Fixed Point Opportunistic Routing in Delay Tolerant Networks

Vania Conan*, Jérémie Leguay,*†, Timur Friedman†
* Thales Communications
† Université Pierre et Marie Curie, Laboratoire LiP6–CNRS

***Abstract*— We propose in this work a single copy and multi-hop opportunistic routing scheme for sparse delay tolerant networks (DTNs). The scheme uses as only input the estimates of the average inter-contact times between the nodes in the network. Defined as the fixed point of a recursive process, it aims at minimizing delivery time in case of independent exponential pairwise inter-contacts. The two properties of loop-free forwarding and polynomial convergence make the scheme workable for routing in DTNs. The routing performances of the scheme are evaluated on three publicly available reference data sets. Comparisons with well known single-copy schemes, including MED and the *two hop relay* strategy, consistently demonstrate improvements for both delivery ratio and delay.**

I. Introduction

In delay tolerant networks (DTNs) [9] nodes are typically mobile and have wireless networking capabilities. They are able to communicate with each other only when they are within transmission range. The network suffers from frequent connectivity disruptions, making the topology only intermittently and partially connected. This means that there is no guarantee that an end-to-end path exists between a given pair of nodes at a given time. Examples from the recent literature include the DieselNet project [24], which features communication devices deployed in a regional bus system, and Pocket Switched Networks (PSNs) [6], which are formed by devices that people carry every day, such as cell phones, PDAs, and music players.

The main contribution of this work is the introduction and evaluation of a novel opportunistic routing scheme for sparse DTNs built upon the fixed point of a recursive process. The scheme is single copy, thus keeping the load in the network low. It uses as only input the estimates of the average inter-contact times between the nodes in the network. The scheme provides an opportunistic version of the *Minimum Expected Delay* (MED) routing introduced by Jain et al. [13], whereby a node relays a message to a neighbor that is closer, in terms of total expected delivery time, to the destination. It is loop-free and convergences polynomially, which make the scheme workable for routing in DTNs. We formally derive the scheme for the case of heterogeneous independent exponential inter-contacts and evaluate it through simulation on three reference data sets publicly available in the CRAWDAD archive [1].

Following the taxonomy introduced by Jones et al. [16], routing propositions for DTNs can be divided into three main categories: *replication* based, *knowledge* based and *hybrid* strategies. Replication based approaches take advantage of node diversity. They address ways several copies of the message can be disseminated among several carriers to increase the chance that it would reach the destination. Knowledge based strategies make use of information that nodes obtain about connectivity or network conditions to make efficient forwarding decisions that improve routing performance. Hybrid approaches, as suggested, combine both the replication and knowledge based strategies.

The knowledge and hybrid based routing schemes aim at reaching high delivery ratio, low average delivery delays and limited overhead. The general principle of these approaches is the following: a node routing a given message applies forwarding rules that tell it, for each other node it encounters, whether it should give it the message (or copies of the message) or whether it should keep it. The rules are heuristics that estimate whether the encountered node is closer to the destination. Different proposals use different such *distance* or *utility* measures.

Chen and Murphy [7] define a utility function that locates the relay node within a connected cluster that is closest to a disconnected destination. Lindgren et al. [20] rely on nodes having a community mobility pattern. Their scheme uses history of encounters and the transitivity of the estimated delivery probability to distinguish between candidate relays. Nodes mainly remain inside their community and sometimes visit the others. As a consequence, a node may transfer a message to a node that belongs to the same community as the destination. Burns et al. [4] use both information of contacts between nodes and of visits to locations for routing. The movement patterns are structured and each node learns the probability that another node can successfully deliver a message to the destination. Burgess et al. [3] have proposed the protocol MaxProp in the context of a real DTN deployment on 40 buses. This protocol uses meeting probabilities to find paths in association to complementary mechanisms for improving performance in terms of delivery ratio and latency such as buffer management and transmission scheduling. Leguay et al. [18] define a high-dimensional Euclidean space, called MobySpace, constructed upon nodes mobility patterns. The specific MobySpace evaluated is based on the frequency of visits of nodes to each possible location.

The present work fully exploits the transitivity that the distance or utility measures introduce. Lindgren et al. [20] take into account transitivity as one of the three components of the heuristics used to compute the *delivery predictability* of any node to any destination that informs routing decisions.

Burns et al. [4] discussed the impact of several relaying steps on their scheme but used only the single relay formula for routing. We propose a workable solution and algorithm that generalises these previous works. It is based on estimating expected message delivery time and faithfully accounts for the effects of transitivity in the case of any number of relaying steps.

The distance that the algorithm computes is thus capable of anticipating the effects of future relaying opportunities to improve routing performances. To derive the scheme, we start from the *two hop relay* strategy introduced by Grossglauser and Tse [11] in which one relay is used to reach the destination. Then, we use the expected delivery time to the destination as the distance to be minimized to define the best set of candidate relays. As we recursively increase the number of relaying steps, the expected delivery time diminishes, and the estimated distance to the destination gets refined. This procedure is shown to converge and provides a fixed point routing strategy that decreases the overall expected delivery time.

The scheme uses the average meeting times between any two nodes to distinguish among candidate relay nodes. As we will see by looking at real life data sets (see Sec. III), these average meeting times span a large spectrum of values in the DTN. To anticipate the effect of successive relaying steps on delivery time, the scheme makes use of a formal model of inter-contact patterns between nodes. Many of the models used for evaluating DTN routing protocols do not explicitly address the inter-contact variability that we observed in the data sets. For example, Spyropoulos et al. model mobility of nodes as independent random walks on a torus, and use it to analyze the performance of different routing schemes [22]. Their model considers all pairs of nodes to follow the same law, with the same parameters. This *homogeneity* hypothesis is common to many DTN models. Groenevelt et al. [10] study a multicopy version of the *two hop relay* strategy for the homogeneous model that captures the characteristics of the network through a single parameter representing the expected inter-contact time between any two pairs of nodes. Chaintreau et al. [6] model the sequence of contacts as a discrete renewal process, and study power-law distributed inter-contacts. Karagiannis et al. [17] analyze mobility traces and explain the observed exponential tail behavior of inter-contact times with a simple random walk on a two dimensional torus followed by all nodes in the network. In the present work, we model pairwise contacts as independent Poisson processes *with different parameters*. Contrary to previous work, we put the focus on capturing the differences in average inter-contact times between any node and its neighbors. The model is crude and we do not claim it matches all real-life inter-contact behaviors, in particular complex ones that would involve a mixture of different pairwise laws (e.g., a mix of light-tail and heavy-tail distributions). Our solution is built upon a memoryless model which is supported by some recent theoretical and experimental results [17], [14]. Its purpose is to capture average inter-contact time heterogeneity and remain tractable for the derivation of the multi-hop routing scheme. We also validate the overall DTN routing scheme on three real life data sets showing that the scheme, although derived formally on the model, provides enhanced performance beyond the original mobility hypothesis.

