Exploiting Space Syntax for Deployable Mobile Opportunistic Networking

Abderrahmen Mtibaa and Khaled A. Harras
School of Computer Science
Carnegie Mellon University
Email: {amtibaa, kharras}@cmu.edu

***Abstract*—Despite the plethora of opportunistic forwarding solutions offered by the research community, we revisit this domain from a new perspective by exploiting the concept of space syntax to enable deployable solutions in large scale urban environments. We present a set of algorithms that build upon space syntax, which predicts natural movement patterns by interacting with pre-built static environments. We design these algorithms for three assumption categories that represent the spectrum of assumptions regarding the underlying environment and node capabilities. We adopt a data-driven approach to evaluate the performance of our algorithms when compared to other state-of-the-art solutions within each representative category that make similar assumptions. Overall, our results show the great promise space syntax based algorithms have for efficiently guiding messages towards the destination. We show 5% to 20% success rate improvement compared to selected well known state-of-the-art forwarding algorithms within each assumption category while reducing the cost in terms of message replicas by up to 10%.**

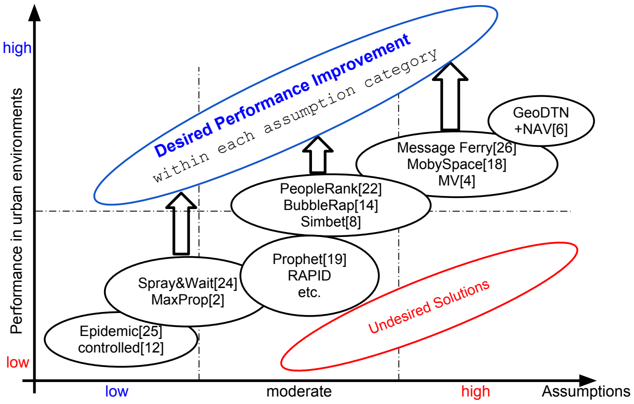


Fig. 1. Desired performance improvement in large scale urban environments under each assumption category

I. INTRODUCTION

There are many cities in countries like India, Pakistan, and Egypt, where urbanization occurs at a faster rate than that of communication infrastructure deployment. Mobile users with sophisticated devices are often dissatisfied with this lag in infrastructure deployment; their connection to the Internet is usually via opportunistic open access points for short durations, or via weak, unreliable, and costly 3G connections. In such situations, as well as in developed cities with increased demands on network infrastructure, we believe that opportunistic networking, where user mobility is exploited to increase capacity and augment Internet reachability [10], can play an active role as a complimentary technology to improve user experience, particularly with delay insensitive data.

In our work, we consider opportunistic mobile to infrastructure transfers in urban city environments. In such environments, delivering data through the infrastructure may create path shortcuts towards final destinations. We therefore consider mobile-to-infrastructure transfers where destinations are public access points acting as Internet hotspots.

Over the past decade, researchers in opportunistic networks have typically focused on efficient forwarding in small scale environments such as conferences, malls, and university campuses. Forwarding solutions were mainly designed using a set of assumptions that have grown in complexity, rendering solutions built on them unusable outside their intended environment categories. Fig. 1 categorizes a set of state-of-the-art opportunistic forwarding solutions based on their required assumptions which include the knowledge of historical contacts with other devices [2], [19], information about device mobility patterns [26], the total number of nodes in the network [24], or static social information [8], [22]. Most of these solutions, however are not designed for large scale urban environments.

In this paper, we exploit the space syntax paradigm [16] to better guide forwarding decisions in large scale urban environments. Space syntax, initially proposed in the field of architecture to model natural mobility patterns by analyzing spacial configurations, offers a set of measurable metrics that quantify the effect of road maps and architectural configurations on natural movement. By interacting with the pre-built static environment, space syntax predicts natural movement patterns in a given area [16]. *Our goal is to leverage space syntax concepts to create efficient opportunistic forwarding distributed solutions for large scale urban environments*. To the best of our knowledge this work represents the first serious attempt to exploit space syntax concepts to improve opportunistic forwarding decisions in large scale urban cities.

We propose a set of space syntax based algorithms that adapt to a spectrum of simplistic assumptions in urban environments. To obtain the best possible performance, we first propose an algorithm at the high end of the assumption spectrum where nodes are aware of the city map, the location of destination Internet hotspots, and node mobility profiles. We then gradually remove assumptions and propose algorithms within the moderate and low categories where awareness of the city map layout is the only remaining assumption. As depicted in Fig 1, our goal in this work is to gain performance improvement across the spectrum, within each assumption category, when compared to other state-of-the-art solutions.

We adopt a data driven approach to evaluate the space syntax based forwarding algorithms we propose, within each of three assumption categories, based on large scale mobility traces capturing vehicle mobility [23] [3]. Overall, our results show that our space syntax based algorithms perform more efficiently within each assumption category. In the low assumption side of the spectrum, our algorithm outperforms other opportunistic forwarding algorithms, such as Prophet [20] and Spray&Wait [24], by 5% to 20% while reducing the cost by roughly 15%. In the medium and high assumption ranges, we outperform similar approaches such as MobySpace [18] and MV [4] by up to 20% while reducing the forwarding cost by more than 10%. Improvements in performance are observed even when comparing our low assumption algorithm to state-of-the-art algorithms that operate under higher assumption category. The significance of these results lies, not only in the percentage improvement in performance, but in the simplified assumptions adopted in our solutions compared to other relevant research.

The remainder of this paper is organized as follows. Section §II describes related work in the fields of mobile opportunistic networks and space syntax. Section §III provides a brief overview of the space syntax concept and introduces the terminology used throughout the paper. Section §IV discusses the different space syntax based forwarding algorithms we propose based on the range of our assumption spectrum. Our data-driven evaluation against state-of-the-art solutions is then detailed in Section §V. Finally, we conclude in Section §VI.

