DTN Routing in a Mobility Pattern Space*

Jérémie Leguay
Univ. P. & M. Curie†, Thales‡

Timur Friedman
Univ. P. & M. Curie

Vania Conan
Thales

ABSTRACT

Routing in delay tolerant networks (DTNs) benefits considerably if one can take advantage of knowledge concerning node mobility. The main contribution of this paper is the definition of a generic routing scheme for DTNs using a high-dimensional Euclidean space constructed upon nodes' mobility patterns. We call this the *MobySpace*. One way of representing nodes in this space is to give them coordinates that correspond to their probability of being found in each possible location. We present simulation results indicating that such a scheme can be beneficial in a scenario inspired by studies done on real mobility traces. This work should open the way to further use of the virtual space formalism in DTN routing.

Categories and Subject Descriptors

C.2 [**Computer-Communication Networks**]: Network Architecture and Design

General Terms

Algorithms, Performance, Experimentation

Keywords

Delay Tolerant Networks, Routing, Mobility

1. INTRODUCTION

The novelty of the work presented here is that we transcribe the problem of routing in a delay tolerant network (DTN) [9] on the basis of mobility patterns into a problem of routing in a virtual space defined by those patterns. By doing so, we can bring the powerful formalism of Euclidean space to bear on the problem of DTN routing.

In one common DTN scenario, nodes are 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 a very low probability that an end-to-end path exists between a given pair of nodes at a given time. End-to-end paths can exist temporarily, or may sometimes never exist, with only partial paths emerging. Due to these disruptions, regular ad-hoc networking approaches to routing and transport do not work, and new solutions must be proposed.

The Delay Tolerant Network Research Group (DTNRG) [1] has proposed an architecture [5] to support messaging that may be used by delay tolerant applications in such a context. The architecture consists mainly of the addition of an overlay, called the bundle layer, above a network's transport layer. Messages transferred in DTNs are called bundles. They are transferred in an atomic fashion between nodes using a transport protocol that ensures node-to-node reliability. These messages can be of any size. Nodes are assumed to have large buffers in which they can store the bundles.

Routing is one of the very challenging open issues in DTNs, as mentioned by Jain et al. [11]. Indeed, since the network suffers from connectivity issues, MANET [7] routing algorithms such as OLSR, based on the spreading of control information, or AODV, which is on-demand, fail to achieve routing. Different approaches have to be found.

Epidemic routing is a possible solution when nothing is known about the behavior of nodes. Since it can lead to buffer overflows and inefficient use of transmission media, one would prefer to limit bundle duplication and instead use routing heuristics that can take advantage of context. In order to move in such a direction, the DTN architecture defines several types of contacts: scheduled, opportunistic, and predicted. *Scheduled* contacts can exist, for instance, between a base station somewhere on earth and a low earth orbiting relay satellite. *Opportunistic* contacts are created simply by the presence of two entities at the same place, in a meeting that was neither scheduled nor predicted. Finally, *predicted* contacts are also not scheduled, but predictions of their existence can be made by analyzing previous observations.

Some work has been done with scheduled contacts, such as the scheme described by Jain et al. [11] that tries to improve

*This work has been co-funded in part by Thales Communications and LiP6 through Euronetlab, and in part by the ANRT (Association Nationale de la Recherche Technique) through the CIFRE grant 135/2004.

†Laboratoire LiP6–CNRS, Université Pierre et Marie Curie
‡TAI Laboratory, Thales Communications

SIGCOMM'05 Workshops, August 22–26, 2005, Philadelphia, PA, USA.

the connectivity to the internet of an isolated village based on knowledge of when a low-earth orbiting relay satellite and a motorbike might be available to make the necessary connections. Also of interest, work by Akyildiz et al. [2] on inter-planetary networking uses predicted contacts, such as the ones between planets, within the framework of a DTN architecture. The case of only opportunistic contacts has been analyzed by Vahdat and Becker [18] using the epidemic routing scheme.

Most of the work concerning routing in DTNs has been performed with predicted contacts, such as the algorithm of Lindgren et al. [12], which relies on nodes having a community mobility pattern. Nodes mainly remain inside their community and sometimes visit others. To route a bundle to a destination, a node may transfer that bundle to a node that belongs to the same community as the destination. In a similar manner, Burns et al. [4] proposed a routing algorithm that uses past frequencies of contacts. Also making use of past contacts, Davis et al. [8] improved the basic epidemic scheme with the introduction of adaptive dropping policies. Recently, Musolesi et al. [14] have introduced a generic method that uses Kalman filters to combine and evaluate the multiple dimensions of the context in order to take routing decisions. The context is made up of measurements that nodes perform periodically, which can be related to connectivity issues, but not necessarily. This mechanism allows network architects to define their own hierarchy among the different context attributes.

