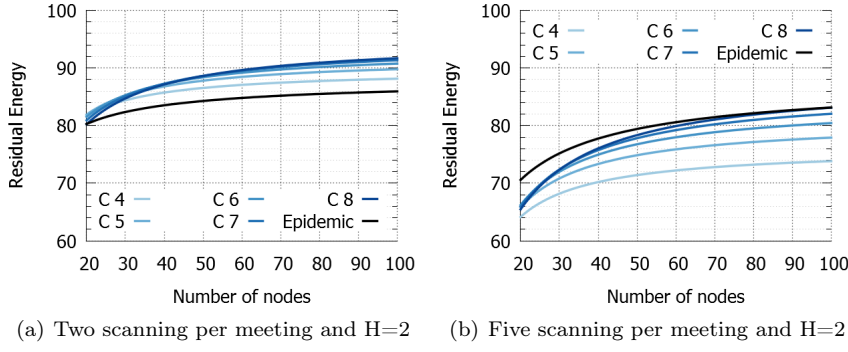


Fig. 10 Percentage of residual energy after 2 MTTA, considering 2 and 5 instances of device discovery between every nodes meeting

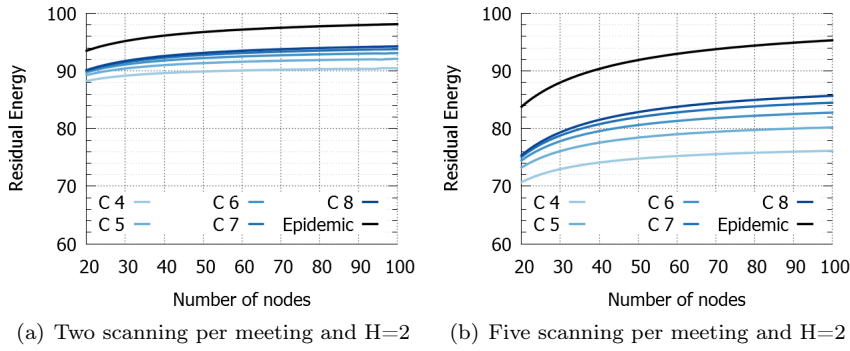


Fig. 11 Percentage of residual energy due device discovery after 2MTTA, considering 2 and 5 scanning between each nodes meeting.

period that is shorter than Spray and Wait (until a message copy reaches the destination node).

Contrarily, Spray and Wait consumes fewer resources in each message delivery round, but it requires more time to reach the destination. Therefore, the device discovery process is repeated frequently, which makes the energy consumption higher. This situation becomes explicit in Figure 10, where Epidemic routing is the most expensive strategy when we perform two device discovery between every nodes meeting. However, if we increase to 5 times the device scanning frequency, the energy consumption for Spray and Wait overcomes the Epidemic.

Figure 11 shows how the discovery process affects the residual energy of the network, depending on the dissemination strategy that is used. These strategies are ranked from lower to higher according to their energy consumption, which corresponds to their *MTTA*; e.g., Epidemic (that has the shortest *MTTA*) is the strategy that has the lowest energy consumption, then follows Spray and Wait with 8 copies, and so on.

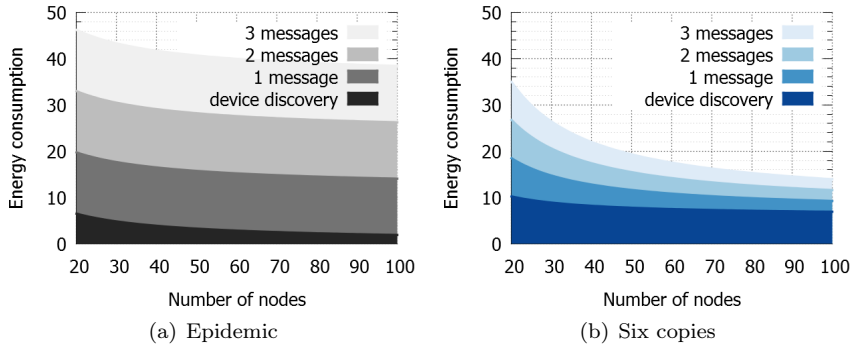


Fig. 12 Energy consumption for concurrent transmissions

After evaluating several frequencies for device discovery, we can say that the energy consumption increases with this frequency. However, the ranking of strategies is remained in the order shown in Figure 11.