II. RELATED WORK

Researchers in opportunistic networks have designed various solutions correlated with assumptions regarding specific applications and environments. With minimal assumptions, naive forwarding protocols based on flooding [25] are extremely inefficient, even with the various controlled flooding techniques developed with simplistic assumptions [12]. More efficient solutions make various assumptions such as the knowledge of ferry routes and contact predictability as in Message Ferrying [26], details about future device mobility [18], [6], the ability to control the mobility of some nodes [4], [26], historic contact information [2], [19], or various social-based information such as rank or centrality [8], [22], [14]. Generally speaking, these solutions perform poorly outside their intended environments.

Position-based opportunistic routing techniques are closest to the environments considered in this paper [4], [18], [6]. *GeoDTN+Nav* exploits on-board vehicular navigation systems to guide delivering packets [6]. Vehicles exchange routes to predict future contact opportunities, assuming every vehicle enters its full route in its navigation system or has fixed routes (*e.g.,* buses). This assumption, however, is unrealistic in the situations and scale we consider since most vehicles do not utilize navigation systems, and if they did, city residents do not typically use them since they know the roads well enough. *Mobyspace* uses Euclidean space as a tool to make better routing decisions [18] where messages will be routed to nodes having similar mobility patterns as the destination. In a similar manner, *MV* learns the movement patterns of the nodes in the network in order to guide message passing [4]. Such techniques require the exchange of a considerable amount of data in order to update a highly dynamic graph of node mobility patterns. We compare our work to Mobyspace and MV and show how space syntax based solutions are more efficient in the environments we consider.

We summarize the correlation between performance and assumptions with respect to some prominent examples in the research body in Fig. 1. The assumptions made behind these solutions do not allow opportunistic networking to operate at a large scale urban environments as a complementary technology to existing networking infrastructures. In this paper, we consider a novel approach, by leveraging space syntax [16] concepts in order to create opportunistic forwarding solutions that would fall in the desired circle of Fig. 1. As opposed to position-aware algorithms such as MobiSpace and MV, space syntax provides techniques to compute the importance (*i.e.,* popularity) of a place based on static information related to map layout, and not the nodes themselves.

Our earlier work represents the first attempt on exploiting space syntax techniques to improve forwarding decisions [1]. This work, however, does not provide scalable algorithms, makes hard assumptions about identifying and calculating popularity indices, and does not thoroughly evaluate the concept. In this paper, we consider a clear spectrum of assumption categories, propose efficient scalable distributed algorithms for each category, and investigate the trade-offs between them. Finally, we evaluate these algorithms using real mobility traces and compare to various state-of-the-art solutions.

III. SPACE SYNTAX FOR MOBILE OPPORTUNISTIC NETWORKS

Space syntax was introduced in 1984 by Hillier and Hanson mainly to build and design the architectural spaces in which people move [13]. The main concept of space syntax relies on providing simple mechanisms to better model and understand how cities operate. Some researchers used it to model user behavior and their interactions with an urban space [15]. More recently, it has been utilized to predict pedestrian and vehicular natural movement within a city [16].

Technically, space syntax provides a unique transformation of a city map, which can be represented as an axial map, into a graph $G(V, E)$ [16]. Each street, boulevard, and highway in the city map is represented as a vertex $v \in V$. Edges in E represent intersections between the corresponding vertices. An example of this representation is provided in Fig. 2. This graph representation is different from how computer scientists typically represent maps in graph theory. Representing streets using vertices, instead of edges, provides an easy representation that allows us to assign concrete parameters

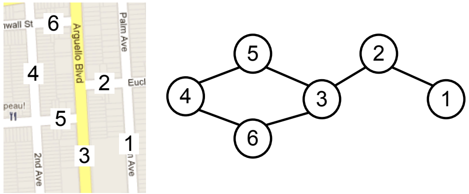


Fig. 2. Space syntax graph representation

and differentiate between street types such as boulevards and highways.

We believe that space syntax introduces a novel way by which opportunistic forwarding decisions can be taken. Space syntax mainly relies on static information such as road maps which are highly unlikely to change over short periods of time. It provides simple tools to compute the importance or popularity of a location in the map based on information related to the map layout, rather than the nodes themselves. Locations can then be ranked according to how attractive a location is for people and vehicles in the city. A normalized ranking of 0.25 would mean that a given location is more popular than 25 percent of all the other locations in the map. This knowledge can be exploited in the opportunistic networking world to provide guidance about when and how to spread messages in popular locations in order to avoid unnecessary overhead. In this paper, *we consider information related to the map layout to guide opportunistic forwarding based on the likelihood of a location to attract mobile nodes*.

In opportunistic networks, whenever two mobile nodes come within communication range, we call such event a *contact*, during which nodes can exchange data messages. Because such networks exhibit intermittent connectivity, data is typically stored in intermediate nodes awaiting appropriate contact formation. We are interested in delivering the maximum number of messages in mobile networks with the least delay and the least cost in terms of number of message replicas. We tackle the typical cases in such networks where: (i) messages may not be fragmented, and (ii) contact duration is long enough to exchange messages between nodes.

We believe that opportunistic communication can play a significant role as a supplementary network that co-exists with the deployed infrastructure, especially for supporting delay insensitive data (*e.g.,* email, multimedia upload or download). Opportunistic communication can therefore support this infrastructure when it is sparsely deployed, or is too costly. In fact, delivering a message through the infrastructure instantaneously creates path shortcuts towards the final destination (*i.e.,* dissemination through the infrastructure is faster). Hence, a category of our solutions in the following sections focuses on mobile-to-infrastructure transfers where destinations are public access points or gateways with reliable connections to the Internet. The main goal is therefore to *ensure successful delivery through the infrastructure*. Routing from such gateways to either a directly connected or opportunistically connected destination is, however, another interesting and challenging research problem [11], [17], [21].

IV. FORWARDING UNDER DIFFERENT ASSUMPTION CATEGORIES

In this section, we first present the spectrum of assumptions that we consider for deployable large scale urban environments. This spectrum traverses three different categories with different assumption complexity all based on core space syntax concepts. Afterwards, we propose various forwarding algorithms within each of these assumption categories. Each of theses algorithms exploit space syntax techniques in a different way in order to reduce the amount of information needed to achieve better forwarding performance.