The case study presented in this paper also relies on contacts that can be characterized as predicted, but the underlying idea is a more generic abstraction compared to previous work, being able to capture the interesting properties of major mobility patterns for routing. The main contribution of this paper is the use, for routing in DTNs, of the formalism of a high-dimensional Euclidean space based on nodes' mobility patterns. We show the feasibility of this concept through an example in which each dimension represents the probability for a node to be found in a particular location. We conduct a simulation that produces promising initial results for this concept.

Applying the formalism of Euclidean space to computer networking problems is not in itself a new idea. To our knowledge, however, it has not previously been used for DTN routing. Furthermore, we believe that the idea of constructing a virtual space based upon mobility patterns is new.

Previous work with Euclidean spaces for networking has included the geolocalization of internet hosts, as in the GeoPing technique developed by Padmanabhan and Subramanian [15]. The position of a host to localize in Euclidean space is compared to the position of well know landmark nodes in order to estimate the host's location. Round trip times are used to determine coordinates. As opposed to what we do here, this Euclidean space is not used for routing.

Euclidean spaces have also been exploited in peer-to-peer architectures, notably by Ratnasamy et al. for CAN [17], in order to construct a robust and scalable mechanism to handle search queries. In this case, the Euclidean space is a virtual space, in which keys describing files are assigned virtual coordinates, and each node in the system governs a portion of the space. Queries are routed from node to node in the direction of the key's coordinates.

The routing scheme presented in this paper is similar to the one of CAN, since messages are routed in a virtual space. As opposed to CAN, in our scheme there is no notion of neighbor in the virtual space. Nodes may be directly connected to nodes that are nearby or that are very far. And these connections arise and dissolve dynamically as a function of node mobility in the physical space. Nodes opportunistically take advantage of connections that promise to advance bundles toward the destination.

The rest of the paper is structured as follows. Sec. 2 describes the mobility pattern based routing scheme. Sec. 3 presents the simulation results. Sec. 4 concludes the paper, discussing directions for future work.

2. ROUTING IN A MOBILITY PATTERN SPACE

This section first presents the idea behind routing in a high-dimensional Euclidean space constructed upon mobility patterns of nodes and then shows how we applied this idea within the framework of a scenario inspired by real observations.

2.1 Concept

The Euclidean virtual space, or *MobySpace*, introduced here is a generalization of ideas that are already current in the DTN literature. The principle is to use a Euclidean space as a tool to help nodes to take routing decisions. These decisions rely on the notion that a node is a good candidate for taking custody of a bundle if it has a mobility pattern similar to that of the bundle's destination. Routing is done by forwarding bundles toward nodes that have mobility patterns that are more and more similar to the mobility pattern of the destination. Since in the MobySpace, the mobility pattern of a node provides its coordinates, its *MobyPoint*, routing is done by forwarding bundles toward nodes that have their MobyPoint closer and closer to the MobyPoint of the destination.

Several questions arise. What type of dimensions do we choose, how many, and what kind of range for values do we define? How do we define the notion of distance? Is straightforward Euclidean distance useful or are other similarity functions more appropriate? Is it possible to have an infinite space in terms of the number of dimensions? What might be the problems with such a scheme?

Note that the objective of this paper is not to answer all these questions. It is to introduce a new concept for routing and to examine some of the interesting problems that the concept presents. In this section, we describe a manner in which mobility patterns can be characterized and discuss other possible alternatives. The question of the choice of a similarity metric is addressed in Sec. 2.2.

2.1.1 *Mobility pattern characterization*

Several methods could be employed to describe a mobility pattern. For instance it could be based upon historic information regarding contacts that the node has already had. If we want to route a bundle from one node to another, we have an interest in considering information on these contacts as forming a virtual space. Each possible contact is an axis, and the distance along that axis indicates a measure of the probability of contact. Two nodes that have a similar set of contacts that they see with similar frequencies are close in this space, whereas nodes that have very different sets of contacts, or that see the same contacts but with very different frequencies, are going to be far. It seems reasonable that one would wish to pass a bundle to a node that is as close as possible to the destination in this space, because this should improve the probability that it will eventually reach the destination.

In the virtual contact space just described, knowledge of all axes of the space requires knowledge of all nodes that are circulating in the space. Note that a full knowledge of the axes might not be required for successful routing. Nonetheless, we might wish to consider an alternative space in which there is a fixed, or at least more limited and well known number of axes. If nodes' visits to particular locations can be tracked, then the mobility pattern of a node can be described by its visits to these locations. In this scenario, each axis represents a location, and the distance along the axis represents the probability of finding a node at that location. We can imagine that nodes that have similar probabilities of visiting a similar set of locations are more likely to encounter each other than nodes that are very different in these respects. This is the kind of MobySpace that we employ in the study described in Sec. 2.2.

