

Opportunistic Spatial Gossip over Mobile Social Networks

Augustin Chaintreau^{*}
Thomson
Paris

Pierre Fraigniaud[†]
CNRS
Universite Paris Diderot

Emmanuelle Lebhar[‡]
CNRS
Universite Paris Diderot

ABSTRACT

This paper investigates how the principles underlying online social network services could be used to take advantage of node mobility in an opportunistic manner. As an example, we show how to take advantage of opportunistic contacts between mobile phones that run an online social network service. Our model includes static nodes, and mobile nodes which follow random walks. As in an online network service, we assume that each node can only communicate with a small subset of others nodes (called its *mates*) in addition to its geographical neighbors. Here we prove that, in such context, a simple connection scheme enables to execute sophisticated tasks (*e.g.*, routing) and mechanisms (*e.g.*, spatial gossip), while using only opportunistic communication and communication between mates. In other words, our results show that future online social networks can exploit mobility as long as they forget connections appropriately.

Categories and Subject Descriptors

C.2.2 [Network Protocols]: Routing protocols

General Terms

Algorithms

Keywords

Spatial gossip, network navigability, delay-tolerant network, pocket-switched networks, online social networks

^{*}augustin.chaintreau@thomson.net

[†]Pierre.Fraigniaud@liafa.jussieu.fr. Additional supports from the COST Action 295 "DYNAMO", from the ANR projects "ALADDIN", and INRIA project "GANG".

[‡]Emmanuelle.Lebhar@liafa.jussieu.fr. Additional supports from the COST Action 295 "DYNAMO", from the ANR projects "ALADDIN", and INRIA project "GANG".

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

WOSN'08, August 18, 2008, Seattle, Washington, USA.
Copyright 2008 ACM 978-1-60558-182-8/08/08 ...\$5.00.

1. INTRODUCTION

It is now a few years since researchers became aware that, in a near future, most users will access the Internet via their cell-phone or another portable device [10, 12]. In such a new networking environment, which includes node mobility by default, routing may sometimes benefit from favorable timing circumstances, because mobility create temporal paths between nodes. Opportunistic routing, based on this principle, has been shown to dramatically increase the capacity of dense networks that rely on local communication to avoid infrastructure [4]. Unfortunately it seems difficult to design algorithms that take advantage of this bandwidth when the mobility pattern is not known *a priori* [5, 2, 7].

This paper claims that, in contrast to the above, Online Social Networks (OSN) can take advantage of nodes mobility by appropriate connections maintenance. Services of OSNs have been widely adopted by millions of Internet users; they assume that each person is more likely to publish, exchange, and retrieve some information through a set of privileged acquaintances (*i.e.*, known as friends) than through anyone else. Following the same principle, we assume that users in a mobile network may only communicate with a few number of nodes (called their *mates*), in addition to their geographical neighbors. We assume that mates are forgotten and replaced by other nodes met in function of time and opportunistic encounters. Our results show that, under some conditions on this process, the induced topology supports well decentralized routing scheme (*i.e.* greedy routing) as well as spatial gossip mechanism.

We model a network composed of two types of nodes:

- a set of *static* nodes positioned at the vertices of a lattice, and that can communicate locally with all their neighboring nodes in the lattice;
- a set of *moving* nodes that are performing random walks from vertex to vertex in the lattice.

We assume that each static node maintains some privileged relationships with one or few mobile nodes that it has seen in the past. We call these latter mobile nodes the *mates* of the static node. A mobile node can be the mate of at most one static node, and a static node can have up to m mates, where $m \geq 1$ is a parameter. A static node is assumed to be able to communicate with its mates, at all times. (For instance by remembering the cellphone numbers of the moving nodes, if nodes are all equipped with such a device). As mobile nodes follow their journey through the lattice, they can be reached by the static nodes for which they are mates, in which case they can give information about their current

locations. However, static nodes are not aware of the positions of all the other mobiles nodes.