Figure 12 shows the energy consumption when several messages are sent in a concurrent way. In this case, the energy required per message to be transmitted in the Epidemic strategy is higher than the energy required by the Spray and Wait with $C=6$ and $H=2$. This result is relevant because in an oppnet usually several messages are being transmitted among the nodes in a concurrent way.

Choosing a strategy for the message dissemination is not easy, since we have to consider the required dissemination speed, network autonomy (in terms of energy) and also the number of nodes that will eventually produce messages for other nodes. For instance, in applications for public advertisement, where energy consumption is ensured, there is usually an information producer and several interim nodes replicating the information and trying to reach the destination. Provided that the producer would like to irradiate as much nodes as possible during each dissemination round, using an Epidemic strategy would be more effective.

For applications that involve many information producers, like those supporting environmental monitoring or first responses during emergencies, a Spray and Wait approach is probably more effective, due these networks have to be alive for long time periods. Although the time for the message delivery is important in these scenarios, much more important is to keep the network alive.

In order to make this type of decisions, it is important that software or protocols designers count on a modeling approach that allows them to perform iterative modeling and evaluations involving a low effort. In this sense, the proposed approach has an advantage over the simulations and the empirical tests.

7 Conclusions and Future Work

This paper proposes a stochastic approach for modeling oppnets based on a Continuous Time Markov Chain model. This approach allows modeling and evaluating message delivery strategies on oppnets from an analytical point of view. Once created the model, it can be easily reused, by reducing thus the effort required to create these models, compared to implementing simulations or conducting empirical tests. Moreover, the analytical models can be instantiated or tuned in a simple way, which allows the iterative modeling and contributes to reduce the time and effort required for evolve the message delivery strategies. Therefore, the proposed approach is particularly useful to prototype solutions in an interactive way and involving a low effort, compared to using simulations or empirical tests. Designers of several types of systems, such as mobile collaborative systems, mobile wireless sensors networks and mobile distributed systems can take advantage of it.

In order to illustrate the usability and usefulness of the proposal, the message dissemination process for Epidemic and Spray and Wait routing strategies were modeled using this approach. The resulting models were instantiated to determine the performance of each routing strategy in particular interaction scenarios. This also allowed us to compare the performance of these strategies in terms of mean time to absorption, expected number of copies and residual energy in the network.

The performance analysis introduced two parameters, C (number of message copies) and H (number of hops used for the message delivery) for the Spray and Wait strategy. By combining these parameters we can represent several strategies based on Spray and Wait (e.g., binary, one copy or more than one).

The results of the performance analysis show that Epidemic routing is always the option with the shortest delay in transmitting a message. However, it is also the strategy that consumes more memory, and depending on the device discovery protocol, it may also consume more energy than any strategy based on Spray and Wait. Contrarily, the Spray and Wait is preferable in applications that need to save as much energy as possible, in order to prolong the network autonomy. This strategy is also useful when the network nodes have little memory. As previously shown, performing this kind of analysis usually involves a low effort if we use the analytical models, opening thus an alternative to the simulations and empirical tests.

The next steps in this initiative considers to evaluate the proposal formally comparing the modeling effort required to represent message dissemination using the proposed stochastic approach and also simulations. This will be done to quantify the reduction of the modeling effort obtained for using this proposal.

Acknowledgements s This work was partially supported by the Spanish government under the contracts TIN2013-47245-C2-1-R and TIN2016-77836-C2-2-R.