2.1.2 Possible limits and issues

DTN Routing in a contact space or a mobility space is based on the assumption that there will be regularities in the contacts that nodes have or their choices of locations to visit. There is always the possibility that we may encounter mobility patterns similar to the ones observed with random mobility models. The efficiency of the virtual space as a tool may be limited if nodes too rapidly change their habits.

Some problems could occur even if nodes have well defined mobility patterns For instance, in the MobySpace, a bundle may reach a local maximum if a node has a mobility pattern that is the most similar in the local neighborhood to the destination node's mobility pattern, but is not sufficient for one reason or another to achieve delivery. In the second type of space, where each dimension represents a location, it can happen if nodes visit similar places, but for timing reasons, such as being on opposite diurnal cycles, they never meet.

Other issues surrounding the use of a virtual contact space or mobility space are discussed in Sec. 4.

2.2 A case study

Recent studies of the mobility of students in a campus [13, 10, 6] or corporate users [3] equipped with PDAs or laptops able to be connected to wireless access networks, show that they follow common mobility patterns. They show that significant aspects of the behavior can be characterized by power-law distributions. Specifically, the session durations and the frequencies of the places visited by users follow power laws. This means that users typically visit a few access points frequently while visiting the others rarely, and that users may stay at few locations for long periods while visiting the others for very short periods. Henderson et al. observed [10] that 50% of users studied spent 62% of their time attached to a single access point and this proportion decreased exponentially. If we take these wireless access network studies to be representative of a class of mobile node behavior, we can consider that these observations are applicable to at least certain DTN scenarios.

For this case study we propose the following mobility model. Let us consider a set of nodes that move among a set of N locations. Two nodes can communicate only if they are at the same location. Node movements are based on power-laws, and each node has a mobility pattern defined by the distribution of P. $P(i)$ is the probability for the node to be at location i and $P(i) = K\left(\frac{1}{d}\right)^{n_i}$ with n_i the preference index of location i, d the exponent of the power-law based mobility pattern and K a constant. $n_j = 0$ means that the location j is the preferred one. Because $\sum_i P(i) = 1$, we have:

$$K = \frac{1 - \frac{1}{d}}{1 - \frac{1}{d^N}} \qquad (1)$$

Under this model, d is the fundamental parameter governing node behavior. As shown in Fig. 1, when d is high, nodes tend to move among a very small subset of locations, having one that they strongly prefer to the others. As d approaches 1, the range of locations that nodes visit regularly becomes wider, while still presenting a hierarchy of preferences. When $d = 1$, we have equiprobability.

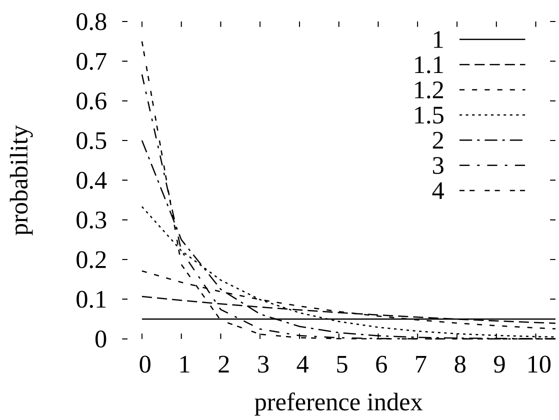

Figure 1: Power-law probability distributions for different values of d

Note that a mobility pattern should also characterize the time that a node remains at each location. We propose to capture this in the mobility model described here by uniformly distributing the resting time at each location over the interval $[t_{\min}, t_{\max}]$ with $t_{\min}$ and $t_{\max}$ quite close together, and by allowing nodes to randomly choose the same location consecutively.

For each of the nodes in this model, there is therefore a well defined probability of finding that node at each of the N locations. This set of probabilities is a node's mobility pattern, and is described by a MobyPoint, in an N dimensional MobySpace. We propose to route bundles in the MobySpace by sending them to nodes having mobility patterns that are successively closer to the mobility pattern of the destination. In simple words, we prefer to give custody of a bundle to a node that has mobility habits similar to those of the bundle's destination. In order to complete our model, we therefore require a similarity function that can be used to compare mobility patterns.