In this setting, one wishes to maintain a favorable connection topology to exploit connection with mobile nodes to improve routing and/or gossip communication protocols. The challenge is to use the mobility of the nodes implicitly, *i.e.*, without keeping track of their current positions, nor being able to predict their future locations. This issue relates with a property known as *navigability* [11], which is the ability of routing efficiently in the absence of global information about the network topology. Our work builds upon a model that we have recently introduced to explain the emergence of the navigability property in social networks (see [1]). Here we show how to use the same property to derive an opportunistic connection scheme that enables routing to perform efficiently. This scheme defines how static nodes select and refresh their set of mates. In a nutshell, our scheme simply requires the static nodes to reinitialize their (mobile) mates according to a simple random process consisting in (1) deciding when a mate should be forgotten, and (2) how a forgotten mate should be replaced by a new one. We show that our scheme not only enables to implement routing efficiently, but can also be used to implement spatial gossip [6, 9]. As a consequence, several distributed computing problems such as resource location, and construction of approximate spanning tree could be tackled using our scheme.

To the best of our knowledge, these results are the first ones proving that complex algorithmic properties of social networks can be constructed from opportunistic contacts between nodes in a mobile environment. They also prove that OSNs, provided that connections are appropriately maintained, can take advantage of mobility and local bandwidth.

2. OPPORTUNISTIC CONNECTION SCHEME

2.1 Description

We study a network where nodes are composed of two sets: a set of static nodes identified by their spatial coordinates, and a set of mobile nodes which follow random walks. Each static node acquires up to m mates in the set of mobile nodes. A mobile node can be the mate of at most one static node (see Section 5). Mates are acquired opportunistically by static nodes when they happen to meet mobile nodes. Precisely, if a static node x that has less than m mates happens to meet a free mobile node y , (*i.e.*, a mobile node that is the mate of no static node), then y becomes the mate of x . Then, y carries on its journey, and can inform x of its current position on demand (this position may for instance be obtained from static nodes nearby). An important feature of our opportunistic connection scheme is that a mate does not remain such forever, and a forgetting process between the static nodes and their mates enables a renewal of the set of mate. This forgetting process is depending on the age of the mate connection. Assume that the mobile node y became the mate of the static node x at time t_0 . Then, every Δt period of time, the static node x determines whether y should remain its mate. More precisely, at time $t_0 + a \cdot \Delta t$, with $a \geq 1$, node x forgets y independently from the past with probability $\phi(a)$. This probability only depends on the age of the connection between x and its mate.

The choice of ϕ has a tremendous impact on the distri-

bution of connections between the static nodes and their mates, and, as a consequence, on the efficiency of protocols taking advantage of these connections. A typical example of such protocol is *greedy routing*, where a message for a static target \mathbf{u} is forwarded by each node to its neighbor that has the smallest spatial distance to the target \mathbf{u} . This algorithm can exploit the shortcuts resulting from connections between static nodes and their mates. This is however true only if these shortcuts are distributed appropriately [11]. Here the distribution of the shortcuts depends on the movements of the mobile nodes, and on the function ϕ . For instance, if the mates connections are forgotten too rapidly, then the distance between a static and a mobile node connected by a shortcut is small, and hence these connections would not speed up routing. On the other hand, if a node keeps its mates too long, then the position of a mate would be distributed uniformly, which is known to be of poor help as far as efficient routing is concerned (see [11]).

2.2 Setting of the forgetting function

We show that it is possible to choose ϕ to obtain the desired distributions of distance between a static node and its mates. We consider the following setting. Static nodes are assumed to be located on the vertices of the k -dimensional lattice \mathbb{Z}^k , while the mobile nodes are assumed to perform random walks in \mathbb{Z}^k . More precisely, at each step, a mobile node at position $\mathbf{x} \in \mathbb{Z}^k$ chooses a vector $\mathbf{r} \in \{-1, 1\}^k$, uniformly at random, and move to $\mathbf{x} + \mathbf{r}$. A shortcut is therefore a connection between a static node \mathbf{x} occupying vertex $(x_1, \dots, x_k) \in \mathbb{Z}^k$ and a mate \mathbf{y} currently occupying vertex $(y_1, \dots, y_k) \in \mathbb{Z}^k$. The length of the shortcut is $d = \sum_{i=1}^k |x_i - y_i|$. In this context, the density of distributions of shortcut lengths that are of specific interest are of the type

$$\frac{1}{d^{\rho \cdot k}}, \quad \text{for } \rho \geq 1.$$

The following results show that one can approach all these distributions by using appropriate forgetting functions ϕ .

First, we assume $\rho = 1$. Let $\|\cdot\|$ denote a norm on \mathbb{R}^k .