A. The Assumption Categories

In order for forwarding solutions to be truly deployable, we only consider forwarding decisions autonomously made using static information (e.g. city map and attraction points) or local information regarding the node itself. We focus on exploiting space syntax concepts coupled with zero to many a priori local information to improve forwarding decisions in mobile opportunistic networks. We consider assumptions restrictively known by a mobile node with regards to its social/mobility behavior, or attractions in its vicinity. We avoid any global and unrealistic assumptions in mobile disconnected and distributed networks, such as the awareness of the total number of nodes, the network topology, or node trajectory.

Fig. 3 shows our proposed space syntax based forwarding algorithms within an assumption spectrum. We classify this spectrum into three categories, each using a different set of assumptions. We propose, in this section, one or many forwarding *algorithms* within each category. The high end category assumes that nodes are aware of a map layout, their positions and the position of destination Internet access point locations (*i.e.,* hotspots in popular city locations), and their own mobility patterns. For this category, we propose ASOF: the Assisted Space-syntax Opportunistic Forwarding algorithms. We remove the mobility pattern awareness in the middle category and offer LASOF: the Location-Aware Space-syntax Opportunistic Forwarding algorithms. Finally, in the low end of the spectrum, we propose SOF: the Space-syntax Opportunistic Forwarding algorithms, where nodes are only aware of the city map.

B. Assisted Space-syntax Opportunistic Forwarding (ASOF)

In this subsection, we assume that a node has a priori knowledge about environmental characteristics of the area

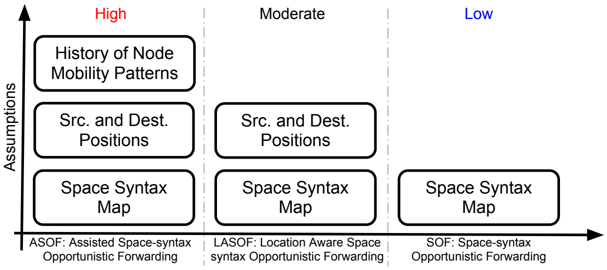


Fig. 3. Space syntax based solution under different assumption categories

which includes the map of the city, the position of public access point within the city (*i.e.,* destinations), and node's own historical mobility patterns. We believe normal high-end phones nowadays can easily keep track of and store such information which can provide targeted recommendations such as attractions and public WiFi access points in the vicinity. This represents the most demanding assumption set within the assumption spectrum as shown in Fig. 3.

Prior to introducing our algorithms, we first present two forwarding utility metrics based on (*i*) the probability of a node to reach the destination (Maximum Probability) and (*ii*) the average delay needed to reach a destination (Earliest Delay).

1) Maximum Probability: The Space Syntax Maximum Probability utility function (SS-MP) ranks all nodes in the network based on their probability to reach a destination. Every node i uses its mobility pattern history in order to update its space syntax graph $G_i(V, E_i)$ as follow: $E_i = \{e_{u,v} = P_i[v|u], u \neq v \in V\}$, where $P_i[v|u]$ is the probability with which node i moves from a street u to another street v within a city represented by the $G_i(V, E_i)$ graph. Therefore, every node computes its forwarding *utility* as the maximum probability to reach the destination $D = \{d_j, j = 1..k\}$, where k is the total number of public access points in the city (*i.e.,* number of destinations).

$$U_i^1(u) = \max_{j=1..k} \left\{ P_i[d_j|u] = \prod_{\{u,..,d_j\}} P_i[u_n|u_{n-1}] \right\}$$

$P_i[d_j|u]$ is the probability computed by node i to reach a destination d_j from a location u, and $\{u, .., d_j\}$ is any path connecting the street u and the street d_j. u_{n-1} and u_n represent two consecutive locations in the path. In Fig 2, we assume that a mobile node (*i.e.,* a driver) is used to take street 5 more than street 6 to drive from street 4 to street 3. We then assign to this node's graph a higher edge weight between $e_{4,5}$ compared to the one for $e_{4,6}$. We essentially assign the probability of node movement in a particular city.

2) Earliest Delay: The Space Syntax Earliest Delay utility function (SS-ED) ranks all nodes in the network based on their minimum remaining delay to reach a destination. We update the edge weights $e_{u,v}$ of the space syntax graph using the mobility pattern history to determine the average delay needed for a particular node to move from a street u to another street v. Utility function is computed as the minimum delay to reach the closest destination from a particular street u:

$$U_i^2(u) = \min_{j=1..k} \left\{ D(d_j|u) = \sum_{\{u,..,d_j\}} D(u_n|u_{n-1}) \right\}$$

Where $D(d_j|u)$ is the average delay needed to reach one destination $d_j \in D$, and $\{u, .., d_j\}$ is any path connecting the street u with street d_j. For example, in Fig 2, every node will update its average amount of time spent to move from one street to another. This will help estimate the distance traveled by a node from one street to another in a given map. The node then computes an average estimated time to reach an Internet relay from any particular street in the city.

3) The ASOF Algorithm: We have described the first step in employing a space syntax forwarding approach which is determining the utility function per node over time. The second step is to compare these utility metrics U_i and U_j whenever two nodes i and j come within wireless range. Node i will then decide to forward the message at a particular street u to node j if and only if $U_i(u) \leq U_j(u)$ (using U^1 or U^2 as described in Alg. 1). Having a directed and weighted graph G_i, a node is responsible for updating its shortest path from its current location (*i.e.,* street) to a set of destinations $D = destination(m) = d_j, j = 1..k$ of a given message m. While this technique is fast and simple, it falls under the highest considered assumption category that may or may not be realistic.

Algorithm 1 ASOF: SS-ED(node i)

```
Require: $G_i(V, E_i)$
Ensure:
$U_i \leftarrow 0$
while 1 do
    $\forall u \in V, update(G_i(V, E_i))$
    while $i$ is in contact with $j$ do
        while $\exists\ m \in buffer(i)$ do
            $U_i \leftarrow U_i^2(current\_street)$
            $exchange(U_i^2, U_j^2)$
            if $U_j^2 \leq U_i^2$ OR $j \in destination(m)$ then
                $Forward(m, j)$
            end if
        end while
    end while
end while
```

C. Location Aware Space-syntax Opportunistic Forwarding (LASOF)

In Location Aware Space-syntax Opportunistic Forwarding category (LASOF), we assume no awareness about the history of node mobility patterns. We note that knowing the position of public access points in a city can be considered as a realistic assumption since it has been used in most of the vehicular networks papers. We now describe two heuristic based approaches to measure the probability of vehicles moving from one street to another. Afterwards, we discuss the LASOF algorithm.