We study the following functions, each of which is a measure of similarity between two points in a Euclidean space:

- *Euclidean distance*: This is the most common distance measure. It returns the root of the sum of the square differences between the coordinates of a pair of points.

$$d_{ij} = \sqrt{\sum_{k=1}^{n} (x_{ik} - x_{jk})^2} \quad (2)$$

- *Canberra distance*: Canberra distance returns the sum of a series of fractional differences between coordinates of a pair of points. Each fractional difference term has a value between 0 and 1. If one of coordinate is zero, the term is defined to be unity regardless the other value.

$$d_{ij} = \sum_{k=1}^{n} \frac{|x_{ik} - x_{jk}|}{|x_{ik}| + |x_{jk}|} \quad (3)$$

- *Cosine angle separation*: This measure represents the cosine of the angle between two vectors. It measures similarity rather than distance or dissimilarity. Thus, a high cosine angle separation value indicates that two vectors are similar.

$$s_{ij} = \frac{\sum_{k=1}^{n} x_{ik}.x_{jk}}{\sqrt{\sum_{k=1}^{n} x_{ik}^2 . \sum_{r=1}^{n} x_{jr}^2}} \quad (4)$$

- *Matching distance*: This measure is simply the raw number of location probabilities that are similar for two nodes. We consider two coordinates on a given axis to be similar if their absolute difference is less than or equal to a defined value δ.

In the routing scheme presented here, the only information that must be flooded by nodes is their mobility patterns. These mobility patterns can be spread in an epidemic fashion. Furthermore, we introduce an optimization based upon the power-law distribution of probabilities. Under this optimization, nodes transmit only the main components of their mobility patterns. The other components are presumably negligible in comparison. We examine the performance and the limits of such an optimization in Sec. 3.2.2.

3. SIMULATION RESULTS

This section presents the manner in which we evaluated the mobility pattern based routing scheme, and the results we obtained.

3.1 Methodology

We have implemented a stand alone simulator to evaluate the mobility pattern based routing scheme presented in this paper. This simulator only implements the transport and network layers and it makes simple assumptions regarding lower layers, for instance allowing for infinite bandwidth.

We studied the use of each of the four functions described in the previous section for routing in the MobySpace. We will refer to them here by the following names: *Euclidean*, using the Euclidean distance metric; *Canberra*, using the Canberra distance metric; *Angle*, using the cosine angle separation distance metric; and *Matching*, using the matching distance metric. We compared these MobySpace routing algorithms against the following:

- *Epidemic*: This is based on epidemic routing, as described by Vahdat and Becker [18]: Each time two nodes meet, they exchange their bundles. The major interest of this algorithm is that it provides the optimum path and thus the minimum bundle delay. We use it here as a lower bound. In practice, epidemic routing suffers from high buffer occupancy and high bandwidth utilization.
- *Opportunistic*: A node waits to meet the destination in order to transfer its bundle. The main advantage of this method is that it involves only one transmission per bundle.
- *Random*: When a node is at a location and the bundle's destination in not there, the node transfers the bundle to a neighbor chosen at random. We have added a rule to avoid local loops: a node can only handle a bundle one time per location visit. This scheme is used in this paper as another basis of comparison. A novel algorithm should perform better than this one in order to be valuable.

All the scenarios simulated in the rest of the paper share common parameters that can be found in Table 1. We considered a set of 25 locations. The MobySpace used for routing thus has 25 dimensions. There are 50 mobile nodes. Every node generates bundles destined toward each of the others every 30 s with the first bundle being sent at a time randomly chosen from a uniform distribution over the interval [0,30s]. Simulations last 4000 s. We generate traffic in the first 500 s of the simulations in order to give enough time for all the bundles to reach their destination. The simulator used a time step of 10 ms.

Parameter	**Value**
Number of nodes	50
Number of locations	25
Simulation duration	4000 s
Traffic generation	until 500 s
Packet interval	30 s
$t_{\min}$	5 s
$t_{\max}$	15 s
δ	210^{-8}
Time step	10 ms

Table 1: Simulation parameters

We have tested two variants of the mobility pattern based routing scheme. In the first, we assume that a node that is sending a bundle has full knowledge of the destination's mobility pattern, and that it addresses the bundle accordingly. In the second, we assume that nodes communicate only the major components of their mobility patterns. This reduces the amount of control traffic exchanged between nodes, but it also means that a node that is sending a bundle can only specify partial information regarding the destination.

3.2 Results

We evaluate routing algorithms on their transport layer performance in the simulation. We consider a good algorithm to be one that yields a low average bundle delay and a low average route length.

3.2.1 *With full knowledge*

We preface our detailed discussion of simulation results with the observation that Euclidean and Angle yielded identical results. This may be explained by the fact that when the number of dimensions of the space is high, there is a strong correlation between those metrics, as shown by Qian et al. [16]. In this section, therefore, the Euclidean and Angle results are reported together.