Theorem 1 *For any $\varepsilon > 0$, there exists a forgetting function $\phi_{1,\varepsilon}$, and three positive constants d_0 , c , and c' such that for any $\mathbf{d} = (d_1, \dots, d_k)$ with $|d_i| \geq d_0$ for all $i \in \{1, \dots, k\}$, the density f of the shortcuts distribution satisfies*

$$\frac{c}{\|\mathbf{d}\|^k \ln^{1+\varepsilon} \|\mathbf{d}\|} \leq f(\mathbf{d}) \leq \frac{c' \cdot \ln^{k/2} \|\mathbf{d}\|}{\|\mathbf{d}\|^k \ln^{1+\varepsilon} \|\mathbf{d}\|}.$$

This result follows from Theorem 1 in [1], where the function $\phi_{1,\varepsilon}$ is defined as:

$$\phi_{1,\varepsilon}(a) = \begin{cases} 0 & \text{for } a = 0, 1 \text{ or } 2, \\ 1 - \frac{a-1}{a} \left(\frac{\ln(a-1)}{\ln a} \right)^{1+\varepsilon} & \text{for } a > 2. \end{cases}$$

Note that the expression of $\phi_{1,\varepsilon}$ for large “age” a , may be written as $\phi_{1,\varepsilon}(a) = \frac{1}{a} + o(\frac{1}{a})$.

We now assume $\rho > 1$, we then have:

Theorem 2 *For any $\rho > 1$, there exists a forgetting function ϕ_ρ , and three positive constants d_0 , c , and c' such that for any $\mathbf{d} = (d_1, \dots, d_k)$ with $|d_i| \geq d_0$ for all $i \in \{1, \dots, k\}$, the density f of the shortcuts distribution satisfies*

$$\frac{c}{\|\mathbf{d}\|^{\rho k}} \leq f(\mathbf{d}) \leq \frac{c' \cdot \ln^{k\rho/2} \|\mathbf{d}\|}{\|\mathbf{d}\|^{\rho k}}.$$

SKETCH OF PROOF. We prove the result for the ℓ_∞ norm, denoted by $\|\cdot\|$. Let $\rho > 1$, and consider one of the shortcuts. Denote by $(A(t))_{t \geq 0}$ the ages of the current mate connection at time t . $A(t)$ follows the evolution of a Markov chain. This Markov chain is recurrent positive if and only if the series $(\prod_{i=1}^a (1 - \phi(i)))_{a \geq 0}$ converges. It then admits the following stationary distribution:

$$\pi(a) = \frac{\prod_{i=1}^a (1 - \phi(i))}{1 + \sum_{j \geq 1} \prod_{i=1}^j (1 - \phi(i))}.$$

Let us consider $\nu > 1$. Later, we should fix the value of ν to $1 + \frac{k(\rho-1)}{2}$. We define

$$\psi_\nu(i) = \begin{cases} 0 & \text{for } i = 1, \\ 1 - (\frac{i-1}{i})^\nu & \text{for } i > 1. \end{cases}$$

We have $\prod_{i=1}^a (1 - \psi_\nu(i)) = a^{-\nu}$. It is a convergent series, and $\pi(a) = \frac{c}{a^\nu}$ with $c = (1 + \sum_{j \geq 1} j^{-\nu})$.

In the stationary regime, as the movement of the mate is independent from the forgetting process of the stationary node, we have:

$$f(\mathbf{d}) = \sum_{a \geq 0} \pi(a) \cdot \Pr\{X(a) = \mathbf{d}\}. \quad (1)$$

where $(X(t))_{t \geq 0}$ is a random walk in \mathbb{Z}^k starting in $(0, 0, \dots, 0)$ at time $t = 0$. The following lemma deals with random walk on \mathbb{Z} , it has been proved in [1]

Lemma 1 *Let X be a random walk in \mathbb{Z} , and $C > 0$. For any $\zeta > 0$, there exists $d_0 > 1$ such that, for any $d \in \mathbb{Z}$, $|d| \geq d_0$, and any $a \geq \frac{d^2}{C \cdot \ln|d|}$, we have*

$$\begin{aligned} \Pr\{X(a) = d\} &\geq (1 - \zeta) \cdot \sqrt{\frac{2}{\pi \cdot a}} \exp\left(-\frac{3d^2}{4a}\right) \\ \Pr\{X(a) = d\} &\leq (1 + \zeta) \cdot \sqrt{\frac{2}{\pi \cdot a}} \exp\left(-\frac{d^2}{4a}\right). \end{aligned}$$

