Random Walks in Time-Graphs

Utku Günay Acer
INRIA
Sophia Antipolis, France
utku.acer@sophia.inria.fr

Petros Drineas
CS Dept, RPI
Troy, NY, USA
drinep@cs.rpi.edu

Alhussein A. Abouzeid
ECSE Dept, RPI
Troy, NY, USA
abouzeid@ecse.rpi.edu

ABSTRACT

Dynamic networks are characterized by topologies that vary with time and are represented by *time-graphs*. The notion of connectivity in time-graphs is fundamentally different than that in static graphs. End-to-end connectivity is achieved opportunistically by store-forward-carry paradigm if the network is so sparse that source-destination pairs are usually not connected by complete paths. In static graphs, it is well known that the network connectivity is tied to the spectral gap of the underlying adjacency matrix of the topology: if the gap is large, the network is well connected and a random walk on this graph has a small hitting time. In this paper, we investigate a similar metric for time-graphs, which indicates how quickly opportunistic methods deliver packets to destinations, speed of convergence in estimating an entity and quickness in the online optimization of protocol parameters, etc. To this end, a time-graph is represented by a 3-mode reachability tensor which yields whether a vertex is reachable from another node within t steps. Our observations from an extensive set of simulations show that the correlation between the expected hitting time of a random walk in the time-graph (following a non-homogenous Markov Chain) and the second singular value of the matrix obtained by unfolding the reachability tensor is significantly large, above 90%.

Categories and Subject Descriptors

C.2.1 [**Computer-Communication Networks**]: Network Architecture and Design—*Wireless Networks*

General Terms

Performance

Keywords

Time-Graphs, Dynamic Networks, Tensors, Structural Properties

1. INTRODUCTION

In wireless mobile networks, end-to-end connectivity is achieved collectively without the need for an established infrastructure using self-configuring applications and protocols (i.e. routing). Because of node mobility and other forms of dynamism in the network topology, the information these protocols use change frequently. Rather than fetching more recent information at the cost of overhead, the protocols may employ opportunistic methods to cope with dynamism [1]. In addition, the density of the network may be low so that source-destination pairs are not connected by complete paths most of the time. In such intermittently connected networks, end-to-end connectivity is achieved *over time* by utilizing the *store-forward-carry* paradigm.

It is useful for many applications to characterize how well the network is connected. For example, in well connected networks epidemic algorithms quickly spread the messages to the network and the minimum and/or maximum time needed to spread information to the whole network is small. Mechanisms that are used to estimate or optimize a parameter quickly converges if the network is well connected and the information flow is fast. Similarly, a random walk based mechanism, in which a random walker moves to neighboring nodes with equal probabilities, quickly terminates with success if the network is well connected. In intermittently connected networks, even though there are no complete paths between source-destination pairs at any given time instants, the messages are delivered relatively quickly to the destinations if the network is well-mixed.

In static networks, the connectivity of the network can be quantified by the *spectral gap* of the graph that represents the network, $\lambda_1 - \lambda_2$, where λ_1 and λ_2 are the two largest eigenvalues of the adjacency matrix of the graph [13]. If the spectral gap of the graph is large, the *mixing time*, the number of steps that a random walk must make on the graph before reaching a distribution that is acceptably close to the stationary distribution is small. A large spectral graph implies that *hitting time*, the number of steps a random walk makes before visiting a particular subset of the graph vertices for the first time, and the *cover time*, the number of steps a random walk must make to visit each vertex in the graph, are small.

In dynamic networks, the network topology constantly evolves typically in a non-deterministic manner, by inserting or deleting edges and/or nodes over time and the notion of connectivity is different from static networks. Consider the snapshots taken from a mobile network, that are depicted in Figure 1. In this example, nodes A and F are never connected. However, node A sends a packet to C at t_1. At this time C has no neighbor that it can forward this packet for delivery to F. C keeps the packet until t_2 in its buffers and send it to E at this time. At t_3, now E transmits the packet to node F. This delivery method exploits the dynamism in the network for end-to-end connectivity. This is called store-carry-forward paradigm and widely used in routing protocols for networks with intermittent connectivity [22]. Even though com-