We performed 5 runs for each set of parameters (the number was limited by the length of time required for simulations). Figures reported in the tables here are mean results with confidence intervals at the 90% confidence level, obtained using the Student t distribution.

d	1.1	1.5	2
Epidemic	10.9 ± 7.3	13.2 ± 0.4	16.2 ± 0.5
Opportunistic	123.3 ± 7.7	287.4 ± 8.4	550.2 ± 15.2
Random	117.8 ± 8.0	160.0 ± 1.9	203.3 ± 17.3
Euclidean & Angle	103.0 ± 7.7	59.1 ± 2.7	54.6 ± 2.0
Canberra	104.8 ± 4.6	113.4 ± 10.4	245.0 ± 41.2
Matching	118.5 ± 5.7	189.5 ± 12.1	352.9 ± 56.0

Table 2: Average bundle delay

Table 2 presents the mean bundle delay obtained for each routing algorithm, and for various exponents, d, of the power law distribution, describing the preferential attachment of nodes toward each location. The notable feature of these results is that Euclidean and Angle show improved performance with an increase in d, whereas performance declines for all other routing algorithms as d increases. Opportunistic performs worst, followed closely by Random, Matching, and Canberra.

The fact that Matching and Canberra are worse than Random is interesting. One hypothesis could be that they are actively making poor choices. However, we have reason to believe that Random has a delay advantage that Matching and Canberra do not share. In Random, bundles will jump to other nodes without any preference ordering. This makes for highly mobile bundles, as is borne out by their extraordinarily high average route lengths, shown in Table 3. One might not necessarily want to pay the price of such processing overhead in order to obtain modest gains in delay. A better standard for comparison might be a random algorithm that shows preferences, as do Matching and Canberra, but preferences that are purely random in nature. Our judgment concerning Matching and Canberra is thus still in reserve pending further study.

Projecting from the results in this section, we might ask what would happen for ever higher values of d. Recall that, the higher the value of d, the higher the probabilities are that nodes will find themselves at a few select locations. For a high enough value, there would be little node movement, and little diversity in their movements. We would expect this to have a negative effect on all routing schemes, including Euclidean and Angle. We have not yet had the opportunity to conduct studies to see if this phenomenon emerges as expected.

Table 4 and Table 5 are representative plots from one of the five experiments conducted. Note that bins of size 1 s have been used to aggregate data and that 38378 bundles have been generated in this run. Table 4 presents the frequency of the delay difference in seconds compared to Epidemic. For each bundle, we have compared the delay we

d	1.1	1.5	2
Epidemic	3,7 ± 0.0	3.7 ± 0.0	3.8 ± 0.1
Opportunistic	1 ± 0.0	1 ± 0.0	1 ± 0.0
Random	44.5 ± 0.7	55.9 ± 1.0	69.8 ± 2.2
Euclidean & Angle	3.3 ± 0.0	3.2 ± 0.0	3.2 ± 0.0
Canberra	3.3 ± 0.0	3.2 ± 0.0	3.2 ± 0.0
Matching	2.5 ± 0.0	2.5 ± 0.0	2.4 ± 0.0

Table 3: Average route lengths

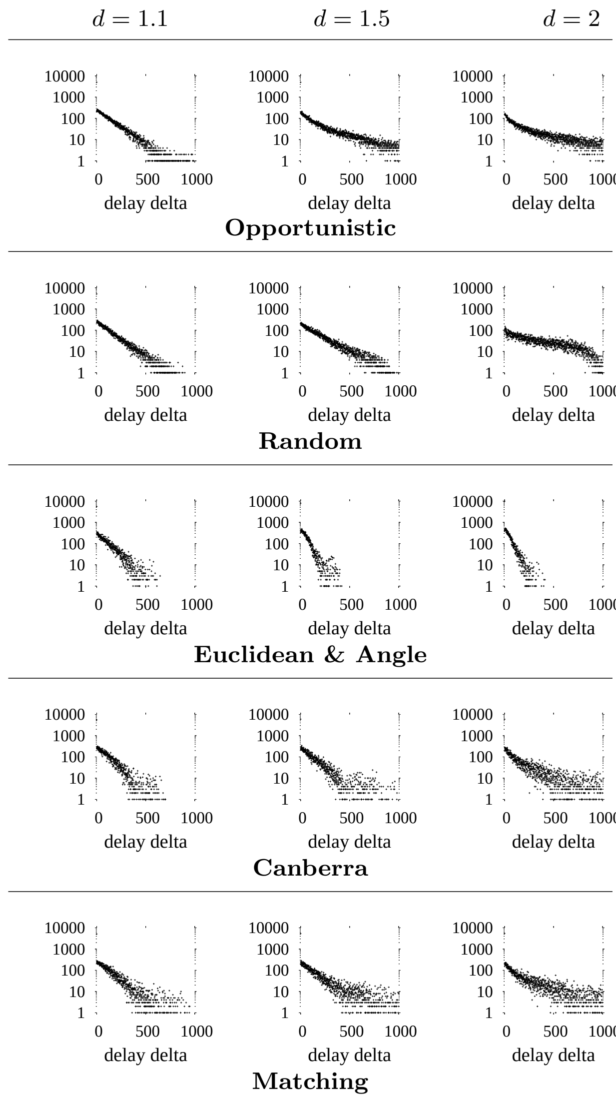

Table 4: Delay frequency compared to Epidemic

obtained with Epidemic to the delay we obtained with the other algorithms. This was feasible since the traffic generation process can be identically reproduced in our simulator with different routing protocols. Since it is not possible in our case to find an algorithm having the same performance as Epidemic, we wish to for an algorithm that approximates its performance to the extent possible. We can observe that the higher d is, the heavier is the tail of the distribution,