The rest of this paper is structured as follows. Sec. II introduces our novel opportunistic routing scheme, and Sec. III provides its evaluation. Sec. IV discusses the routing schemes that we proposed. Sec. V concludes the paper.

II. Fixed point opportunistic routing

In this section we present a single copy and multi-hop routing strategy that is obtained as a fixed point of a recursive algorithm that minimizes the delivery delay of messages across the DTN. Central to the establishment of this result and to the derivation of the routing scheme and algorithm is a memoryless hypothesis akin to the well known properties of independent exponential variables. This inter-contact model hypothesis is introduced in Sec. II-A. Sec. II-B derives the formula for expected delivery delay for the *two hop relay* scheme of Grossglauser and Tse [11]. Sec. II-C introduces a variation of the *two hop relay* scheme leading to minimum delay and Sec. II-D extends the later to the multi-hop case using a recursion.

A. Inter-contact pattern model

For the purpose of our study we model the inter-contact patterns between any two nodes in the network as a set of independent Poisson processes. In other words inter-contact times between any two nodes i and j are independent identically distributed *(iid)* random variables and the *(iid)* variables follow exponential distributions with parameters λ_{ij}.

The average inter-contact time is given by $1/\lambda_{ij}$, so different values of the parameter account for the heterogeneity of inter-contacts in the network. In real data sets, a node i only encounters a subset of all other nodes in the network. We call these nodes the *neighbors* of i, and each of these neighbors is encountered with a different average time. In the model, if nodes i and j never meet we set $\lambda_{ij} = 0$ and the corresponding average meeting time $1/\lambda_{ij}$ is infinite.

The model is used in the derivation of the opportunistic minimum delay scheme to evaluate the first encounter time between any node and any subset of its neighbors. This is where the exponential and independence hypotheses are used (see Sec. II-B). The equations that define the routing scheme only make use of the average meeting times between the nodes. This makes it possible to estimate the required average meeting time inputs from real data, and to analyze the performances of the scheme in cases that do not fully comply with the hypotheses that were used to derive it.

The formal derivations and results below still hold if we replace the model above by what Carreras et al. designate the Marks Memoryless model [5]. In that case the exponential distribution hypothesis can be dropped and replaced by any distribution with finite expectation. One could then consider truncated power laws (e.g., power laws with an exponential cut-off as in [17]) for pairwise inter-contact distributions. Note that the scheme would nevertheless not apply to pure power-laws with shape parameter lower than 2 as they do not have finite expectations. The Marks Memoryless model keeps a strong independence hypothesis since the identity of the node pair that meets at each encounter instant is independent of the encounter time. Carreras et al. show on several simulation examples that this is verified by some instances of the most common mobility models used in the MANET community, namely Random Walk, Random Direction and Random Waypoint. But this is probably still too strong to cover all real data sets as we will discuss in Sec. IV.

B. two hop relay *strategy*

The *two hop relay* strategy works in two phases [11]. In the first phase, the source node waits for its first encountered neighbor. If this neighbor is not the destination, the source uses it as a relay, gives it the message and does not keep a copy. In the second phase, the relay node waits until it meets the destination to deliver the message. In the rest of this paper, by convenience of notation, the *two hop relay* scheme is denoted 2-MH (two-hop version of a potentially Muti-Hop routing scheme).

Let us first consider the first phase. The message is injected at source s at time instant t. The first node r it encounters may be any of the $n-1$ other nodes $d, r_1, r_2, ..., r_{n-2}$ and the time X it takes to meet this first node is the infinum of the inter-contact times with all other nodes:

$$X = inf(R^t_{sd}, R^t_{sr_1}, ..., R^t_{sr_{n-2}}) \tag{1}$$

where R^t_{sr} is the remaining inter-contact time, after t, before the next contact between nodes s and r. By the memoryless property of the exponential distribution, R^t_{sr} is also exponential with parameter λ_{sr}.

Since all $R^t_{sr_i}$ are independent exponentials with parameters λ_{sr_i}, we have (see [2, p.328]):

- The random index r of the first node encountered is independent of the first encounter time X
- X is exponentially distributed, with parameter: $\Lambda_s = \lambda_{sd} + \sum_{i=1}^{n-2}(\lambda_{sr_i})$
- $\Pr(\textit{First node encountered is } r) = \frac{\lambda_{sr}}{\Lambda_s}$

This means that we can represent the first phase as independently identifying the encountered node (with probability $\frac{\lambda_{sr}}{\Lambda_s}$) and adding an exponential waiting time with parameter Λ_s. [1]

Two cases may arise: either the first node encountered r equals d, and s delivers the message with expected time $\frac{1}{\Lambda_s}$, or $r \neq d$ and node r waits to meet node d to deliver the message.

Let's evaluate the time it takes for r to meet d and deliver the message. If the message is received by node r at time t (let's say), its delivery time is equal to R^t_{rd}, the remaining inter-contact time before the next contact between nodes r and d. The memoryless nature of exponentials implies that R^t_{rd} follows an exponential distribution with the same parameter λ_{rd} as the inter-contact time. The mean expected delivery time for a message at node r awaiting delivery to d is thus given by:

$$E[D^w_{rd}] = 1/\lambda_{rd} \tag{2}$$

The total delivery time Z_r along path r, i.e., conditioned on using node r as a relay, is thus the sum of the first encounter time X and $E[D^w_{rd}]$ the remaining delivery time between nodes r and d and thus:

$$E[Z_r] = \frac{1}{\Lambda_s} + \frac{1}{\lambda_{rd}} \tag{3}$$

The total delivery time Z is computed by conditioning on all possible first encountered nodes $d, r_1, r_2, ..., r_{n-2}$, events whose probabilities are given by $\frac{\lambda_{sr}}{\Lambda_s}$.

After simplification, this leads to the following mean delivery time for 2-MH:

$$E[D^{2-MH}_{sd}] = \frac{(1 + \sum_{r \neq s, r \neq d} \frac{\lambda_{sr}}{\lambda_{rd}})}{\sum_{r \neq s} \lambda_{sr}} \tag{4}$$

C. *Optimal* two hop relay *strategy*

The minimum delay routing strategy is based on the previous 2-MH scheme. We derive 2-MH* which transfers the message only to a *subset* of neighbors of the source that minimizes the expected delivery time in case of independent pairwise exponential inter-contacts.

Instead of considering all neighbors of source node s as candidate relays, as in the 2-MH scheme, let's consider that the source node s forwards the message only to nodes in a subset R. We call this a 2-MH^R scheme. Following the same line of reasoning as in Sec. II-B, and defining $1/\lambda_{dd} = 0$, one finds that the expected delivery time is given by:

$$E[D^{2-MH^R}_{sd}] = \frac{(1 + \sum_{r \in R} \frac{\lambda_{sr}}{\lambda_{rd}})}{\sum_{r \in R} \lambda_{sr}} \tag{5}$$

We define 2-MH* to be a 2-MH^R scheme which uses a subset R that minimizes $E[D^{2-MH^R}_{sd}]$.