(a) Network Topology at t_1

(b) Network Topology at t_2

(c) Network Topology at t_3
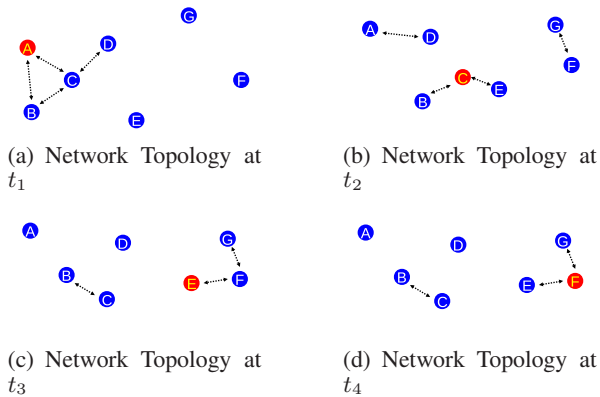
(d) Network Topology at t_4

Figure 1: Evolution of a network over time. A **and** F **are never connected. Still, end-to-end connectivity can be maintained between these nodes over time.**

plete paths between a majority of node pairs time can do not exist in this scenario, the links between the nodes might be formed so that one node quickly reaches another over time.

In this paper, we propose a structural metric that is similar to the spectral gap of the static networks to characterize the connectivity of dynamic networks. Other metrics can characterize the properties such as expected cluster numbers, average connected cluster size in a way (similar to the property in static graphs that when the spectral gap is 0, the network has more than one connected clusters). These metrics could also be useful in determining the optimum data dissemination protocols, quantifying the network vulnerability, etc.

A dynamic, variable topology network is represented by a *time-graph*, which indicates the creation and deletion of the vertices and/or edges in time. In particular, we use 3-mode adjacency tensors or 3-dimensional arrays and relate the structure of the time-graph to the network connectivity. The tensor yields whether two nodes are connected at a given time, similar to the adjacency matrix of a static graph, i.e. the entry $\mathcal{A}_{ijk} = 1$ in the adjacency tensor if vertex i is connected to vertex j at time k.

As a general rule, if the network is well connected, an opportunistic method such random walk performs better. This quality stems from the fact that random walk is able to sample nodes in a network with respect to a (typically uniform) probability distribution in a small number of steps in well connected networks. Hence, the performance of the random walk indicates how well the network is connected. In dynamic networks, the forwarding probabilities are derived from the adjacency tensor and continuously change in time. Therefore, random walks follow non-homogenous Markov Chains. On the other hand, the structural elements of the network is deduced through a series of operations on the adjacency tensor. Using the information given by the adjacency tensor, we can obtain the reachability tensor, which reports if a node can be reachable from another node within t time steps. We normalize the rows of the matrices of this tensor, and unfold it around the "distinguished" mode or dimension [15], which in this case is the dimension that depicts the time. This operation yields a two dimensional matrix. We use the singular values of this matrix as the structural metrics of the time-graphs.

Our observations based on the data from extensive simulations show that the correlation between the second singular value of this matrix and the expected hitting time is very high, above 0.9, which is a very large value for correlation. Hence, the second singular value is a good indicator of network connectivity. Performing

these experiments, we used a variety of node density values and node speed values so that the data forwarding can be in the form of store-and-forward as well as store-carry-forward. Tensors have been drawing a lot of interest recently; however, researchers still have a very long way of investigating the algebraic properties of the tensors. Therefore, it is not possible to support these observations with theoretical proof yet. Still, our experiments show that the proposed singular value can be used to evaluate the connectivity of dynamic networks.

The rest of this paper is organized as follows. In the next section we review the related work. In Section 3, we introduce our time-graph model, explain how the expected hitting time on time-graphs is derived and present the notion of reachability tensor. In Section 4, we show that the hitting time is highly correlated to the structural properties of reachability tensor via data obtained from simulations with various mobility models. Section 5 concludes the paper.

2. RELATED WORK

Random walks and the related Markov Chain Monte Carlo method are predominant in many areas of Computer Science, Mathematics, Engineering, Physics, Biology, Economics, etc. Random walks have been proposed as key algorithmic ingredients in protocols for various aspects of network design and maintenance. Existing literature [3, 6, 14] reports that a task for which independent sampling would be a good algorithmic primitive, such as searching, will benefit from random walks.