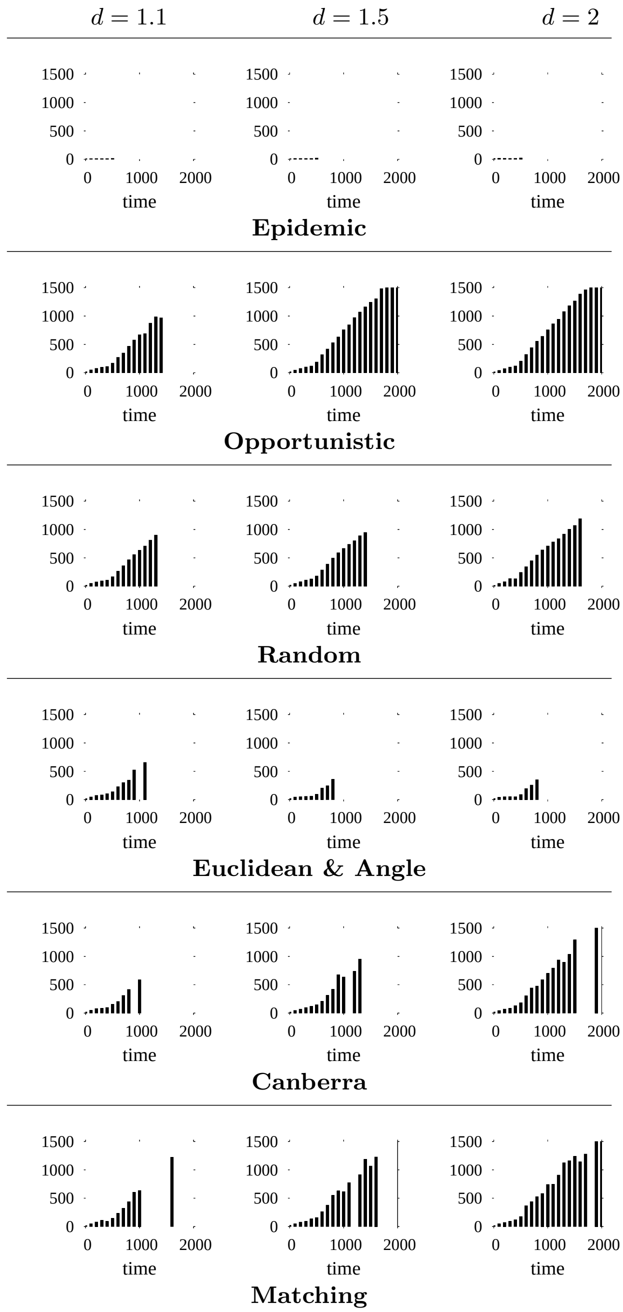

Table 5: Average delay evolution by steps of 100 s

except for Euclidean and Angle, which show much better behavior.

Table 5 shows the evolution of the average delay in seconds over time, by steps of 100 s, for the different algorithms. Epidemic performs the best in terms of delay and route length. It leads to very low delays that remain constant over time. No bundles are delivered after 500 s because we stop generating traffic at that time. For the other algorithms, bundle delay increases slightly before 500 s and then linearly with time after that point because remaining bundles are waiting in buffers to be delivered. When some bundles are difficult to deliver, it results in delivery in batches, such as for Matching at time 1600 s for the case $d = 1.1$. Euclidean and Angle seem to be the best solutions, resulting in low delays and fast deliveries when the mobility patterns have power-law distributions with d greater than 1.5.

3.2.2 With partial knowledge

Here, we analyze the performance and the limitations of a scheme by which communication overhead is reduced by having nodes only diffuse the main components of their mobility patterns.

We ran simulations for the different similarity metrics with values of d going from 1.1 to 2, and by taking into account only the principal 1^{st}, 2^{nd}, 3^{rd}, or 4^{th} components of a node's mobility pattern.