Brute force minimization amounts to testing all subsets R of neighbors of source node s. The complexity of the algorithm is exponential in the degree d_s of node s. The structure of Eq. 5 allows for the definition of an algorithm which is linear in d_s (see Sec. VI-A). To find the subsets R of neighbors of source node s that minimize $E[D^{2-MH^R}_{sd}]$ (Eq. 5), we propose the following algorithm:

for *every destination d* **do**
  Sort its neighbors in increasing mean inter-contact times, in which case we have: $0 \leq \frac{1}{\lambda_{1d}} \leq \frac{1}{\lambda_{2d}} \leq ... \leq \frac{1}{\lambda_{nd}}$
  Initialise the result set $I = \oslash$ and corresponding minimal mean delivery time (using set I) $c_I = \frac{1}{\lambda_{1d}}$
  for $i = 1, ..., n$ **do**
    Add node i to set I and compute $E[D^{2-MH^I}_{sd}]$ (as in Eq. 5)
    If this value is strictly larger than c_I, remove node i from I and stop
    Otherwise, place this value in c_I
  end
end

At the end, the optimal set of nodes is found in I and the corresponding minimal delay in c_I. Proof of the algorithm is provided in Sec. VI-A.

[1] The decoupling of mean waiting time and identity of the encountered node is the key to the derivation of the scheme. In the case of the *Marks Memoryless* class, this result is provided by Wald's Lemma.

D. Multi-hop relay strategy

In this section, we show that recursively applying the 2-MH* scheme leads to a fixed point that minimizes the delay in the case of an arbitrary number of intermediate relay nodes.

Let's introduce one more hop in the single relay scheme 2-MH*. Node s now chooses the best first encountered neighbors based on the assumption that they will relay the messages following the 2-MH relay scheme. For a given set of first encountered neighbors R, the total expected delay is given by:

$$E[D_{sd}^{3-MH^R}] = \frac{(1+\sum_{r\in R}\lambda_{sr}E[D_{rd}^{2-MH}])}{\sum_{r\in R}\lambda_{sr}} \quad (6)$$

minimizing Eq. 6 for all sets R of neighbors of s is obtained by applying the same algorithm as for 2-MH, since Eq. 6 is deduced from Eq. 5 by replacing $\frac{1}{\lambda_{rd}}$ by $E[D_{rd}^{2-MH}]$. The latter value represents an improved expected delivery time to the destination. Intuitively this increases the attractiveness of routes going through node r, thus increasing its chances of being added to the list of relaying neighbors.

Gradually introducing further relaying steps amounts to recursively applying the process. The sequence of values $E[D_{sd}^{2-MH}], E[D_{sd}^{3-MH}], ..., E[D_{sd}^{n-MH}]$ thus created is decreasing and positive (see Sec. VI-B), so it converges to, let's say, $E[D_{sd}^{MH*}]$. $E[D_{sd}^{MH*}]$ is necessarily attained in a finite number of steps (since there are only a finite number of possible intermediate nodes) and is a fixed point for the recursive process. Because the set R_{sd}^* realises the fixed point, the forwarding strategy simply amounts, for node r, to relaying any message with destination d to any first encountered neighbor in R_{rd}^*.

One can establish the following two properties of loop-free forwarding and polynomially bounded convergence that make the scheme workable for routing in DTNs.

First *forwarding a message with MH* is loop free*. Remember that a message with destination d is transfered by any node i to the first node j in set R_{id}^* it meets. The reason why there is no loop is because transfers always go to nodes that are strictly closer to the destination. All next relay nodes j have an expected delivery time to d strictly lower than that of the current relay node i (see Sec. VI-C). This implies that for any route a bundle takes, the sequence of nodes it visits, $i, i_1, i_2, ..., i_p$, necessarily verifies $E[D_{id}^{MH*}] < E[D_{i_1d}^{MH*}] < ... < E[D_{i_pd}^{MH*}]$. If there were a loop, it would mean that one of the visited nodes $i_1, ..., i_p$ is i, so we would have $E[D_{id}^{MH*}] < ... < E[D_{i}^{MH*}]$, which is a contradiction.

Second *the routing algorithm of Sec. II-C has complexity of* $O(L.n^2.D)$. L is the diameter of the *binary connectivity graph* in which two nodes share a link whenever they have been in contact, n is the total number of nodes and D is the average node degree. On the data sets we considered, L is equal to 10 for Dartmouth, 5 for MIT, and 3 for iMote. The average node degree D is 60.5 in Dartmouth, 22.3 in MIT and 22.8 in iMote. As we can see D and L are very small compared to n^2 in the three data sets. More generally, worst case complexity is $O(n^4)$, but if connectivity graphs have scale free properties, as one can expect in large networks, we would have $L = log(n)$ and $D \ll n$, and the complexity would scale as $O(n^2.log(n))$.

III. Comparing routing protocols

This section looks at the routing behavior and performances of the protocols described above and presents the results of simulations we performed using mobility traces to study how they compare with some well known approaches.

A. Experimental data sets

We describe here the contexts in which the data sets we used have been collected. All of these data sets are publicly available in the CRAWDAD archive [1].

Dartmouth data

This connectivity data set has been inferred from traces collected in the Wi-Fi access network of Dartmouth College [12]. These traces were pre-processed by Song et al. [21]. They track users' sessions in the wireless network, noting the time at which nodes associate and dissociate from access points. Although the Dartmouth data is not from a DTN network, they are perhaps the richest data set publicly available that tracks users in a campus setting. Jones et al. [15], Leguay et al. [18], and Chaintreau et al. [6] have recently used these traces in a similar way.

We only consider the subset of users who were present in the network every day between January 26$^{\text{th}}$ 2004 and March 11$^{\text{th}}$ 2004, an academic period during which we expect nodes' activity to be fairly stationary. This data set contains 834 users, or nodes. A few judicious assumptions are required to adapt the Dartmouth data for DTN studies. First, we assume that two nodes are in contact if they are attached at the same time to the same access point (AP). Then, we filter the data to remove the well known *ping-pong* effect. Wireless nodes, even non-mobile ones, can oscillate at a high frequency between two APs. To counter this, we filter all the inter-contact times below 1,800 seconds (30 minutes). Note that defining better filtering methods, although challenging, would be of interest for the community. As this is not the purpose of this work, we choose here the threshold that Yoon et al. [25] used for the same purpose. We use this inferred data set for the remainder of this paper.

iMote data

Chaintreau et al. [6] used iMotes (Bluetooth contact loggers from Intel) to acquire proximity contacts that occurred between participants in the student workshop at the *Infocom 2005* research conference. Students were asked to carry one of these sensors in their pocket at all times. Due to Bluetooth's short range, authors logged instances when people were close to each other (typically within 10 meters). They collected data from 41 iMotes over 3 days. The devices performed Bluetooth inquiry scans every 2 minutes. For each pair of nodes (i, j), we considered that i and j were in contact if either one saw the other.