References

1. Boldrini, C., Conti, M., Passarella, A.: Modelling data dissemination in opportunistic networks. In: Proceedings of the third ACM workshop on Challenged networks, CHANTS, pp. 89–96. ACM (2008). DOI 10.1145/1409985.1410002
2. Boldrini, C., Conti, M., Passarella, A.: Modelling social-aware forwarding in opportunistic networks. In: K. Hummel, H. Hlavacs, W. Gansterer (eds.) Performance Evaluation of Computer and Communication Systems. Milestones and Future Challenges, *LNCIS*, vol. 6821, pp. 141–152. Springer Berlin Heidelberg (2011). DOI 10.1007/978-3-642-25575-5_12
3. Boldrini, C., Conti, M., Passarella, A.: Performance modelling of opportunistic forwarding under heterogenous mobility. *Computer Communications* **48**, 56 – 70 (2014). DOI 10.1016/j.comcom.2014.03.028
4. Cai, H., Eun, D.Y.: Crossing over the bounded domain: From exponential to power-law intermeeting time in mobile ad hoc networks. *Networking, IEEE/ACM Transactions on* **17**(5), 1578–1591 (2009). DOI 10.1109/TNET.2008.2011734
5. Cho, D.K., Lee, S.H., Chang, A., Massey, T., Chang, C.W., Tsai, M.H., Sarrafzadeh, M., Gerla, M.: Opportunistic medical monitoring using bluetooth p2p networks. In: World of Wireless, Mobile and Multimedia Networks. WoWMoM 2008. International Symposium on a, pp. 1–6 (2008). DOI 10.1109/WOWMOM.2008.4594895
6. Ciobanu, R.I., Dobre, C.: Predicting encounters in opportunistic networks. In: Proceedings of the 1st ACM Workshop on High Performance Mobile Opportunistic Systems, HP-MOSys, pp. 9–14. ACM (2012). DOI 10.1145/2386980.2386983
7. Drula, C., Amza, C., Rousseau, F., Duda, A.: Adaptive energy conserving algorithms for neighbor discovery in opportunistic bluetooth networks. *Selected Areas in Communications, IEEE Journal on* **25**(1), 96–107 (2007). DOI 10.1109/JSAC.2007.070110
8. Ergen, M., Varaiya, P.: Decomposition of energy consumption in ieee 802.11. In: Communications, 2007. ICC '07. IEEE International Conference on, pp. 403 –408 (2007). DOI 10.1109/ICC.2007.73
9. Garcia-Saavedra, A., Serrano, P., Banchs, A., Bianchi, G.: Energy consumption anatomy of 802.11 devices and its implication on modeling and design. In: Proceedings of the 8th international conference on Emerging networking experiments and technologies, CoNEXT, pp. 169–180. ACM (2012). DOI 10.1145/2413176.2413197
10. Garg, K., Giordano, S., Förster, A.: A study to understand the impact of node density on data dissemination time in opportunistic networks. In: Proceedings of the 2Nd ACM Workshop on High Performance Mobile Opportunistic Systems, HP-MOSys, pp. 9–16. ACM (2013). DOI 10.1145/2507908.2507919
11. Grasic, S., Lindgren, A.: An analysis of evaluation practices for dtn routing protocols. In: Proceedings of the seventh ACM international workshop on Challenged networks, CHANTS, pp. 57–64. ACM (2012). DOI 10.1145/2348616.2348629
12. Groenevelt, R., Nain, P., Koole, G.: The message delay in mobile ad hoc networks. *Performance Evaluation* **62**, 210–228 (2005). DOI 10.1016/j.peva.2005.07.018
13. Keränen, A., Ott, J., Kärkkäinen, T.: The one simulator for dtn protocol evaluation. In: Proceedings of the 2nd International Conference on Simulation Tools and Techniques, Simutools, pp. 55:1–55:10. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering) (2009). DOI 10.4108/ICST.SIMUTOOLS2009.5674
14. La, R.: Distributional convergence of intermeeting times under the generalized hybrid random walk mobility model. *Mobile Computing, IEEE Transactions on* **9**(9), 1201–1211 (2010). DOI 10.1109/TMC.2010.66
15. Lambert, J.D.: Numerical Methods for Ordinary Differential Systems: The Initial Value Problem. John Wiley & Sons, Inc. (1991)
16. Monares, Á., Ochoa, S.F., Santos, R., Orozco, J., Meseguer, R.: Modeling iot-based solutions using human-centric wireless sensor networks. *Sensors* **14**(9), 15,687–15,713 (2014). DOI 10.3390/s140915687
17. Mota, V.F., Cunha, F.D., Macedo, D.F., Nogueira, J.M., Loureiro, A.A.: Protocols, mobility models and tools in opportunistic networks: A survey. *Computer Communications* **48**, 5 – 19 (2014). DOI 10.1016/j.comcom.2014.03.019