Dynamic networks constantly evolve by inserting or deleting edges and/or vertices over time. The notion of time evolving graphs was introduced by Kumar *et al.* in [9] as a novel combinatorial object to depict dynamic networks. In this model, a time graph $G = (V, E)$ consists of a set V of nodes where each node v_i has an associated interval $D(v_i)$ on the time axis, called the duration of v_i, and a set E of edges. A node v_i is said to be alive at time t, if $t \in D(v_i)$. Each edge is a triplet (v_i, v_j, t) where v_i and v_j are nodes in V and t is a point in time. The interpretation is that each edge is created at a point in time at which two end-points are alive. In [5], the author describes a combinatorial reference model capturing characteristics of time varying networks. The proposed time-graph model gives rise to several different metrics that may serve as objective functions in routing strategies, such as "earliest time to reach one or all the destinations". Scherrer *et al.* propose methods to describe dynamic graphs in [18] using properties such as the number of links and average degree as a function of time. Other relatively few studies that investigate dynamic graphs include [2, 12, 16]. The works generally investigate properties such as existence of communities and community size as well as how these properties change with time. On the other hand, we are interested in the structure of the entire dynamic graph, which we relate to the end-to-end connectivity.

There has been significant progress in understanding the linear algebraic properties of multi-mode tensors. Many researchers have focused on tensor decompositions, which have been successfully applied in data analysis [8, 11, 15, 21]. However, we are not aware of any attempt to connect tensors with the random walks or the notion of network connectivity.

In a wireless network, the dynamism stems from the node mobility. Mobility of nodes can be exploited to deliver packets to destination nodes that are not immediately connected to source nodes. In their seminal paper, Grossglauser and Tse show that if the network topology changes over time, the mobility can increase wireless network capacity assuming delay can be traded-off, and unlimited storage is available [7]. This result has inspired the design of
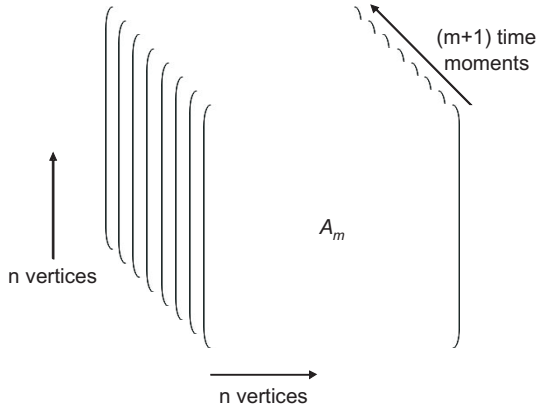
Figure 2: A tensor representation of time-graphs. A_i represents the adjacency matrix obtained from the network snapshot obtained at time t_i. A_i is the i-th slab of the adjacency tensor $\mathcal{A}$.

routing protocols for Delay Tolerant Networks where the connectivity is intermittent and it is not possible to form immediate paths between source-destination pairs. Instead, intermediate nodes have to store and carry the packets until they encounter the destination node or another node that is more likely to deliver the packet. Examples include [19, 20]. However, the question of how node mobility affects the evolving connectivity graph of the network remains unanswered.

3. METHODOLOGY

In this section, we introduce our time-graph combinatorial object. Then, we derive the expected hitting time for a random walk in an evolving time-graph. Finally, we present the notion of reachability tensor which we will use to denote the structure of the time-graph.

3.1 Our Time-Graph Model

In constructing the time-graphs, we discretize the continuous time intervals and focus on the edges instead of vertices. Let $t_0, t_1, \ldots, t_m$ denote discrete time moments. In this simple scenario, an edge between two nodes v_i and v_j might be present at t_0 and t_1 but disappear at t_2 and t_3, and reappear at t_4. We emphasize that the vertices on the graph are fixed. In our setting, deletion of a node v_i corresponds to making the node disconnected from the rest of the network, e.g. all edges adjacent to v_i disappear, and insertion of a node corresponds to adding an edge between v_i and other nodes that v_i contacts with. Clearly, at any time moment, any number of nodes might be disconnected or isolated.