MIT data

The Reality Mining experiment [8] conducted at MIT captured proximity, location, and activity information from 97 subjects (mainly students) over the course of an academic year. Each participant had an application running on their mobile phone to record proximity with others through periodic Bluetooth scans (every 5 minutes) in a similar fashion to that of the iMote experiment. Locality information comes from knowing which GSM network cell the phone is attached to. We only make use of the Bluetooth proximity data to determine whether two nodes were in contact. We selected 95 days of data corresponding to the first semester of the academic year 2004-2005 where activity was high in the traces in terms of the number of phones that collected data and the number of contacts that were recorded.

We will refer to these data sets as *Dartmouth*, *iMote* and *MIT*.

B. Routing protocols considered

We simulate the following protocols:

- *Wait*: The source node waits to encounter the destination to transfer the message. The main advantage of the scheme, also known as direct transmission, is to perform only one transmission per message.
- *2-MH*: this is the *two hop relay* scheme of Grossglauser and Tse [11]. The source gives the message to the first encountered node. If it is not the destination, this node is used as a relay and it keeps the message until it encounters the destination.
- *2-MH**: this scheme is similar to 2-MH but relays are only chosen among the set of nodes R that minimizes Eq. 5, as seen in Sec. II-C.
- MED (Minimum Expected Delay): this scheme was introduced by Jain et al. [13]. The strategy, similar to source routing, defines which path the message will follow from s to d, that is, the ordered list of intermediate relay nodes it will have to go through. The list is chosen to provide minimum expected end-to-end delay. Each relay node, upon receiving the message, will not be free to choose the next relay: it will have to follow the initial plan. Finding the optimal path thus amounts to finding a lowest-weight path between nodes s and d in a graph in which the weight on each link (i, j) is defined as $1/\lambda_{ij}$. Dijkstra's algorithm is used.
- *MH**: this is the fixed point opportunistic routing strategy that we introduced in Sec. II-D.
- *Epidemic*: each time two nodes meet, they exchange their messages. The algorithm provides the optimal path and thus the minimum delay.

We slightly modified 2-MH, to better compare it with 2-MH*: a node i is a potential relay only if $\lambda_{id} > 0$, i.e., if it has a chance of meeting the destination. In MED, we authorized intermediary relays to directly transfer messages to the destination whenever met.

In each of the simulation series, we choose at random 100 different source destination pairs (s, d) and replay the contacts between nodes present in the data to see how, for each pair, a message, generated at the beginning of the two months period, is delivered. For each data set, because of computational issues, we used a different constant message generation rate between source destination pairs.

We have implemented a stand alone simulator to evaluate the routing scheme. This simulator only implements the transport and network layers and it makes simple assumptions regarding lower layers, allowing infinite bandwidth between nodes and contention free access to the medium. Nodes are also supposed to have infinite buffers and to have inherent knowledge of all other nodes' mobility patterns. Because in ambient networks, nodes may have limited resources and capabilities, routing solutions should also be evaluated with limited buffers and more realistic models for the MAC and physical layers. One way in which we address the problem of limited resources is to examine, in Sec. IV-D, the possibility of limiting the amount of information that is sent regarding nodes' connectivity patterns. However, our aim here is principally to validate our routing proposition. We leave to future work a detailed study of the modifications that would be required to accommodate resource limitations.

The λ values used for route selection in 2-MH*, MH* and MED, and to determine theoretical delays for Wait, 2-MH*, MH* and MED are computed over the whole data set in a preliminary step. The inter-contact time averages are estimated from the data for all $n(n-1)/2$ pairs of nodes, and the λ values are set to the inverse of these averages.

C. Simulation results

This section presents results for the different routing strategies on each of the data sets.

Dartmouth

In Dartmouth, the λ values are computed over the filtered data set to avoid the bias introduced by the ping-pong effect (see Sec. III-A). The simulations replayed the original contacts and the messages between source-destination pairs were generated every 20 days.

	Del. (%)	**A. delay** (days)	**M. delay** (days)	**Th. delay** (days)	**A. hops** (#)	**Overhead** (trans.)
Wait	8.6 ±1.0	12.2 ±2.7	7.2 ±4.4	11.9 ±3.1	1.0 ±0.0	25.8 ±3.1
2-MH	57.4 ±2.0	16.5 ±0.7	14.0 ±1.6	-	1.9 ±0.0	427.8 ±15.3
2-MH*	61.4 ±1.1	13.5 ±0.6	10.0 ±0.9	8.4 ±0.6	1.9 ±0.1	416.8 ±12.8
MED	34.2 ±1.2	17.9 ±1.0	15.2 ±1.8	1.0 ±0.1	6.1 ±0.2	724.8 ±20.4
MH*	82.4 ±1.4	7.8 ±0.4	4.3 ±0.3	1.4 ±0.1	5.7 ±0.1	1993.6 ±793.4
Epidemic	99.0 ±0.8	1.0 ±0.2	0.9 ±0.0	-	9.8 ±0.2	123851 ±3687.8

TABLE I
SIMULATION RESULTS WITH DARTMOUTH DATA.

Table I presents the simulation results averaged over 5 runs with 90% confidence levels that are obtained using the Student t distribution. It presents, for each of the protocols, the average delivery ratio, the average delay ("A delay") and the median delay ("M delay") computed over the delivered messages, the average theoretical delay over all the messages generated (infinite delay is assumed to be the length of the simulated period, i.e., 45 days), and the average hop count, also obtained on delivered messages. We also measured the protocol overhead, considering the total number of transmissions that occurred before message delivery (or nondelivery for those that never reached their destination).

Wait and Epidemic are the two extreme schemes that we simulated. They respectively deliver 8.6% and 99.0% of messages with a mean delay of 12.2 and 1.0 days and with a median delay of 7.2 and 0.9 days. Wait only delivers 8.6% of messages because most of the source-destination pairs, selected at random, satisfy $\lambda_{sd} = 0$ (i.e., they never met). Wait only involves 1.0 hop while Epidemic attains a high average hop-count of 9.8. Naturally, Epidemic plots the highest overhead with 123,851 transmissions in total while Wait only realizes 25.8 transmissions.

2-MH and 2-MH*, which are the two one-relay algorithms that we simulated, deliver respectively 57.4% and 61.4% of messages with an average delay of 16.5 and 13.5 days. 2-MH* outperforms 2-MH while only requiring 416.8 transmissions instead of 427.8 on average. 2-MH gives the message to the first node it encounters while 2-MH* may be more selective, as it uses only a subset of its neighbors as relays. The comparison shows that the strategy used by 2-MH* of minimizing Eq. 5 allows to reduce delivery delay and to increase delivery ratio.

MH* delivers more messages than MED (82.4% of messages are delivered against 34.2%), and does it faster (average delay of 7.8 days against 17.9 days). MH* has performance close to that of Epidemic in delivery ratio while only involving 1,993 transmissions. The hop-by-hop opportunistic nature of MH* is the main reason for its superiority over MED, in which messages follow a strict sequence of relays. A node cannot take advantage of an opportunistic contact with a node that has a lower cost path than does the predesignated next hop node. This weakness has already been mentioned by Jain et al. [13] and MH* overcomes it.