1) SS-IA: The *Implicit Approach* does not rely on any additional information to estimate these probabilities. Street popularity can be determined using Google traffic, history of the place, or additional information such as WikiPedia or Facebook. We choose street widths to reflect the popularity or expected demand and traffic on that street. We consider the following street types: St: street, Av: avenue, $Blvd$: boulevard, and Fwy: freeway. Our idea is motivated by the following intuition: streets are designed to accommodate the traffic of a city. The width of a street is then proportional to the relative traffic load in general. Therefore, we assume that $width(St) \leq width(Av) \leq width(Blvd) \leq width(Fwy)$ where $width()$ is a function to measure the average width of a street. Let $P[St|St] = P$, $P[Av|St] = 2 \times P$, $P[Blvd|St] = 3 \times P$, and $P[Fwy|St] = 4 \times P$. Having a directed and weighted graph, we know that: $\sum_{e_{u,v} \in E} P[v|u] = 1$. We then compute P, and the weights of the graph edges using this implicit approach (*SS-IA*).

2) SS-EA: The *Explicit Approach* consists of a statistical method relying on online traffic volume data reports. Such information is available either online or could be extracted from historical data reports or other sources [7]. We aggregate all available reports about a specific region during long periods of time and compute the transition probabilities from one street to another. Using this space syntax explicit approach (*SS-EA*), we generate an undirected and weighted graph $G_i(V, E_i)$ for all vertices in the network.

3) The LASOF Algorithm: Having a space syntax graph defined using either the implicit or explicit approach, each node will then compute its average distance to all the destination nodes in the networks. Particularly, a node u measures $cc(u)$ its closeness centrality to all $d_i \in D$. Closeness centrality is considered a measure of how long it will take to spread information from a particular node to all other nodes in the graph. The more central an node is, the lower its total distance is to all other nodes. $cc(u) = \frac{\sum_{d_i \in D} \pi(u,d_i)}{|D|}$ where $\pi(u, d)$ denotes the shortest path from the current street u to a particular destination d_i. We then propose a forwarding scheme that relies on a non-decreasing utility function $cc()$; whenever a node u carrying a message m meets another node v, u will decide to replicate the message and forward it to v if and only if $cc(v) \leq cc(u)$. In this case, node u will keep a copy of m and delegate another copy to v. The algorithm is similar to the one described by Alg. 1 with the only difference of updating the G_i graph with one of the two approaches Explicit (SS-EA) or Implicit (SS-IA).

D. Space-syntax Opportunistic Forwarding (SOF)

We now assume no awareness about the source and destination positions or historical information. The idea is then to spread a minimum number of message replicas across *popular regions* and maximize the chance that one of them reaches the destination with minimal delay. In this assumption category, we highlight the gain given by space syntax to guide data transfers in large scale urban environments.

Having a space syntax graph, we compute $C(v)$, the betweenness centrality of a node v in the graph as follows:

$$C(v) = \frac{|\{\ s, d \in V \setminus \{v\}, s \neq d\ \ |\ v \in \pi(s,d)\ \}|}{(|V|-1)(|V|-2)}$$

where $\pi(s, d)$ denotes the shortest path from s to d in G.

We then consider a non-decreasing betweenness centrality rule to forward messages from one node to another. In other words, whenever two nodes meet, they exchange their corresponding C values. Messages will then be forwarded only if the receiving node has a higher betweenness centrality. It is a heuristic method to select amongst all the opportunistic contacts the ones that are crucial in order to connect a source and destination quickly over time. Our objective is to select the most popular places within which a node should forward a message in order to reduce the contacts used, while allowing quasi-optimal delay. We call this, the space syntax based forwarding algorithm SOF.

V. EVALUATION

Realistically evaluating the performance of opportunistic network forwarding protocols in large scale urban environments is a challenge. We overcome this challenge by evaluating our algorithms using real data sets reflecting the mobility of vehicles in urban environments. Also, in each of the assumption categories, we select and implement representative state-of-the-art algorithms. These data sets, chosen state-of-the-art algorithms, metrics used, along with the compilation of a space syntax axial city map over which we conduct our evaluation, are all discussed in the evaluation methodology and setup section. Afterwards, we present our results within each of the assumption categories, followed by a brief discussion section.

A. Evaluation Setup and Methodology

1) Data Sets: We have chosen to evaluate our algorithms using analysis on real mobility traces. Specifically, we use the SanFrancisco [23] and the DieselNet [3] data sets.

SanFrancisco Data Set: Our analysis is based on the SanFrancisco data set which contains mobility traces of taxi cabs in SanFrancisco [23]. It contains GPS coordinates of approximately 500 taxis collected over 30 days. We select only 486 taxis that present consistent GPS records over 10 days. Each trace contains the reported time and location for each Cab in latitude and longitude. We take these traces for the duration of 10 days, interpolate the movement of the cabs, and then generate the contacts between the cabs. We consider the communication range of each cab to be 100 meters; *i.e.,* a contact has occurred when a cab comes in proximity of 100 meters of another cab.

DieselNet Data Set: The DieselNet data set consists of 30 to 40 buses operating from the UMass Amherst campus and its surroundings covering an area of approximately 150 square miles. Each bus takes a different path every day and can leave or change its path any time [3]. We select only 22 buses that provide consistent GPS data because other node logs contain gaps where GPS data was lost.

2) Selected State-of-the-Art Algorithms: We now describe the selected state-of-the-art forwarding algorithms we consider in our evaluation, along with the changes we made in each algorithm to effectively operate in an urban city environment where destinations are Internet hotspots. We compare our space syntax based algorithms with these opportunistic forwarding algorithms representative of different parts of the assumption spectrum. These algorithms include MobySpace [18], MV [4], Prophet [20], and Spray&Wait [24].