We focus on undirected, unweighted time-graphs, which provide information of whether there is a bidirectional connection between two nodes at a time instant. Let V be the set of all the nodes of the time-graph, whose cardinality is n; $|V| = n$. Let $T = \{t_1, t_2, \ldots, t_m\}$ denote the set of all time moments of interest. At each time, $t \in T$, we take a snapshot of the dynamic network. Let $\mathcal{G}$ represent the time-graph and G_k, $k = 1, \ldots, m$, denote the snapshots of the dynamic graph obtained at time t_k. A_k denotes the adjacency matrix of G_k. Clearly, A_k is an $n \times n$ matrix. This representation implies a 3-dimensional array or a 3-mode $n \times n \times m$ tensor $\mathcal{A}$, which consists of all m matrices A_k. $\mathcal{A}_{ijk}$ is equal to 1 if there is an edge between nodes v_i and v_j at time t_k, otherwise it is zero. We call $\mathcal{A}$ the adjacency tensor of the time-

graph $\mathcal{G}$. The snapshot of the dynamic network at a specific time moment corresponds to a *slab* of the tensor [15]. Figure 2 depicts the tensor representation of time-graphs.

3.2 Expected Hitting Times of the Random Walks on Time-Graphs

Given the abstraction of time-graphs, we can visualize and formally define random walks on dynamic networks. Random walks in fixed graphs proceed in discrete steps: at time t_0 the walk takes a step from vertex $v(t_0)$ to a vertex $v(t_1)$ that is adjacent to $v(t_0)$ and is chosen uniformly at random. This node makes a similar decision at time t_1. The transitions of the random walk is modeled by a Markov Chain where the transition probabilities are the same at all times. In other words, the Markov Chain is homogeneous. Random walks in time-graphs are essentially the same, except for that even though $v(t_1)$ is adjacent to $v(t_0)$ at time t_0, they might not be adjacent to each other at time t_1 (when the next transition actually takes place). The transition probabilities also change with time. As a result, the Markov Chain becomes non-homogeneous. The state space of the Markov Chain is fixed and each state corresponds to a particular node in the network.

The connectivity of the time-graph can be evaluated by the expected hitting time of a random walk on that graph. The hitting time is the number of steps a random walk takes to reach a particular node for the first time. If the network is well connected, the expected hitting time is small. We now summarize the results given in [17] to derive the expected hitting time in a non-homogeneous Markov Chain.

For notational convenience, we only use j instead of v_j and k instead of t_k. Let X_k denote the state of the Markov Chain, the node at which the random walk resides at time k. Multiple step transition probabilities from time k to time $k + a$ are defined as

$$p_{k,k+a}(i,j) = P\{X_{k+a} = j | X_k = i\} \quad (1)$$

where $a \geq 1$. For presentation simplicity, let $p_k(i,j) = p_{k,k+1}(i,j)$. Single step transition probabilities are thus

$$p_k(i,j) = \begin{cases} \frac{1}{\zeta_k^i} & \text{if } i \text{ and } j \text{ are neighbors at time } k \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

where ζ_k^i is the number of neighbors node i has at time k. The entries $p_k(i,j)$ with $0 \leq i, j \leq n - 1$ constitute the probability transition matrix at time k, P_k. Note that P_k is obtained by normalizing the rows of A_k.

Transition probabilities have the following properties:

- $p_k(i,j) \geq 0$ for all $i, j \in V$ and $k \geq 0$.

- $\sum_{j \in V} p_k(i,j) = 1$ for all $i \in V$ and $k \geq 0$.

- $p_k(i,i) = 0$ if $\zeta_k^i > 0$ and 1 otherwise.