metric		$d = 1.1$	$d = 1.5$	$d = 2$
Euclidean	$l = 25$	103.0 ±7.7	59.1 ±2.7	54.6 ±2.0
	$l = 4$	106.2 ±4.3	60.0 ±2.4	54.5 ±2.3
	$l = 3$	107.2 ±7.0	60.0 ±2.4	54.9 ±1.8
	$l = 2$	107.2 ±7.0	62.2 ±2.3	57.2 ±1.9
	$l = 1$	110.7 ±4.6	69.2 ±3.0	75.1 ±9.4
Angle	$l = 25$	103.0 ±7.7	59.1 ±2.7	54.6 ±2.0
	$l = 4$	106.2 ±4.3	60.0 ±2.4	54.5 ±2.3
	$l = 3$	107.4 ±7.2	60.0 ±2.4	54.5 ±1.8
	$l = 2$	107.4 ±7.2	62.3 ±2.2	57.2 ±1.9
	$l = 1$	110.5 ±4.8	69.0 ±3.0	75.1 ±9.0
Canberra	$l = 25$	104.8 ±4.6	113.4 ±10.4	245.0 ±41.2
	$l = 4$	106.4 ±5.2	68.6 ±4.1	98.5 ±8.5
	$l = 3$	106.4 ±7.5	68.6 ±4.1	80.2 ±3.3
	$l = 2$	106.4 ±7.5	67.5 ±1.6	66.2 ±3.2
	$l = 1$	109.9 ±4.5	69.9 ±2.8	75.2 ±9.4
Matching	$l = 25$	118.5 ±5.7	189.5 ±12.1	352.9 ±56.0
	$l = 4$	116.2 ±6.4	109.0 ±4.5	225.1 ±16.0
	$l = 3$	113.3 ±4.8	109.0 ±4.5	168.8 ±8.8
	$l = 2$	113.3 ±4.8	103.5 ±3.3	164.7 ±8.2
	$l = 1$	112.7 ±4.1	136.8 ±6.3	265.6 ±10.7

Table 6: Average delay, with l the number of components taken into account

Table 6 shows the average bundle delays and Table 7 the average route lengths. The case with $l = 25$ in the tables comes from the results obtained with full knowledge. We provide separate results for Angle and Euclidean because they are not exactly the same, even if they remain similar. We can see two kind of behaviors for the metrics. On one hand, for Angle and Euclidean, the less information is available, the higher is the delay. Route lengths remain the same, except for $l = 1$ where it decreases a bit. On the other hand, for Canberra and Matching the less information we have, the lower are the delays (except for $l = 1$ where it increases a bit) and the lower are the route lengths. These results remain higher in term of delay than when nodes have full knowledge of destination's mobility patterns but show that we can obtain a diminution of route lengths on average.

This is especially the case for Matching where route lengths are lower (up to 1 hop less) without leading to dramatic delays for small values of d.

The fact that Canberra and Matching performs better in most cases with low values of l than for $l = 25$ tends to confirm that these metrics were certainly not used at their best in this study. Further analysis of these metrics is required.

metric		$d = 1.1$	$d = 1.5$	$d = 2$
Euclidean	$l = 25$	3.3 ±0.0	3.2 ±0.0	3.2 ±0.0
	$l = 4$	3.3 ±0.0	3.2 ±0.0	3.2 ±0.0
	$l = 3$	3.3 ±0.0	3.2 ±0.0	3.2 ±0.0
	$l = 2$	3.3 ±0.0	3.2 ±0.0	3.2 ±0.1
	$l = 1$	3.1 ±0.0	3.1 ±0.0	3.0 ±0.0
Angle	$l = 25$	3.3 ±0.0	3.2 ±0.0	3.2 ±0.0
	$l = 4$	3.3 ±0.0	3.2 ±0.0	3.2 ±0.0
	$l = 3$	3.3 ±0.0	3.2 ±0.0	3.2 ±0.0
	$l = 2$	3.3 ±0.0	3.2 ±0.0	3.2 ±0.1
	$l = 1$	3.1 ±0.0	3.1 ±0.0	3.0 ±0.0
Canberra	$l = 25$	3.3 ±0.0	3.2 ±0.0	3.2 ±0.0
	$l = 4$	3.3 ±0.0	3.2 ±0.0	3.2 ±0.0
	$l = 3$	3.3 ±0.0	3.2 ±0.0	3.2 ±0.0
	$l = 2$	3.3 ±0.0	3.2 ±0.0	3.2 ±0.0
	$l = 1$	2.9 ±0.0	2.9 ±0.0	3.0 ±0.0
Matching	$l = 25$	2.5 ±0.0	2.5 ±0.0	2.4 ±0.0
	$l = 4$	1.9 ±0.0	1.8 ±0.0	1.9 ±0.0
	$l = 3$	1.7 ±0.0	1.8 ±0.0	1.8 ±0.0
	$l = 2$	1.7 ±0.0	1.7 ±0.0	1.7 ±0.0
	$l = 1$	1.5 ±0.0	1.5 ±0.0	1.5 ±0.0

Table 7: Average route lengths, with l the number of components taken into account

4. CONCLUSION AND FUTURE WORK

The main contribution of this paper has been the definition of the concept of MobySpace, a generic routing scheme using the formalism of a high-dimensional Euclidean space constructed upon mobility patterns. We have shown, in a scenario inspired by real-world observations, that the scheme can applied to DTNs and that it can bring benefits, through reduced bundle delay and through lower communication costs. We believe that the scheme is sufficiently general that it opens broad perspectives for DTN routing. Much work remains to be done.