MobySpace is an opportunistic position-aware forwarding algorithm that uses Euclidean space as a tool to make better routing decisions. It provides nodes with probabilities of being found at different coordinates that correspond to their probability of being found in each possible location (assuming that we have a finite and predefined set of positions). We have

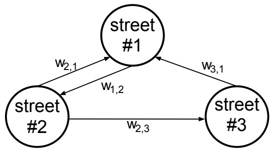


Fig. 4. A space syntax graph consisting of 3 streets

modified the implementation of the MobySpace algorithm to accept more than a single destination node for each message. Messages will then be forwarded to nodes having higher probabilities to meet at least one hotspot (*i.e.,* destination).

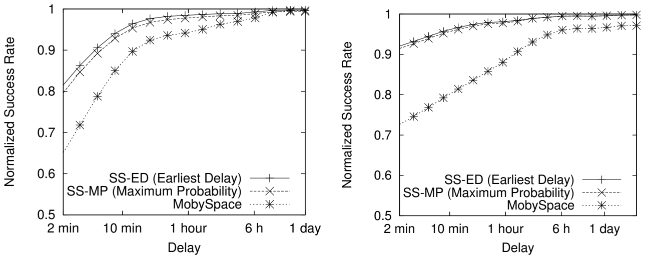
The *MV* forwarding algorithm guides message passing by estimating the likelihood of delivery per node. It computes a probability, $P_n^s(d)$ for a node s to successfully deliver a message to a destination d within n hops. It also uses additional nodes called autonomous agents that adapt their movements in response to variations in network capacity and demand. In this paper, we do not assume controlled mobility, and therefore, we implement a version that does not use autonomous agents.

Finally, we also modified Prophet [20], and Spray&Wait [24] by accommodating multiple destinations in order to fairly compare them with SOF. We implement the Binary Spray&Wait version which consists of setting a strict upper bound on the number of copies per message allowed in the network.

3) Space Syntax Axial Map: Since data sets provide GPS coordinates of mobile vehicles, we use Google's *reverse geocoding* API [9] to determine the street names and addresses from the GPS coordinates of the taxis. We then aggregate all the addresses and create a space syntax directed weighted graph $G(V, E)$. Each street in the SanFrancisco map is represented as a single node in the graph as shown in Fig. 4. In order to better model the natural movement of the cabs in the city, we link streets by weighted edges approximating the flow rate between the street pairs in both directions. Edges in the graph are weighted according to one of the methods previously described in Section §IV.

4) Metrics and Experimental Setup: The performance of a forwarding algorithm is typically determined by two conflicting factors: (*i*) the success rate within a maximum message delivery delay; and (*ii*) the cost induced by the forwarding mechanism in terms of the number of message replicas in the network. We evaluate the performance of our three space syntax based algorithms, designed for each of the assumption categories, using these two performance indicators. Within each assumption category, we compare our forwarding algorithms to representative state-of-the-art algorithms within the same category.

In our evaluation, we compute the sequence of delay-optimal paths and deduce the delay obtained by the optimal path at all time. We uniformly combine all the observations of a trace amongst all sources, and a set of destinations D, and for every starting time (the time in seconds when the message m was generated by the source node S). We fix the number of destinations to 5 access points connected to the Internet and randomly placed on the map. We present an aggregated sample of observations as an average of 5 different runs via its empirical CDF (*i.e.,* different destination sets).



(a) SanFrancisco data set (b) DieselNet data set

Fig. 5. Normalized success rate distribution of SS-ED, SS-MP, and MobySpace using (a) SanFrancisco, and (b) DieselNet data sets

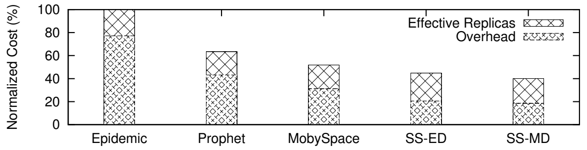
We plot the success rate of a forwarding algorithm normalized by the success rate of flooding (*i.e.,* Epidemic forwarding with unlimited storage/bandwidth) as a function of the message delivery delay. The detailed computation process can be found in [5]. In our experimental evaluation, we utilize the following metrics to evaluate a given forwarding algorithm: (*i*) the *normalized success rate within time* t: the probability to successfully deliver a message to all destination nodes within time t normalized by the same probability given by Epidemic routing (optimal success rate within the same time t), and (*ii*) the *normalized cost*: the fraction of contacts (*i.e.,* number of replica copies) used to deliver a message to all destinations normalized by the fraction of contacts used by the epidemic forwarding algorithm (the most expensive).

B. Results

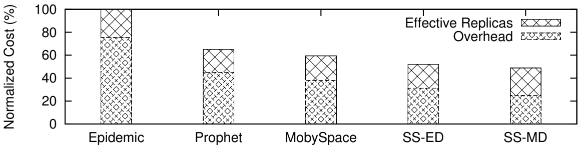
1) Evaluation of ASOF Algorithms: In this subsection, we compare our proposed space syntax based forwarding algorithms in the high assumption category, SS-ED and SS-MP with *MobySpace* [18] that falls within the same category.

Fig. 5 plots the CDF distribution of the normalized success rate using the two data sets, SanFrancisco (Fig. 5-a) and DieselNet (Fig. 5-b). We show that SS-ED (using the Earliest Delay utility) and SS-MP (using the Maximum Probability utility) achieve more than 95% success within 10 minutes. They outperform MobySpace in both data sets by roughly 10% and 21% for the SanFrancisco and the DieselNet experiments respectively. These results shows that under the same assumptions, space syntax techniques are more suitable to guide forwarding decisions in mobile opportunistic networks. This is mainly because space syntax helps in predicting the natural movement of the vehicles towards.

After evaluating the success rate of our space syntax algorithms, it is important to show the overhead injected in the network to achieve such performance. The cost of a forwarding algorithm, which we define as the fraction of contacts involved in the forwarding process, is as important as the success rate and the end-to-end delay metrics in opportunistic networks. We further classify the cost metric defined in the previous section into two categories: (*i*) effective message replicas



(a) SanFrancisco data set



(b) DieselNet data set

Fig. 6. Normalized effective and overhead cost generated by SS-ED, SS-MP, MobySpace, and Prophet using (a) SanFrancisco and (b) DieselNet data sets.