18. Neglia, G., Zhang, X.: Optimal delay-power tradeoff in sparse delay tolerant networks: a preliminary study. In: Proceedings of the 2006 SIGCOMM workshop on Challenged networks, CHANTS, pp. 237–244. ACM (2006). DOI 10.1145/1162654.1162661
19. Reina, D.G., Askalani, M., Toral, S.L., Barrero, F., Asimakopoulou, E., Bessis, N.: A survey on multihop ad hoc networks for disaster response scenarios. *Journal of Distributed Sensor Networks* **2015**, 16 (2015). DOI 10.1155/2015/647037
20. Sakaguchi, H., Izumi, T., Nakatani, Y.: An opportunistic tourism navigation system using photographing point recommendation. In: Technologies and Applications of Artificial Intelligence (TAAI), 2013 Conference on, pp. 318–323 (2013). DOI 10.1109/TAAI.2013.69
21. Santos, R., Mosse, D., Znati, T., Comfort, L.: Design and implementation of a witness unit for opportunistic routing in tsunami alert scenarios. *Safety science* **90**, 75–83 (2016). DOI 10.1016/j.ssci.2015.09.014
22. Santos, R., Ochoa, S.: Disseminating shared information in disaster relief efforts: A communication computable model. In: IEEE International Conference on Systems, Man and Cybernetics 2011, pp. 3446–3451 (2011)
23. Santos, R., Orozco, J., Ochoa, S.F.: A real-time analysis approach in opportunistic networks. *ACM SIGBED Review* **8**(3), 40–43 (2011). DOI 10.1145/2038617.2038626
24. Sermpezis, P., Spyropoulos, T.: Understanding the effects of social selfishness on the performance of heterogeneous opportunistic networks. *Computer Communications* **48**, 71 – 83 (2014). DOI 10.1016/j.comcom.2014.03.016
25. Spyropoulos, T., Turetli, T., Obraczka, K.: Routing in delay-tolerant networks comprising heterogeneous node populations. *Mobile Computing, IEEE Transactions on* **8**(8), 1132 –1147 (2009). DOI 10.1109/TMC.2008.172
26. Wang, W., Srinivasan, V., Motani, M.: Adaptive contact probing mechanisms for delay tolerant applications. In: Proceedings of the 13th annual ACM international conference on Mobile computing and networking, MobiCom, pp. 230–241. ACM (2007). DOI 10.1145/1287853.1287882
27. Whitbeck, J., Conan, V., de Amorim, M.: Performance of opportunistic epidemic routing on edge-markovian dynamic graphs. *Communications, IEEE Transactions on* **59**(5), 1259 –1263 (2011). DOI 10.1109/TCOMM.2011.020811.090163
28. Yao, H., Huang H.and Zeng, D., Li, B., Guo, S.: An energy-aware deadline-constrained message delivery in delay-tolerant networks. *Wireless Networks* **20**, 1981–1993 (2014)
29. Yoon, S.K., Haas, Z.J.: Efficient tradeoff of restricted epidemic routing in mobile ad-hoc networks. In: Military Communications Conference, 2007. MILCOM 2007. IEEE, pp. 1 –7 (2007). DOI 10.1109/MILCOM.2007.4455272
30. Zhang, X., Neglia, G., Kurose, J., Towsley, D.: Performance modeling of epidemic routing. *Computer Networks* **51**(10), 2867–2891 (2007). DOI 10.1016/j.comnet.2006.11.028