Since $|d_i| \geq d_0$ for all $i \in \{1, \dots, k\}$, we can apply Lemma 1 to every dimension because the X_i s are mutually independent random variables. As a consequence, if $a \geq \frac{3}{4} \|\mathbf{d}\|^2$ then, for every i ,

$$\Pr\{X_i(a) = d_i\} \geq \frac{1 - \zeta}{e} \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{a}}.$$

Therefore,

$$\begin{aligned} \Pr\{X(a) = \mathbf{d}\} &= \Pr\{X_1(a) = d_1, \dots, X_k(a) = d_k\} \\ &\geq \left(\frac{1 - \zeta}{e} \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{a}}\right)^k. \end{aligned}$$

From Eq. (1), we get

$$\begin{aligned} f(\mathbf{d}) &\geq \left(\frac{1 - \zeta}{e} \sqrt{\frac{2}{\pi}}\right)^k \sum_{a \geq \frac{3}{4} \|\mathbf{d}\|^2} \frac{c}{a^{\nu + (k/2)}} \\ &\geq \left(\frac{1 - \zeta}{e} \sqrt{\frac{2}{\pi}}\right)^k \frac{c}{\nu + (k/2) - 1} \frac{1}{(3/4 \cdot \|\mathbf{d}\|^2)^{\nu + (k/2) - 1}} \\ &\geq \left(\frac{1 - \zeta}{e} \sqrt{\frac{2}{\pi}}\right)^k \frac{c}{\nu + (k/2) - 1} \left(\frac{4}{3}\right)^{\nu + (k/2) - 1} \frac{1}{\|\mathbf{d}\|^{2\nu + k - 2}} \end{aligned}$$

which proves the lower bound by choosing $\nu = 1 + \frac{k(\rho-1)}{2}$. The proof of the upper bound is similar. It uses Chernoff bound, and Lemma 1, in a way similar to the proof in [1]. We omit it due to space constraints. \square

Note that in the two results above, there is a polylogarithmic amount of slack between the lower and the upper bound on the distribution f . The multiplicative term in the upper bound comes from the way the Chernoff bound is used in the proof. As shown in the next section, this slack does not impact the results on routing or gossip mechanism, which rely only on the lower bound.

3. OPPORTUNISTIC ROUTING

In this section, we show how the opportunistic connection scheme described above allows the same navigability property as the one shown in [11] for a lattice enhanced with shortcuts chosen with a probability that depends on the distance. We analyze the performances of *greedy routing*, that simply selects the next hop towards a target as the node that is connected to the current node and is the closest to the target according to its position in the lattice. More precisely, the target is a fixed static node \mathbf{u} identified by its coordinate (u_1, \dots, u_k) in the lattice \mathbb{Z}^k . A static node $\mathbf{x} = (x_1, \dots, x_k)$ selects among all its neighbors (*i.e.*, among its $2k$ neighbors in \mathbb{Z}^k and its m mates) the one that is the closest to \mathbf{u} in the lattice, and forwards the message to it. This simply requires \mathbf{x} to ask for the current positions of its mates. If the node receiving the message is a mobile node \mathbf{y} (*i.e.*, a mate \mathbf{y} of \mathbf{x}), then, upon reception of the message from \mathbf{x} , \mathbf{y} transmits it immediately to the static node \mathbf{x}' occupying the vertex currently occupied by \mathbf{y} . (In practice, \mathbf{y} would transmit the message to the static node that is closest to \mathbf{y}). The routing then carries on at \mathbf{x}' .

By a similar analysis as in [1], one can show that greedy routing finds source-destination paths whose length grows polylogarithmically as a function of the distance between the source and the destination in the lattice, insuring both efficiency and scalability. Precisely, we have:

Theorem 3 *For any $\epsilon > 0$, with one mate per static node in \mathbb{Z}^k , if mates are forgotten with probability $\phi_{1,\epsilon}$, then the expected number of steps of greedy routing from any source node to any target node at distance d is $O(\ln^{2+\epsilon} d)$.*

4. OPPORTUNISTIC SPATIAL GOSSIP

Gossip-based communication protocols are efficient methods for designing robust and scalable communication schemes in large distributed systems [3]. In essence, a gossip-based communication protocol, or simply gossip protocol for short, is based upon two mechanisms: a *node selection* mechanism and a *message selection* mechanism. Roughly, the node selection mechanism guides the way rumors can be spread over a network, by choosing appropriate recipients for communications, from every node. The message selection mechanism guides the choice, performed by every node, of which information has to be transferred to the recipients selected by the node selection mechanism. Hence, at an abstract level, a gossip protocol performs as follows: every node \mathbf{u} repeats infinitely the following:

- Node selection: \mathbf{u} selects q nodes $\mathbf{v}_1, \dots, \mathbf{v}_q$;
- Message selection: \mathbf{u} sends \mathcal{M}_i to \mathbf{v}_i , $i = 1, \dots, q$;

where \mathcal{M}_i denotes a message prepared by \mathbf{u} for node \mathbf{v}_i , and $q \geq 1$ is a fixed parameter. To enforce robustness against faults, at a node \mathbf{u} , the choices of the nodes $\mathbf{v}_1, \dots, \mathbf{v}_q$ are

oblivious from the past, and these choices are typically random. The communications between each node \mathbf{u} and its selected nodes \mathbf{v}_i s are achieved via some underlying point-to-point communication protocol which allows these pairs of nodes to communicate, regardless of their distance.

Spatial gossips [6, 9, 8] are gossip protocols in which the node selection mechanism uses the spatial distribution of the nodes (*e.g.*, over a k -dimensional Euclidean space). In [6] a spatial gossip protocol is described for resource location. This protocol allows any node \mathbf{x} to find a copy of a resource present in the network in time polylogarithmic as a function of the distance d between \mathbf{x} and the nearest copy of the resource. Our opportunistic connection scheme enables to mimic this protocol. Actually, we focus only on the node selection protocol. The message selection is identical to the one in [6]. In the following, we use the definition of propagation time defined in [6]. We present two node selection mechanisms: the first, called *geographic gossip*, uses the links of the lattice as well as the connections with the mates, while the second is called *social gossip* as it only uses the connections with the mates.

4.1 Geographic gossip

A straightforward manner to perform spatial gossip is to flood the network. We call *geographic gossip* the resulting protocol. At each step of geographic gossip, every node \mathbf{x} selects one of its neighbors in the lattice or one of its m mates uniformly at random, and communicates with it. Since any neighbor of a node is chosen on average after a constant time, the propagation time of a message in this protocol is a constant multiplied by the distance between the nodes in the lattice augmented with all the shortcuts. Thus, as a consequence of Theorem 3, we have.

Theorem 4 *Assume there is one mate per static node in \mathbb{Z}^k and that mates are forgotten with probability $\phi_{1,\epsilon}$, for some $\epsilon > 0$. Geographic gossip satisfies that, whenever a node initiates a request, its expected propagation time is $O(\ln^{2+\epsilon} d)$ to any node at distance d .*

Actually, generalizing the results in [14, 15] to infinite lattices would enable to show that whenever a node initiates a request for a resource whose one copy is at distance d from it, the expected time to receive this resource is $O(\ln^{1+\epsilon} d)$.

Flooding may not be desirable, especially for lattices of high dimension. We describe hereafter a more practical spatial gossip protocol using our opportunistic connection scheme.

4.2 Social gossip

The *social gossip* protocol uses only connections between the mates (as, for instance, in a social network where information is only exchanged among friends). We assume that the mates of each node are ordered and that this order is preserved by the choices over time. That is, every node \mathbf{x} keeps an array of length m . The i th cell of this array contains the address of the i th mate. If this mate is forgotten, then the content of the cell is replaced by the address of the new mate replacing the former one.

At step t of the social gossip protocol, every node \mathbf{x} communicate with its mate whose address is contained in the cell $t \bmod m$ of its array, *i.e.*, in a round-robin manner.

Theorem 5 *Assume that mates are forgotten with probability ϕ_ρ in \mathbb{Z}^k , $\rho \in (1, 2)$. Social gossip satisfies that, whenever a node initiates a request, its propagation time is $O(\ln^2 d)$ to any node at distance d w.h.p. if there are $\Omega(\ln^2 d)$ mates per static node.*

SKETCH OF PROOF. The proof of this theorem follows the proof of the spatial gossip protocol of [6]. This latter proof indeed applies to our connection scheme scenario for the following three reasons.