Table I shows a discrepancy between the theoretical and the experimental delays. This can be explained by the presence of node pairs that do not have an exponential behavior. Inter-contacts following distributions with fatter tails than the exponential, a likely event, result in increased average delay and contribute to explaining the underestimates we observe. This is particularly true for 2-MH*, MH* and MED that should show average theoretical delays of respectively 8.4, 1.4 and 1.0 days while they achieve 13.5, 7.8 and 17.9 days. In this case the computation of expected delays on mean inter-contact times can also miss possible inter-dependencies of node contacts.

iMote

In simulations with the iMote data set, we generated messages between source destination pairs every 5 hours. Table II shows the simulation results.

	Del. (%)	**A. delay** (h)	**M. delay** (h)	**T. delay** (h)	**Hops** (#)	**Overhead** (trans.)
Wait	81.9 ±2.8	10.5 ±0.6	7.2 ±0.3	5.3 ±0.5	1.0 ±0.0	1146.6 ±39.6
2-MH	83.5 ±1.2	10.6 ±0.6	7.5 ±0.6	-	1.9 ±0.0	2476.4 ±16.6
2-MH*	87.2 ±1.3	9.0 ±0.6	6.3 ±0.5	2.0 ±0.1	1.7 ±0.0	2255.0 ±34.8
MED	82.1 ±3.4	10.3 ±0.5	7.3 ±0.1	2.8 ±0.1	1.3 ±0.0	1669.6 ±31.2
MH*	88.3 ±1.4	8.6 ±0.6	6.1 ±0.7	1.7 ±0.1	2.7 ±0.1	3644.2 ±96.6
Epidemic	91.8 ±1.3	6.5 ±0.4	4.2 ±0.3	-	4.1 ±0.1	27470.6 ±950.8

TABLE II

SIMULATION RESULTS WITH IMOTE DATA.

We first observe that the delivery ratios are closer to each other varying from 81.9% for Wait and to 91.8% for Epidemic. The fact that Wait delivers a large number of messages is another illustration of the high level of interactions that occurred between participants. We observe similar results to those with Dartmouth in ranking of protocol performance.

MIT

In the simulations we performed on MIT, messages between sources and destinations were generated every 15 days. Table III shows the simulation results.

	Del. (%)	**A. delay** (days)	**M. delay** (days)	**T. delay** (days)	**Hops** (#)	**Overhead** (trans.)
Wait	35.6 ±3.6	15.0 ±2.0	4.9 ±1.6	9.15 ±1.2	1.0 ±0.0	249.4 ±25.4
2-MH	67.7 ±2.4	11.2 ±0.5	0.8 ±0.6	-	1.8 ±0.1	1185.6 ±15.5
2-MH*	88.0 ±1.1	10.0 ±0.7	2.3 ±0.6	3.6 ±0.2	1.8 ±0.1	1080.2 ±14.5
MED	46.6 ±4.0	14.6 ±1.0	3.2 ±0.8	3.0 ±0.1	1.5 ±0.1	633.8 ±39.7
MH*	96.4 ±0.3	5.0 ±0.4	0.1 ±0.1	2.2 ±0.1	2.8 ±0.1	1994.6 ±65.5
Epidemic	99.0 ±0.2	1.4 ±0.4	0.1 ±0.1	-	2.5 ±0.1	50344.6 ±897.7

TABLE III

SIMULATION RESULTS WITH MIT DATA.

Results are closer to the ones we obtained with Dartmouth. Furthermore, we observe similar ranking of protocol performance to those with Dartmouth and iMote.

Through all these simulations, we validate the natural sense that we should take into account the heterogeneity of average inter-contact times in the design of routing solutions for DTNs and we show that MH^* achieves good performance in terms of delivery ratio, delay and overhead.

Generally speaking MH^* provides the most significant performance improvements for the Dartmouth and MIT data sets indicating that the scheme appears to be better suited for scenarios where connectivity is sparse.

IV. DISCUSSION

This section discusses the specific factors that could have impacted the results, the assumptions on which is built MH^* and some implementation choices.

A. Impact of traffic generation

The results that we presented show performance that we believe to be underestimated because of the way we generated traffic. In our simulations, as we did not have any knowledge of social relationships between participants, we selected source destination pairs at random and generated traffic with a constant rate. However, in a real deployment of DTN applications, we conjecture that those two parameters would be highly driven by social relationships (most of people would only communicate with friends with who they might also have a high level of interactions) and environmental factors such as specific events or periodic schedules.

Furthermore, as in the Internet, we expect congestions in DTNs. While we did not address this issue in this work, congestion control could be integrated to MH* in two ways. First, one could change the weights (e.g., λ values) so that they would be correlated to inter-contact times and to the levels of congestion of links. The second way would be to slightly change the forwarding algorithm. Instead of selecting the first eligible node met as a relay, one could decide to proceed or not with message transfers depending on traffic conditions. This way, MH* would be used as a means to obtain an interesting subset of paths that the data will follow or not depending on contact opportunities and traffic load. In the first approach, we could still calculate an estimation of the time needed to deliver messages but one should note that weights would have to be updated in the network which might not be practical.

B. Data sets used

The data sets may represent partial or biased real life interactions as sampling methods were used for their collection. The iMote and MIT data sets have been collected using periodic Bluetooth scans which may have underestimated the overall number of contacts or the contact times between nodes. In Dartmouth, we infer that two persons are in contact whenever they are connected to the same AP which might create unrealistic interactions, and more generally mobility of laptops is not really representative of human mobility. More accurate data sets are needed for DTN protocol evaluations.

In real systems, we also expect that these inaccuracies in the sampling of real life interactions could be due to usage of handheld devices (e.g., phones that are turned off, lack of batteries). Such biases in the calculation of λ values might mislead MH* in making the right routing decisions. However, we conjecture that MH* can tolerate a certain level of inaccuracies but we leave this study for future work.

C. Keeping copies at source nodes

An application that might send data over DTN networks would probably keep a copy of a sent message until it gets an acknowledgment asserting that the message has been correctly delivered. In that case, if the source node meets the destination, it would be able to transfer the message directly.

In order to study the impact of such a behavior, we performed simulations using exactly the same parameters and source destination pairs as in Sec. III-C. Table. IV presents the results with the three connectivity data sets. We can see that keeping one copy at the source almost preserves the relative order observed previously. As expected, it improves overall performances, but only slightly. For instance, in Dartmouth, 2-MH, 2-MH*, MED and MH* deliver respectively 58.3%, 61.6%, 34.8% and 82.7% of messages instead of 57.4%, 61.4%, 34.2% and 82.4%.