One possible line of future work concerns the impact of the structure of the MobySpace, where by structure we mean both the number and type of dimensions that define the space, and the similarity function that is used to calculate distances within the space. As described in Sec. 2.1, an alternative set of dimensions might be based on the frequency of contacts directly between pairs of nodes, rather than the frequency of node visits to locations.

Another line of future work concerns the impact on performance of different types of mobility patterns. Dynamic patterns are of particular concern: what happens if mobility patterns are constantly evolving and shifting? How can signaling keep pace with these changes, so that the mobility patterns that are used for routing have sufficient predictive power?

In this paper, we have assumed that nodes have full knowledge of their mobility patterns. But typically we would expect that nodes will have to learn their mobility patterns over time. The impact of imperfect estimation will need to be taken into account.

The Euclidean spaces that we have studied here have a known finite number of dimensions. In reality, the number of dimensions might not be known in advance, and each node can have its own separate, and ever-growing, list of dimensions. What semantics will allow nodes to exchange mutually-intelligible information about mobility patterns under such circumstances?

In this work, we have considered that the essential characteristics of a node's mobility or contact patterns are fully captured by the frequency with which nodes find themselves in certain locations or the probability that they will be in proximity to certain nodes. However, prior work [11], has demonstrated the interest of capturing temporal information as well. It is well known that network usage patterns follow diurnal and weekly cycles. We could easily imagine two nodes that visit the same locations with the same frequencies, but on different days of the week. Though it still might make sense to route to one node in order to reach the other, especially if there is a relay node at the commonly visited location, the Euclidean space description that we have provided does not capture this essential dissimilarity between the nodes.

We can, however, imagine ways in which the dimensional representation could capture temporal information as well. For instance, visit patterns could be translated into the frequential domain. Here, we mean frequency in cyclical terms, not in terms of probability, as we have used the word elsewhere in this paper. A node's visits to a location could be represented by a point on a frequency axis, capturing the dominant frequency of those visits, and a point on a phase axis. These axes could be added to the axis already described, that represents the overall probability of visiting the location.

5. REFERENCES

[1] Delay Tolerant Network Research Group (DTNRG). Website: www.dtnrg.org.

[2] I. F. Akyildiz, O. Akan, C. Chen, J. Fang, and W. Su. Interplanetary internet: state-of-the-art and research challenges. *Computer Networks*, 43(2):75–112, 2003.

[3] M. Balazinska and P. Castro. Characterizing Mobility and Network Usage in a Corporate Wireless Local-Area Network. In *Proc. MobiSys*, 2003.

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

[5] V. Cerf, S. Burleigh, A. Hooke, L. Torgerson, R. Durst, K. Scott, K. Fall, and H. Weiss. Delay tolerant network architecture, draft-irtf-dtnrg-arch-02.txt, July 2004.

[6] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott. Pocket switched networks: Real-world mobility and its consequences for opportunistic forwarding. Technical Report UCAM-CL-TR-617, Intel Research Cambridge, February 2005.

[7] S. Corson. Mobile ad hoc networking (MANET): Routing protocol performance issues and evaluation considerations. RFC 2501, January 1999.

[8] J. A. Davis, A. H. Fagg, and B. N. Levine. Wearable computers as packet transport mechanisms in highly-partitioned ad-hoc networks. In *Proc. ISWC*, 2001.

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

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

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

[12] A. Lindgren, A. Doria, and O. Schelen. Probabilistic routing in intermittently connected networks. *SIGMOBILE Mob. Comput. Communications Rev.*, 7(3):19–20, 2003.

[13] M. McNett and G. M. Voelker. Access and mobility of wireless PDA users. Technical report, UC San Diego, 2004.

[14] M. Musolesi, S. Hailes, and C. Mascolo. Adaptive routing for intermittently connected mobile ad hoc networks. In *Proc. WOWMOM*, 2005.

[15] V. N. Padmanabhan and L. Subramanian. An investigation of geographic mapping techniques for internet hosts. In *Proc. SIGCOMM*, 2001.

[16] G. Qian, S. Sural, Y. Gu, and S. Pramanik. Similarity between euclidean and cosine angle distance for nearest neighbor queries. In *Proc. ACM Symposium on Applied Computing*, 2004.

[17] S. Ratnasamy, P. Francis, M. Handley, R. Karp, and S. Shenker. A scalable content addressable network. In *Proc. SIGCOMM*, 2001.

[18] A. Vahdat and D. Becker. Epidemic routing for partially connected ad hoc networks. Technical Report CS-200006, Duke University, April 2000.