- During a fixed time interval whose length is not exceeding the number of mates m , the node selection mechanism in [6] and the one in this paper follow the same distribution.
- Although the model in [6] assumes nodes selected according to a given distribution $\propto 1/d^{\rho k}$, the proof naturally extends if this distribution is $\Omega(1/d^{\rho k})$, which is the lower bound of Theorem 2.
- Although the lower bound of Theorem 2 applies only for $\mathbf{d} = (d_1, \dots, d_k)$ with $|d_i| \geq d_0$ for all i , we show how to overcome this restriction.

For any $\mathbf{c} \in \mathbb{Z}^k$, and $r \geq 0$, let $B(\mathbf{c}, r)$ be the ball of center \mathbf{c} and radius r , *i.e.*,

$$B(\mathbf{c}, r) = \{\mathbf{x} \in \mathbb{Z}^k, \|\mathbf{c} - \mathbf{x}\|_1 \leq r\}.$$

Let d_0 be the constant defined for ρ in Theorem 2.

Claim 1. If $r \geq 3kd_0$ then for any nodes $\mathbf{u}, \mathbf{v} \in \mathbb{Z}^k$,

$$|B(\mathbf{v}, r) \cap \{\mathbf{w} \in \mathbb{Z}^k : \min_i |w_i - u_i| \geq d_0\}| \geq \frac{1}{(3k)^k} \frac{2^k}{k!} r^k.$$

To establish the claim, let

$$S = \{\mathbf{x} = (x_1, \dots, x_k) \mid x_i \in \{-1, 0, +1\}\}.$$

For any $\mathbf{x} \in S$, let $B_{\mathbf{x}} = B(\mathbf{v} + \frac{2r}{3k}\mathbf{x}, \frac{r}{3k})$. We have $B_{\mathbf{x}} \subseteq B(\mathbf{v}, r)$ for any $\mathbf{x} \in S$. We have

$$|v_i - u_i + \frac{2r}{3k} \text{sign}(v_i - u_i)| \geq \frac{2r}{3k} \geq 2d_0,$$

where $\text{sign}(a)$ equals 1 if $a \geq 0$ and equals 0 otherwise.

Let $\mathbf{x} \in S$, with $x_i = \text{sign}(v_i - u_i)$. We have

$$B_{\mathbf{x}} \subseteq \{\mathbf{w} \in \mathbb{Z}^k : |w_i - u_i| \geq d_0, \forall i = 1, \dots, k\}.$$

Indeed, $w_i - v_i = v_i - u_i + \frac{2r}{3k} \text{sign}(v_i - u_i) + \xi_i$ where $|\xi_i| \leq \frac{r}{3k}$, which completes the proof of the claim.

We can now sketch the proof scheme of [6] to highlight how these changes still allows us to get a polylogarithmic propagation time w.h.p.. Due to lack of space, we do not enter into the details of parameter tuning required for the complete proof.

Consider two nodes \mathbf{u} and \mathbf{v} at some distance d in the lattice. The proof proceeds by induction on d to show that the message has arrived from \mathbf{u} to \mathbf{v} within less than $O(\ln^2 d)$ time steps with probability at least $1 - \alpha/\ln^\kappa d$, for some constants $\alpha, \kappa > 0$. If $d = (3kd_0)^2$, the claim holds for appropriate constants in big- O . Assume now $d > (3kd_0)^2$.

We consider the set of diameter d that contains \mathbf{u} and \mathbf{v} and two balls $B_{\mathbf{u}}$ and $B_{\mathbf{v}}$ of radius $d^{\rho/2}$ that contain respectively \mathbf{u} and \mathbf{v} . Since $\rho < 2$, we have $d^{\rho/2} < d$. By induction hypothesis, any node \mathbf{x} in $B_{\mathbf{u}}$ received the communication of \mathbf{u} within less than $O((\rho/2)^2 \ln^2 d)$ steps, and \mathbf{v} received the

communication of any node in B_v in $O((\rho/2)^2 \ln^2 d)$ steps with probability at least $1 - \alpha/\ln^\kappa d$, each on distinct time intervals, and therefore independently. Let I be the time interval starting at time t , divided in three intervals I_1, I_2, I_3 of length $\beta(\rho/2)^2 \ln^2 d$ as follows:

$$\begin{aligned} I_1 &= [t, t + \beta(\rho/2)^2 \ln^2 d] \\ I_2 &= (t + \beta(\rho/2)^2 \ln^2 d, t + 2\beta(\rho/2)^2 \ln^2 d] \text{ and} \\ I_3 &= (t + 2\beta(\rho/2)^2 \ln^2 d, 3\beta(\rho/2)^2 \ln^2 d], \end{aligned}$$