D. Overhead reduction

Handling information on contact patterns for MH* could lead to high processing and network overhead even if only summary information such as the λ values is used. Nodes would have to perform tasks such as monitoring the inter-contact times they have with the others, disseminating this information to the other nodes (using a centralized architecture or not) and computing periodically the sets of relays that have

	Dartmouth		**iMote**		**MIT**	
	Del. (%)	**A. delay** (days)	**Del.** (%)	**A. delay** (h)	**Del.** (%)	**A. delay** (h)
2-MH	58.3 ±1.9	16.0 ±0.8	88.9 ±1.4	8.8 ±0.5	73.8 ±2.6	10.5 ±0.6
2-MH*	61.6 ±1.1	13.5 ±0.6	88.7 ±1.3	8.7 ±0.6	88.3 ±1.3	9.7 ±0.8
MED	34.8 ±1.2	17.9 ±1.0	84.5 ±3.0	10.0 ±0.4	48.7 ±3.7	15.1 ±1.2
MH*	82.7 ±1.7	7.8 ±0.2	89.5 ±1.3	8.1 ±0.5	96.6 ±0.3	4.8 ±0.4

TABLE IV

SIMULATION RESULTS WHEN A COPY IS KEPT AT THE SOURCE.

to be used for forwarding. We let this detailed analysis for future work, but study here the impact of reducing the amount of information distributed among nodes. In the scenario nodes only disseminate λ values satisfying $1/\lambda < L$. Table V shows the simulation results obtained with the same parameters and source destination pairs as in Sec. III-C on iMote and MIT data sets.

iMote				**MIT**			
L (h)	**Del.** (%)	**A. delay** (h)	**M. delay** (h)	**L** (h)	**Del.** (%)	**A. delay** (days)	**M. delay** (days)
1	81.9 ±2.8	10.5 ±0.6	7.2 ±0.4	**1**	35.6 ±3.6	15.0 ±2.0	4.8 ±1.6
2	86.2 ±1.5	8.9 ±0.5	6.4 ±0.4	**24**	48.1 ±3.4	10.3 ±0.9	2.6 ±0.6
5	87.3 ±1.6	8.5 ±0.5	6.1 ±0.6	**36**	68.4 ±2.2	6.2 ±0.7	1.3 ±0.3
8	87.5 ±1.7	8.6 ±0.5	6.2 ±0.6	**72**	84.2 ±0.9	5.1 ±0.3	0.6 ±0.3
10	88.3 ±1.4	8.7 ±0.6	6.3 ±0.6	**168**	95.3 ±0.1	5.7 ±0.3	0.4 ±0.1
∞	88.3 ±1.4	8.6 ±0.6	6.1 ±0.7	∞	96.4 ±0.3	5.0 ±0.4	0.1 ±0.1

TABLE V

SIMULATION RESULTS OF MH* WITH PARTIAL KNOWLEDGE.

We can see that, as expected, as we increase the threshold L, performance are closer to those observed in Sec. III-C denoted by $L = \infty$ here. The value of L for which performances are reasonably degraded is 10 hours in iMote and 168 hours in MIT leading to a reduction of the shared routing information of respectively 9.8% and 35.2%. These figures depend on the overall density of interactions. Because in iMote node interactions are more homogeneous we are not able to reduce the overhead as we could do in MIT. We expect this reduction to be much higher in Dartmouth data but we were not able to perform simulations for computational reasons. This result is promising regarding the scalability of routing algorithms that would involve summary information on pairwise contacts such as average inter-contact times.

E. Complexity of inter-contact times processes

Furthermore, evaluating schemes that use summary information such as the average inter-contact times on real data have to deal with two factors: the presumed stationarity of inter-contact processes and the short and long terms dependencies in interactions between nodes. As a consequence, average values might not be sufficiently precise because processes are not stable over time and their burstiness is not well accounted for. As an illustration, we have seen that the theoretical value of delivery delay for MH* underestimates the values observed from the simulations on real data sets. This indicates that the hypothesis of pairwise independent exponential inter-contacts processes is not sufficiently faithful to model accurately the actual behaviors of the nodes.

Simulation artifacts also come into play. The routing simulation is carried out on a limited time scale. The λ values are computed over the entire data set in a prior pass, so a relay node may meet the destination for the last time before having met the source for the first time. A more realistic estimation could use on-line predictive or learning methods. However, as they are challenging to define, we let this study for future work and intend here to provide early validation results to motivate research in the domain. MED clearly suffers from this simulation artefact, it would deliver more messages otherwise.

V. Conclusion and future work

In this paper, we have presented a new single copy and multi-hop routing strategy, MH*, which waits for only a subset of relays at each hop to improve routing performance, measured in terms of average delay. The scheme uses as only input the estimates of the average inter-contact times between the nodes in the network. Defined as the fixed point of a recursive process, it provides the minimum delivery time in case of independent exponential pairwise inter-contacts. It is loop-free and convergences polynomially, which make the scheme workable for routing in DTNs. We show, by replaying real connectivity traces, that MH* achieves good performance, in terms of delivery ratio and delay, while keeping the overhead low. We also discussed factors and implementation issues that might have impacted the results.

Future work along these lines might include formal and simulation studies to elaborate more complex schemes in terms of number of copies distributed and knowledge considered to make routing decisions. Indeed, single-copy protocols suffer clearly in DTNs from reliability issues as nodes can disappear altogether at any time for various reasons. Redundancy approaches have then to be considered. There are several ways in which MH* could work with multiple copies. One could for instance use a utility function in the forwarding process to decide how many copies should be transmitted and to which nodes in the eligible set. Also of interest, similar to multi-copy approaches introduced by Leguay et al. [19] and Spyropoulos et al. [23], messages could be first spread to some nodes and then routed independently.

Furthermore, more realistic evaluations and modelling would be needed to better take into account interactions between nodes (e.g., considering realistic contact durations) and the limitations inherent to real systems such as buffers size and available bandwidth. This work would allow studying a number of associated mechanisms to improve performance such as transmission scheduling, buffer management and congestion control.

References

[1] CRAWDAD: A community resource for archiving wireless data at dartmouth. `http://crawdad.cs.dartmouth.edu`.

[2] P. Bremaud. *Markov Chains, Gibbs Fields, Monte Carlo simulation, and queues*. Springer, 1999.

[3] J. Burgess, B. Gallagher, D. Jensen, and B. N. Levine. MaxProp: Routing for Vehicle-Based Disruption-Tolerant Networks. In *Proc. IEEE INFOCOM*, 2006.

[4] B. Burns, O. Brock, and B. N. Levine. MV routing and capacity building in disruption tolerant networks. In *Proc. Infocom*, 2005.

[5] I. Carreras, D. Miorandi, and I. Chlamtac. A framework for opportunistic forwarding in disconnected networks. In *Proc. of MOBIQUITOUS*, 2006.

[6] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott. Impact of human mobility on the design of opportunistic forwarding algorithms. In *Proc. INFOCOM*, 2006.

[7] X. Chen and A. L. Murphy. Enabling disconnected transitive communication in mobile ad hoc networks. In *Proc. of Workshop on Principles of Mobile Computing*, 2001.

[8] N. Eagle and A. Pentland. Reality mining: Sensing complex social systems. *Personal and Ubiquitous Computing*, 2005.

[9] K. Fall. A delay-tolerant network architecture for challenged internets. In *Proc. SIGCOMM*, 2003.