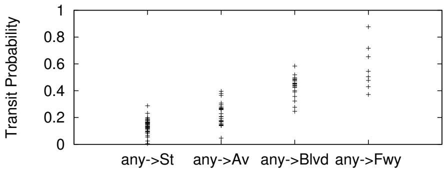


Fig. 7. Correlation between road types and transit rate

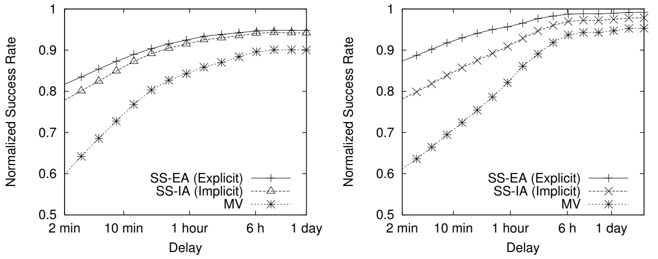
which represent the copies of the messages that have been used to deliver the message to any destination within a TTL of 1 hour, and (*ii*) the overhead defined as the copies of the message injected into the network without being effective within the same hour.

Figure 6 compares the normalized cost of our space syntax based algorithms with epidemic, Prophet, and MobySpace. We show that space syntax algorithms reduce cost by more than 50% compared to epidemic. MobySpace provides similar results by costing 56% of epidemic message exchanges, while Prophet reduces cost by more than 60% consisting of more than 40% overhead. Another interesting result given by space syntax forwarding algorithms is the percentage of effective message replicas compared to the total cost; space syntax algorithms generate roughly 50% of effective message replicas as compared to epidemic (effective replicas represent only fifth of the total cost), and MobySpace where effective message replicas represent more than the third of the total cost. Overall, space syntax techniques are effective in guiding messages towards the more suitable vehicles to reach the destinations since they avoid sending unnecessary messages to nodes moving away from the optimal paths to the destination.

2) Evaluation of LASOF Algorithms: We first evaluate our intuition in estimating transit rate between streets which consist of using street width as an estimate for its popularity. Fig. 7 plots the correlation between road types and the transit probability between streets in the SanFrancisco data set. We show that in general the probability to transit from any street type to a boulevard is greater than the same probability to transit to an avenue, which confirms our intuition discussed in the previous section (see section IV-C). We expect that a vehicle tends to take larger and faster roads to move from one place to another. However, there is no reason to believe that these two probabilities are linearly related. Indeed we are more interested in ranking (*i.e.,* prioritizing) transitions between streets. Hence, we define the *Ordering Error* (OE) as:

$$OE = \frac{|\{\ e_u \neq e_v \in E\ , P_I(E_u) \leq P_I(e_v)\ \text{ and }\ P_E(e_u) > P_E(e_v)\ \}|}{|E|(|E|-1)}$$

where P_I and P_E denote the implicit and explicit probabilities defined previously in Section IV-C. In the SanFrancisco data set, the ordering error is smaller than 0.22: the two probabilities provide the same ranking for more than 78% of total cases.



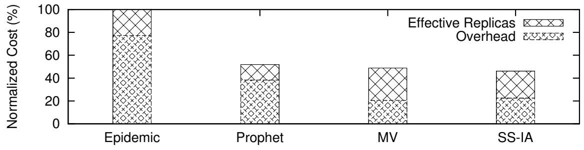
(a) SanFrancisco data set (b) DieselNet data set

Fig. 8. Normalized success rate distribution of SS-EA, SS-IA, and MV using (a) SanFrancisco, and (b) DieselNet data sets

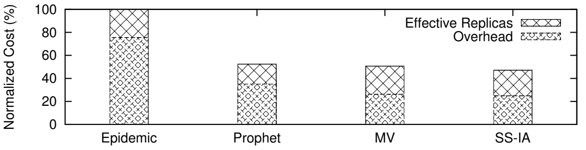
We now compare the performance of our space syntax based forwarding techniques to MV [4]. MV computes the probability for a node to successfully deliver a message to a set of destinations D within n hops based on the node's mobility pattern. We note that MV operates under more assumptions than ASOF since it assumes that nodes learn and store user mobility patterns.

Fig. 8 plots the distribution of normalized success rate of two space syntax LASOF algorithms SS-EA, SS-IA and the MV forwarding algorithm using two real-life data sets, SanFrancisco (Fig. 8-a) and DieselNet (Fig. 8-b). We show that the space syntax based forwarding schemes outperform the MV algorithm that operate under more assumptions. They achieve 5% to 8% improvement in the SanFrancisco data set and 17% to 25% using the DieselNet data set. SS-EA, which uses explicit traffic information to compute the transit rate between streets outperforms the implicit approach (*i.e.,* SS-IA). While there is a minor improvement using the SanFrancisco data set, an improvement of more than 8% was shown in the DieselNet data set where the explicit approach achieves 91% success rate within a 10 minutes timescale and the implicit approach does 83%. This can be explained by the size of the data set, where the DieselNet data set consists of only 22 nodes while the SanFrancisco data set contains more 486. Overall, the implicit approach which relies on environmental information such as street names achieves very comparable performance compared to the explicit approach.

We compare the normalized cost of the space syntax based forwarding schemes SS-EA and SS-IA with the MV and Prophet cost performance in Fig 9. We show similar results compared to the distribution cost of the ASOF algorithms; the



(a) SanFrancisco data set



(b) DieselNet data set

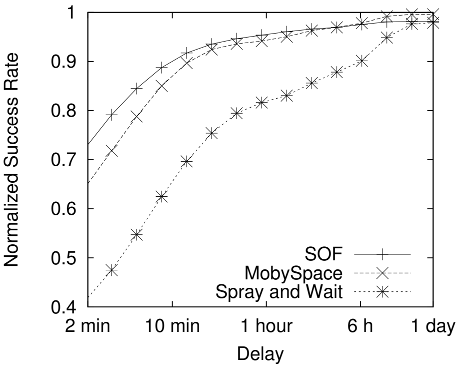
Fig. 9. Normalized effective and overhead cost generated by SS-EA, SS-IA, Prophet, and MV using (a) SanFrancisco and (b) DieselNet data sets

space syntax concepts help improve opportunistic forwarding decisions and guide the message to follow the shortest path to the destinations. In both SanFrancisco and DieselNet data sets, the effective message replicas of space syntax forwarding techniques represent roughly 50% of the traffic generated. Prophet however costs more than 60% which consists of more than 40% overhead. This result shows the relative efficiency of the space syntax forwarding techniques.