for some appropriate constant $\beta > 0$. Let P be the probability that a node \mathbf{s} of B_u call a node \mathbf{t} of B_v , with \mathbf{s} and \mathbf{t} satisfying that $\min_i |s_i - t_i| \geq d_0$. Since the distance from \mathbf{s} to \mathbf{t} is at most $3d$, P is at least $1/(3d)^{k\rho}$. From Claim 1, the number of such nodes \mathbf{t} in B_v is at least $\frac{1}{(3k)^k} \frac{2^k}{k!} d^{k\rho/2}$ because $d^{\rho/2} \geq 3kd_0$. Therefore, the probability that no node of B_u call a node of B_v during I_2 is at most :

$$\left(1 - \frac{1}{(3k)^k} \frac{2^k}{k!} \frac{d^{k\rho/2}}{(3d)^{k\rho}}\right)^{|B_u| \cdot \beta(\rho/2)^2 \ln^2 d}.$$

Indeed, the total number of mates per node is $\Omega(\ln^2 d)$ and the trials are thus independent during this time interval. Since B_u is of size $\Theta(d^{k\rho/2})$, we obtain that this probability is at most $e^{-\gamma \ln^2 d}$, for some constant γ depending on k, ρ and β

The three events that

1. all nodes in B_u received the communication of \mathbf{u} during I_1 ,
2. some node of B_u communicated to a node of B_v during I_2 , and
3. \mathbf{v} received all communications from B_v during I_3

are independent. The probability that these three events happen is at least $(1 - \alpha(\rho/2)^2/\ln^\kappa d)^2 \cdot (1 - e^{-\gamma \ln^2 d})$. Since $\rho < 2$, one can show that this is at least $1 - \alpha/\ln^\kappa d$ for appropriate $\alpha, \kappa, \beta, \gamma$. This proves the induction and the theorem since the length of the time interval I is $O(\ln^2 d)$. \square

5. DISCUSSION

This paper demonstrates that a mobile social network can benefit from efficient decentralized routing, and gossip communication, when nodes maintain connections with other nodes that they have met in the past. The key ingredient for this property to hold consists in the combination of two techniques:

1. mobile nodes are assumed to follow random walks, and
2. a specific forgetting scheme is tuned in order to target appropriate shortcut distributions.

As a consequence, nodes may benefit from a communication scheme whose propagation time preserves locality while exploiting the fast expansion of gossip, in a manner similar to spatial gossip [6].

These results indicate that an Online Social Network with mobile users is a promising environment for distributed computing. In future OSNs, mobile users could take advantage

of their mobility and opportunistic bandwidth between their devices to become mates or exchange information [13]. As different devices come into communication range, the user (or an application) can decide not to communicate with all of these devices, but rather to update a set of local mates to contact for timely and relevant information about the immediate environment. Our work points out that, if mates are forgotten and replaced appropriately, a device may efficiently find all relevant information through its mates. In particular, one does not need to control the mobility or even to have a global view of users's positions. The applications of spatial gossip (resource location, permutation routing, approximate minimum spanning tree, etc.), as shown in [6], are beyond the scope of this note. We have nevertheless demonstrated strong evidences that most of them should hold within the same context.

The results we present have been obtained under a few assumptions that we know discuss in more details. First, we assume that a mobile node can be the mate of at most one static node, and that as soon as a mate is forgotten, a new mate can be found immediately. This is needed to insure that the mates for different static nodes are always distinct, so that their positions at any time are independent variables. In practice, we expect that if the density of mobile nodes is sufficiently large, choosing the first available mate or simply the first mate seen would still create a favorable topology. We now discuss other assumptions related to mobility.

5.1 What about if all nodes are mobiles?

In this paper, our model comprises static nodes (distributed on a lattice) and mobile nodes (which follow random walks). In practice however, all nodes may be moving, which makes the problem of using opportunistic shortcut even more complex. Indeed, if all nodes are moving, one cannot anymore assume that nodes are in contact with a fixed set of local neighbors. It does not seem straightforward to construct a decentralized routing scheme as the positions of the targets are evolving. Nevertheless, other results presented in this paper are likely to extend. First, by assuming that the movements of the nodes are mutually independent, then the shortcut distribution would simply result from the forgetting mechanism applied to the combination of two independent random walks. Properties satisfied by the combination of two random walks actually resemble those of a single random walk, and therefore the resulting shortcut distribution should be similar. Social gossip (similarly to spatial gossip) does not require using lattice connections, neither targeting a specific fixed destination. Therefore we strongly believe that social gossip could be implemented the same if all nodes would be mobile.