[10] R. Groenevelt, P. Nain, and G. Koole. The message delay in mobile ad hoc networks. In *Posters ACM SIGMETRICS*, 2005.

[11] M. Grossglauser and D. Tse. Mobility increases the capacity of ad-hoc wireless networks. *Transactions on Networking*, 10(4):477–486, August 2002.

[12] T. Henderson, D. Kotz, and I. Abyzov. The changing usage of a mature campus-wide wireless network. In *Proc. MobiCom*, 2004.

[13] S. Jain, K. Fall, and R. Patra. Routing in a delay tolerant network. In *Proc. SIGCOMM*, 2004.

[14] A. Jindal and K. Psounis. Fundamental mobility properties for realistic performance analysis of intermittently connected mobile networks. In *Proc. of Pervasive Computing and Communications Workshop on International Workshop on Intermittently Connected Mobile Ad hoc Networks*, 2007.

[15] E. P. C. Jones, L. Li, and P. A. S. Ward. Practical routing in delay-tolerant networks. In *Proc. WDTN*, 2005.

[16] E. P.C. Jones and P. A.S. Ward. Routing strategies for delay-tolerant networks. In *Submitted to Computer Communication Review*, 2006.

[17] T. Karagiannis, J-Y. Le Boudec, and Milan Vojnovic. Power law and exponential decay of inter contact times between mobile devices. In *Proc. ACM MOBICOM*, 2007.

[18] J. Leguay, T. Friedman, and V. Conan. Evaluating mobility pattern space routing for DTNs. In *Proc. INFOCOM*, 2006.

[19] J. Leguay, T. Friedman, and V. Conan. Evaluating mobyspace based routing strategies in delay tolerant networks. In *WCMC Special Issue on Disruption Tolerant Networking*, 2007.

[20] A. Lindgren, A. Doria, and O. Schelen. Probabilistic routing in intermittently connected networks. In *Proc. SAPIR*, 2004.

[21] L. Song, D. Kotz, R. Jain, and X. He. Evaluating location predictors with extensive Wi-Fi mobility data. In *Proc. Infocom*, 2004.

[22] T. Spyropoulos, K. Psounis, and C. Raghavendra. Single-copy routing in intermittently connected mobile networks. In *Proc. IEEE SECON*, 2004.

[23] T. Spyropoulos, K. Psounis, and C. Raghavendra. Spray and focus: Efficient mobility-assisted routing for heterogeneous and correlated mobility. In *Proc. of Pervasive Computing and Communications Workshop on International Workshop on Intermittently Connected Mobile Ad hoc Networks*, 2007.

[24] UMassDieselNet. A Bus-based Disruption Tolerant Network, `http://prisms.cs.umass.edu/diesel/`.

[25] J. Yoon, B. Noble, M. Liu, and M. Kim. Building realistic mobility models from coarse-grained traces. In *Proc. MobiSys*, 2006.

VI. Annex

A. Result 1

For an exponential DTN of n nodes and with parameters (λ_{ij}), and for any matrix (ε_{ij}) which represents the estimated message travel time between node i and node j (given a certain routing policy), introducing an intermediate relay selected among a subset I of first encountered neighbor nodes of the source s, generates, for the delivery of a message to node d, an expected travel time $C(I)$, given by (see Sec.II-D):

$$C(I) = \frac{(1 + \sum_{r \in I} \lambda_{sr} \varepsilon_{rd})}{\sum_{r \in I} \lambda_{sr}} \tag{7}$$

Let I_{min} be one subset that minimizes $C(I)$ for all subsets I of neighbors of s. We consider without loss of generality that $\varepsilon_{1d} \leq \varepsilon_{2d} \ldots \leq \varepsilon_{nd}$.

We are going to establish the following, quite remarkable, result on the structure of the minimal set I_{min}:

THEOREM. If I_{min} is a subset of neighbor nodes of s that minimize Eq. 7, then either $I_{min} = \oslash$ or there is a $p, 1 \leq p \leq n$, for which $I_{min} = [1, 2, ..., p]$.

This result is derived from the special shape of criterion of Eq. 7. More precisely, we will need the following lemma:

LEMMA ϕ. Let's introduce the bivariate function $\phi(x,y) = \frac{b+xy}{a+x}$ (compare it to Eq. 7 to see how it comes into play), we have:

$$\forall b > 0,\ a > 0 \text{ and } \forall x,y \geq 0,\ \phi(x,y) \leq \frac{b}{a} \iff y \leq \frac{b}{a} \quad (8)$$

PROOF. This is straightforward to check.

Let's now proceed to the proof the main result.

PROOF. Consider neighbor nodes of s and whether they ever meet destination d or not:

i) if none of the neighbors of s ever see destination d, this means that using any of them as a relay introduces infinite delivery time, criterion (Eq. 7) becomes infinite. In other words none of the neighbors of s are valid relaying candidates, so $I_{min} = \oslash$
ii) if at least one of the neighbors of s sees d, there exists a node with index m such that $\varepsilon_{md} < \infty$ so $I_{min} \neq \oslash$.

Let p be the index of the largest ε_{id} for nodes i in set I_{min}. We are going to show that all nodes i which satisfy $\varepsilon_{id} \leq \varepsilon_{pd}$ also belong to I_{min}.

Let's note ratio $C(I_{min})$ by $\frac{b}{a} = \frac{(d+\lambda_{sp}\varepsilon_{pd})}{c+\lambda_{sp}}$. Since I_{min} minimizes criterion in Eq. 7, $\frac{b}{a} \leq \frac{d}{c}$: the second term represents the value of the criterion for I_{min} minus p, which is by definition of I_{min} suboptimal.

Rewriting the inequality $\frac{(d+\lambda_{sp}\varepsilon_{pd})}{c+\lambda_{sp}} \leq \frac{d}{c}$, and from the property of ϕ in Eq. 8, we then have $\varepsilon_{pd} \leq \frac{d}{c}$. Now:

$$\varepsilon_{pd} \leq \frac{b}{a} \iff \varepsilon_{pd} \leq \frac{(d+\lambda_{sp}\varepsilon_{pd})}{c+\lambda_{sp}} \quad (9)$$
$$\iff (c+\lambda_{sp})\varepsilon_{pd} \leq d+\lambda_{sp}\varepsilon_{pd} \iff \varepsilon_{pd} \leq \frac{d}{c}$$

Suppose there exists a node m such that $\varepsilon_{md} \leq \varepsilon_{pd}$ and $m \notin I_{min}$. Let's add it to set I_{min}, and consider the value of the criterion for this new set of neighbors of s, $I' = I_{min} \cup m$, $C(I') = \frac{(b+\lambda_{sp}\varepsilon_{md})}{a+\lambda_{sm}}$ which is lower than or equal to $\frac{b}{a}$ (from the property of ϕ in Eq. 8 and the fact that $\varepsilon_{md} \leq \varepsilon_{pd} \leq \frac{b}{a}$); I' would then perform better than I_{min} in minimizing the criterion, which is in contradiction with the definition of I_{min}.