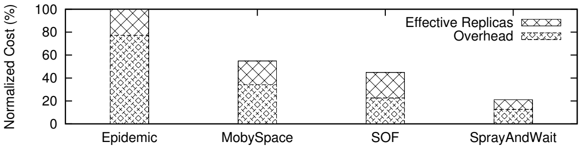
3) Evaluation of SOF Algorithm: We evaluate our space syntax based forwarding under the low assumption category. We study how the structure of the space syntax graph helps build opportunistic forwarding paths.

We compare the performance of our SOF algorithm to MobySpace [18] and Spray&Wait [24]. We note that the MobySpace algorithm operates under more assumptions than our SOF algorithm. We choose to compare with MobySpace to stress the gain achieved by using only environmental information, space syntax, to guide opportunistic forwarding. Spray&Wait, which operates under less assumptions than our SOF algorithm, is chosen to stress the cost gain achieved by SOF.

We plot in Fig. 10 the normalized success rate distribution (Fig. 10-a) and the normalized cost (Fig. 10-b) of our SOF algorithm. Note that we are comparing two algorithms using two different categories of assumptions. We show that SOF outperforms all other schemes in almost all considered timescales. It achieves 4% more success rate that the MobySpace algorithm success rate and more than 20% improvement compared to Spray&Wait success rate within a 10 minutes timescale. However, we show that MobySpace outperforms our SOF algorithm for large timescales. This is because MobySpace operates with more assumptions and uses nodes mobility patterns to find paths that SOF fails to find. In Fig. 10-b, we show that Spray&Wait forwarding algorithm generates roughly 50% less message replicas than our SOF algorithms and yields the best cost performance amongst all considered algorithms. This is an expected result since the Spray&Wait algorithm sets an upper bound on the number of copy replicas in the network and therefore controls the overall cost. But we observe in Fig. 10-a that this upper bound was not enough or not used efficiently to forward messages. Furthermore, Spray&Wait's assumption of knowing the exact number of nodes in such large scale network is unrealistic in large-scale environments considered in this work.

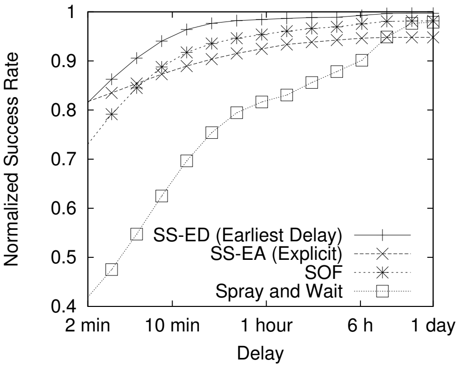


(a) Normalized success rate

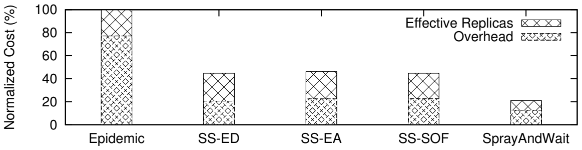


(b) Normalized effective and overhead cost

Fig. 10. Evaluation of the SOF performance using the SanFrancisco data set



(a) Normalized success rate



(b) Normalized effective and overhead cost

Fig. 11. Comparison of space syntax based techniques with Spray&Wait using the SanFrancisco data set

C. Summary and Discussions

Fig. 11 summarizes all our results and compares the performance of different space syntax algorithms across different assumption categories. We compare our forwarding performances to Spray&Wait. Comparing the three space syntax based forwarding schemes considered, we notice that the more assumptions we make the better the performance achieved. This property is not verified for larger timescales when the SOF algorithm outperforms the SS-EA algorithm. This shows the considerable gain introduced by space syntax that operates using information related to the map layout to reduce the cost and achieve efficient forwarding performances. In addition, for very large timescales (*i.e.,* after 8 hours), even the Spray&Wait success rate overpasses that of the SS-EA algorithm. This can be explained by the fact that the explicit approach (valid also for the implicit approach) estimates the best trajectory for the message to reach the destination and helps the forwarding algorithm achieve very good performances in very short delays (within 5 minute timescales). However, when this heuristic based approach fails to reach the destination, a more random technique such as SOF or even Spray&Wait fits better. Fig. 11-b confirms this explanation by showing a high percentage of effective message replicas; *i.e.,* less randomness in the forwarding decision.

VI. CONCLUSION

In this paper, we have proposed exploiting the space syntax paradigm to efficiently formulate message forwarding algorithms that adapt to the assumption spectrum of information available in large scale urban opportunistic networks. Space syntax helps predict natural movement based on the structure of the urban grid, and provides an accurate representation of domain popularity for any point in any given map by incorporating map characteristics and interest point location. We proposed three algorithms, all based on space syntax, that adapt to three categories representative of the spectrum of information available with respect to large scale urban environments. This spectrum ranges from node awareness of the static city map and its current location, to the location of destination Internet hot-spots, and node mobility profile. In the low assumption side of the spectrum, we have shown that our algorithm outperforms other opportunistic forwarding algorithms by 5% to 20% while reducing the cost by roughly 15%. In the high assumption range, we outperformed similar approaches such as MobySpace and MV by up to 20% while reducing the forwarding cost by more than 10%. Finally, we have also shown that more efficiency can be gained by only using space syntax even when compared with state-of-the-art solutions in the higher end of the spectrum.