5.2 What about if nodes follow other mobility processes?

The choice of the forgetting function ϕ to be used in our model is set as a function of the targeted task to be performed. Actually, it also depends on the mobility model followed by the (mobile) nodes. Random walks are starting points of reference. Beside the fact that they are simple to deal with, they also seem to capture some aspects of the topological properties satisfied by the networks. For instance, observe that, for $\rho = 1$, and for any $\varepsilon > 0$, the function $\phi_{1,\varepsilon}$ does **not** depend on the dimension of the lattice, while still generating the desired distribution $1/d^k$. An

interesting direction of research would consist in considering other mobility models, such as biased random walks, or walk with random step size (*e.g.*, Levy walk).

One challenge is to determine which aspects of mobility enable simple forgetting scheme like ours to generate appropriate shortcut distribution. As an example, we have observed that a biased random walk may use the same function $\phi_{1,\epsilon}$ to obtain distribution with $\rho = 1$. However, this holds only in dimension 1, and, for larger dimension (*i.e.*, for $k \geq 2$), it is impossible to construct a function ϕ such that the forgetting process yields $\rho = 1$ when the random walk is biased.

Last, the mobility of different nodes may be correlated, especially as people in the same area may participate to a common event, or visit the same places at different times of the day. It is possible that this does not change the distribution of the shortcut length. Nevertheless, shortcuts would not be independent any more. We choose to ignore this issue here as we are not aware of any analytical results on the navigability of networks with correlated shortcuts.

6. REFERENCES

- [1] A. Chaintreau, P. Fraigniaud, and E. Lebar. Networks become navigable as nodes move and forget. *Proc. of ICALP*, 2008.
- [2] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, J. Scott, and R. Gass. Impact of human mobility on opportunistic forwarding algorithms. *IEEE Trans. Mob. Comp.*, 6(6):606–620, 2007.
- [3] A. Demers, D. Greene, C. Hauser, W. Irish, J. Larson, S. Shenker, H. Sturgis, D. Swinehart, and D. Terry. Epidemic algorithms for replicated database maintenance. *Proc. of ACM PODC*, 1987.
- [4] M. Grossglauser and D. Tse. Mobility increases the capacity of ad hoc wireless networks. *IEEE/ACM Trans. on Net.* 10(4):477–486, 2002.
- [5] S. Jain, K. Fall, and R. Patra. Routing in a delay tolerant network. *Proc. of ACM SIGCOMM*, 2004.
- [6] D. Kempe, J. Kleinberg, and A. Demers. Spatial gossip and resource location protocols. *Proc. of ACM STOC*, 2001.
- [7] T. Karagiannis, J.-Y. L. Boudec, and M. Vojnovic. Power law and exponential decay of intercontact times between mobile devices. *Proc. of ACM MobiCom*, 2007.
- [8] D. Kempe, J. Kleinberg, and A. Demers. Spatial gossip and resource location protocols. *J. ACM* 51(6):943–967, 2004.
- [9] D. Kempe and J. M. Kleinberg. Protocols and impossibility results for gossip-based communication mechanisms. *Proc. of IEEE FOCS*, 2002.
- [10] S. Keshav. Why cell phones will dominate the future internet. *SIGCOMM Comput. Commun. Rev.* 35(2):83–86, 2005.
- [11] J. Kleinberg. The small-world phenomenon: An algorithmic perspective. *Proc. of ACM STOC*, 2000.
- [12] J. Kleinberg. The wireless epidemic. *Nature* 449(7160):287–288, 2007.
- [13] A. Mtibaa, A. Chaintreau, J. Lebrun, E. Oliver, A.-K. Pietilainen and C. Diot. Are you moved by Your Social Network Application?. *Proc. of Workshop ACM SIGCOMM WOSN*, 2008.
- [14] C. Martel and V. Nguyen. Analyzing Kleinberg’s (and other) Small-World Models. *Proc. of ACM PODC*, 2004.
- [15] V. Nguyen and C. Martel. Analyzing and characterizing small-world graphs. *Proc. of ACM-SIAM SODA*, 2005.