In other words, all (reordered) nodes 1 through p belong to set I_{min}, which provides the announced result. This further leads to the linear time algorithm for minimizing the criterion: once sorted in the appropriate order, it suffices to add each node one after the other and stop when the criterion does not diminish anymore.

B. Result 2

We are going to show that for a given source destination pair s, d, the sequence of expected delivery times $E[D_{sd}^{n-MH^*}]$ for the strategy with n relays decreases as the number of relaying steps n increases.

Let's first introduce some notations. For a given destination d, let's consider, for all source nodes s, the sequence of values $E[D_{sd}^{n-MH^*}]$, defined recursively by:

$$\forall s \neq d,\ E[D_{sd}^{1-MH^*}] = \frac{1}{\lambda_{sd}} \text{ and } E[D_{dd}^{1-MH^*}] = 0 \text{ and} \quad (10)$$
$$\forall s, \forall n > 1, E[D_{sd}^{n-MH^*}] = Min_{R \subset P(n)}\left(\frac{(1+\sum_{r\in R}\lambda_{sr}E[D_{rd}^{(n-1)-MH^*}])}{\sum_{r\in R}\lambda_{sr}}\right)$$

THEOREM. The sequence $E[D_{sd}^{n-MH^*}]$ defined in Eq. 10 is decreasing, i.e.:

$$\forall n \geq 0,\ \forall s,\ E[D_{sd}^{(n+1)-MH^*}] \leq E[D_{sd}^{[n-MH^*}] \quad (11)$$

PROOF. We proceed by induction on the number of relays n.

For $n = 0$, we have: $\forall s, E[D_{sd}^{1-MH^*}] = \frac{1}{\lambda_{sd}} \in \Re \cup \infty$

Two cases may occur, depending on whether s meets d or not:

- If $\lambda_{sd} = 0$, $E[D_{sd}^{2-MH^*}] \leq E[D_{sd}^{1-M^*}] = \infty$.
- If $\lambda_{sd} \neq 0$, let's consider the one relay strategy for which R reduces to singleton d. Its delivery delay is given by $\frac{1}{\lambda_{sd}}$. By definition $E[D_{sd}^{2-MH^*}]$ gives a lower delay, so we have $E[D_{sd}^{2-MH^*}] \leq \frac{1}{\lambda_{sd}} = E[D_{sd}^{1-MH^*}]$.

So this proves the result in the case of $n = 1$.

Let's suppose that the result holds at rank $n-1$, $\forall s, E[D_{sd}^{n-MH^*}] \leq E[D_{sd}^{[n-1)-MH^*}]$.

Let's consider $E[D_{sd}^{(n+1)-MH^*}]$ for a given s,

By definition we have: $E[D_{sd}^{n-MH^*}] = \frac{(1+\sum_{r\in I_{min}^n}\lambda_{sr}E[D_{rd}^{(n-1)-MH^*}])}{\sum_{r\in I_{min}^n}\lambda_{sr}}$, for a given set I_{min}^n of neighbors of s.

If one uses this set of nodes when introducing another relay node (i.e., at rank $n+1$), the expected delay is higher than $E[D_{sd}^{(n+1)-MH^*}]$ (by definition), so we have:

$$E[D_{sd}^{(n+1)-MH^*}] \leq \frac{(1+\sum_{r\in I_{min}^n}\lambda_{sr}E[D_{rd}^{n-MH^*}])}{\sum_{r\in I_{min}^n}\lambda_{sr}} \quad (12)$$

But we have by hypothesis, $\forall i, E[D_{id}^{n-MH^*}] \leq E[D_{id}^{[n-1)-MH^*}]$, so this leads to:

$$E[D_{sd}^{(n+1)-MH^*}] \leq \frac{(1+\sum_{r\in I_{min}^n}\lambda_{sr}E[D_{rd}^{(n-1)-MH^*}])}{\sum_{r\in I_{min}^n}\lambda_{sr}} = E[D_{sd}^{n-MH^*}] \quad (13)$$

and this is true for all nodes s, i.e., $\forall s, E[D_{sd}^{(n+1)-MH^*}] \leq E[D_{sd}^{[n-MH^*}]$. This is the induction hypothesis at rank $n+1$. So the result follows by induction.

C. Result 3

We are going to show that a message with destination d is relayed by MH* to a node with lower expected delivery time to d, i.e., we have the following.

THEOREM. For any node s following the routing strategy MH*, we have:

$$\forall r \in R^*_{sd} E[D^{MH^*}_{rd}] < E[D^{MH^*}_{sd}] \tag{14}$$

PROOF. $E[D^{MH^*}_{sd}]$ is the fixed point of the $E[D^{n-MH}_{sd}]$ sequence, so it satisfies:

$$E[D^{MH^*}_{sd}] = \frac{(1 + \sum_{r \in R^*_{sd}} \lambda_{sr} E[D^{MH^*}_{rd}])}{\sum_{r \in R^*_{sd}} \lambda_{sr}} \tag{15}$$

Singling out a giving relay node r, and applying Lemma ϕ of Eq. 8, we have:

$$E[D^{MH^*}_{rd}] \leq E[D^{MH^*}_{sd}] \tag{16}$$

We now have to check that the inequality is strict. Singling out node r in Eq.15, we have:

$$E[D^{MH^*}_{sd}] = \frac{(d + \lambda_{sr} E[D^{MH^*}_{rd}])}{c + \lambda_{sr}} \tag{17}$$

It is straightforward to check that $E[D^{MH^*}_{sd}] = E[D^{MH^*}_{rd}]$ if and only if $E[D^{MH^*}_{sd}] = \frac{d}{c}$. But $\frac{d}{c}$ corresponds to the criterion with set of neighbor nodes of s R^*_{sd} minus r, which is in contradiction with the definition of R^*_{sd}. So the inequality is strict.

Vania Conan received an Engineering Degree (1990) and Ph.D. in Computer Science (1996) from Ecole des Mines de Paris, France, and a Masters Degree in Artificial Intelligence from Universit Pierre et Marie Curie, Paris (1991). He is presently senior research engineer at Thales Communications, and manager of collaborative research projects. His current research topics include ad hoc, delay tolerant networking and distributed system design.

Jérémie Leguay received a Master of Science in Computer Science at Linkoping University in Sweden. He worked on the description and simulation of Internet routes during his master thesis. He received a Ph.D. in 2007 from the Computer Science laboratory (LIP6) of Pierre & Marie Curie University and is woking now at Thales Communications where he conducts research in ad hoc and delay tolerant networking.

Timur Friedman received the A.B. degree in philosophy from Harvard University and the M.S. degree in management from Stevens Institute of Technology. He received the M.S. and Ph.D. degrees in computer science from the University of Massachusetts Amherst, in 1995 and 2001, respectively. He is currently a Matre de Confrences (assistant professor) of computer science at UPMC Paris Universitas, and a researcher at the LIP6 computer science laboratory. His research interests include delay tolerant networking and large scale network measurement systems