This work represents the first steps towards leveraging space syntax concepts for enabling message forwarding in mobile opportunistic networks. We plan to extend this work by investigating the impact of space syntax at different network scales, examining mobile-to-mobile forwarding applications, and studying algorithms for infrastructure-to-mobile transfers.

REFERENCES

[1] Amr Al Jarhi, Hend Adel Arafa, Khaled A. Harras, and Sherif G. Aly. Rethinking opportunistic routing using space syntax. In *Proceedings of the 6th ACM workshop on Challenged networks*, CHANTS '11, pages 21–26, New York, NY, USA, 2011. ACM.

[2] John Burgess, Brian Gallagher, David Jensen, and Brian Neil Levine. Maxprop: Routing for vehicle-based disruption-tolerant networking. In *Proceedings of IEEE Infocom*, 2006.

[3] John Burgess, Brian Neil Levine, Ratul Mahajan, John Zahorjan, Aruna Balasubramanian, Arun Venkataramani, Yun Zhou, Bruce Croft, Nilanjan Banerjee, Mark Corner, and Don Towsley. CRAWDAD data set umass/diesel (v. 2008-09-14), September 2008.

[4] B. Burns, O. Brock, and B.N. Levine. Mv routing and capacity building in disruption tolerant networks. In *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*, volume 1, pages 398–408. Ieee, 2005.

[5] A. Chaintreau, A. Mtibaa, L. Massoulié, and C. Diot. The diameter of opportunistic mobile networks. In *Proceedings of ACM CoNext*, 2007.

[6] P.C. Cheng, K.C. Lee, M. Gerla, and J. Härri. Geodtn+nav: Geographic dtn routing with navigator prediction for urban vehicular environments. *Mobile Networks and Applications*, 15(1):61–82, 2010.

[7] San Francisco chronicle editors. Sf gate: Traffic: Live traffic reports, articles and information. http://www.sfgate.com/traffic/, April 2012.

[8] Elizabeth M. Daly and Mads Haahr. Social network analysis for routing in disconnected delay-tolerant manets. In *MobiHoc '07: Proceedings of the 8th ACM international symposium on Mobile ad hoc networking and computing*, pages 32–40, New York, NY, USA, 2007. ACM.

[9] Google Developers. The google geocoding api - google maps api web services mdash; google developers. https://developers.google.com/maps/documentation/geocoding/, 2012.

[10] M. Grossglauser and D. Tse. Mobility increases the capacity of ad hoc wireless networks. *IEEE/ACM Trans. on Net.*, 10(4):477–486, 2002.

[11] Bo Han, Pan Hui, and Aravind Srinivasan. Mobile data offloading in metropolitan area networks. *SIGMOBILE Mob. Comput. Commun. Rev.*, 14(4):28–30, November 2010.

[12] K.A. Harras and K.C. Almeroth. Controlled flooding in disconnected sparse mobile networks. *Wireless Communications and Mobile Computing*, 9(1):21–33, 2009.

[13] B. Hillier and J. Hanson. *The social logic of space*, volume 281. Cambridge University Press Cambridge, 1984.

[14] Pan Hui, Jon Crowcroft, and Eiko Yoneki. Bubble rap: Social-based forwarding in delay tolerant networks. *IEEE Transactions on Mobile Computing*, 99(PrePrints), 2010.

[15] V. Kostakos, E. O'Neill, and A. Penn. Designing urban pervasive systems. *Computer*, 39(9):52–59, 2006.

[16] Vassilis Kostakos. Space syntax and pervasive systems. In Bin Jiang and Xiaobai Yao, editors, *Geospatial Analysis and Modelling of Urban Structure and Dynamics*, volume 99 of *GeoJournal Library*, pages 31–52. Springer Netherlands, 2010. 10.1007/978-90-481-8572-6-3.

[17] Kyunghan Lee, Joohyun Lee, Yung Yi, Injong Rhee, and Song Chong. Mobile data offloading: how much can wifi deliver? In *Proceedings of the ACM SIGCOMM 2010 conference*, SIGCOMM '10, pages 425–426, New York, NY, USA, 2010. ACM.

[18] J. Leguay, T. Friedman, and V. Conan. Dtn routing in a mobility pattern space. In *Proceedings of the 2005 ACM SIGCOMM workshop on Delay-tolerant networking*, pages 276–283. ACM, 2005.

[19] Anders Lindgren, Avri Doria, and Olov Schelén. Probabilistic routing in intermittently connected networks. *SIGMOBILE Mob. Comput. Commun. Rev.*, 7(3):19–20, 2003.

[20] Anders Lindgren, Avri Doria, and Olov Schelén. Probabilistic routing in intermittently connected networks. In *Proceedings of The First International Workshop on Service Assurance with Partial and Intermittent Resources (SAPIR 2004)*, August 2004.

[21] C. Mayer and O. Waldhorst. Offloading Infrastructure using Delay Tolerant Networks and Assurance of Delivery. In *Proceedings of IFIP Wireless Days*, October 2011.

[22] Abderrahmen Mtibaa, Martin May, Mostafa Ammar, and Christophe Diot. Peoplerank: Social opportunistic forwarding. In *Proceedings of IEEE INFOCOM (mini conference)*. IEEE, 2010.

[23] Michal Piorkowski, Natasa Sarafijanovic-Djukic, and Matthias Grossglauser. CRAWDAD data set epfl/mobility (v. 2009-02-24). Downloaded from http://crawdad.cs.dartmouth.edu/epfl/mobility, February 2009.

[24] Thrasyvoulos Spyropoulos, Konstantinos Psounis, and Cauligi S. Raghavendra. Spray and wait: an efficient routing scheme for intermittently connected mobile networks. In *Proceedings of the 2005 ACM SIGCOMM workshop on Delay-tolerant networking*, WDTN '05, pages 252–259, New York, NY, USA, 2005. ACM.

[25] Amin Vahdat and David Becker. Epidemic routing for partially-connected ad hoc networks. Technical Report CS-2000-06, University of California San Diego, Jul 2000.

[26] W. Zhao, M. Ammar, and E. Zegura. A message ferrying approach for data delivery in sparse mobile ad hoc networks. In *Proceedings of MobiHoc*, pages 187–198, 2004.