- $p_{h,k}(i,j) = \sum_{l \in V} p_{h,r}(i,l) p_{r,k}(l,j)$ for all r such that $h < r < k$

Let Θ denote the time at which a random walk first hits node d when the random walk starts at another node $s \in V \backslash \{d\}$. Without loss of generality, assume that the random walk is initiated at time $t = 0$. The hitting time is

$$\Theta = \inf\{k > 0; X_k = d\}. \quad (3)$$

To find the hitting time, a new non-homogeneous Markov Chain can be defined where d is now an absorbing state. State d always transitions to itself with probability 1 no matter how many neighbors node d has. All the other transition probabilities remain the

same. In both Markov Chains, the expected time a random walk first hits state d is the same. The probability transition matrix at time k for this Markov Chain is

$$Q_k = \begin{bmatrix} R_k & H_k \\ 0 & 1 \end{bmatrix} \qquad (4)$$

where the $(n-1) \times (n-1)$ matrix R_k represents the transition probabilities between the states in the subset $V \backslash \{d\}$ at time k (i.e. R_k is P_k with dth row and column removed) and H_k is the $n-1$ length vector denoting single step transition probabilities from $V \backslash \{d\}$ to d at time k.

Using the non-homogeneous Markov Chain whose transition probability matrices in time are given in (4), the tail distribution for the hitting time is

$$P(\Theta > \eta) = P(\forall i, i \leq \eta, X_i \in V \backslash \{d\})$$
$$= \alpha \left(\prod_{k=0}^{\eta-1} R_k \right) \mathbf{1}_{n-1} \qquad (5)$$

where $\mathbf{1}_{n-1}$ is a length $n-1$ column vector whose entries are all 1 and α is the $n-1$ entry row vector indicating the initial distribution of the random walk. Equation 5 indicates the probability that random walk still remains in $V \backslash \{d\}$ after η time steps.

Since Θ is a discrete random variable,

$$E[\Theta] = \sum_{\eta \geq 0} \eta P(\Theta = \eta) = \sum_{\eta \geq 0} P(\Theta > \eta) \qquad (6)$$

with $P(\Theta > 0) = 1$ since $s \in V \backslash \{d\}$. Hence, the expected hitting time can be rewritten as

$$E[\Theta] = \alpha \left[I + \sum_{\eta \geq 1} \prod_{k=0}^{\eta-1} R_k \right] \mathbf{1}_{n-1} \qquad (7)$$

where I is the $n-1 \times n-1$ identity matrix.

Consider two matrices Q_k and Q_{k+1} of type (4). Their product is

$$Q_k Q_{k+1} = \begin{bmatrix} R_k R_{k+1} & R_k H_{k+1} + H_k \\ 0 & 1 \end{bmatrix}.$$

Since all the terms of the matrices are positive,

$$[H_k](i) \leq [R_k H_{k+1}](i) + [H_k](i) \qquad (8)$$

$\forall i \neq d$, which indicates that probability of transition to d increases and the probability of random walk remaining at a state other than the destination decreases at every time step. A more detailed discussion can be found in [17].

3.3 Reachability Tensor

Using the initial time-graph model, we define another tensor, reachability tensor, $\mathcal{B}$. This tensor indicates whether a node is reachable from another node after a certain number of steps. Let T_k be defined as

$$T_k = \min \left(\mathbf{1}, \prod_{i=0}^{k} A_i \right) \qquad (9)$$

where $\mathbf{1}$ is an all 1 $n \times n$ matrix. The product $\prod_{i=0}^{k} A_i$ is $n \times n$ matrix. The entry that corresponds to the i-th row and the j-th column of this matrix gives the number of k-step paths that start at i and end at j. If this value is above 0, a random walk starting from node i at time 0 can end up at node j at time k and there is at least one k-step path connecting i to j. The k-th slap of the reachability tensor $\mathcal{B}$ is defined as follows:

$$B_k = \min \left(\mathbf{1}, \sum_{i=0}^{k} T_i \right). \qquad (10)$$

When $B_k(i,j) = 1$, node j is reachable from node i in at most k steps, i.e. j is reachable from i after k time steps. The number of slabs in $\mathcal{B}$ is defined as

$$\tau = \inf(k; B_k = \mathbf{1}) \qquad (11)$$

if there exists such $k \leq m$. In this case, every vertex in the graph is reachable from every other one in τ time steps. Otherwise, $\tau = m$. τ is named the *time-diameter* of the time-graph. Note that if this procedure is applied on a fixed graph, the time-diameter becomes the diameter of the graph, which is the maximum distance between two nodes in terms of hop count.

$\mathcal{B}$ is a $n \times n \times \tau$ tensor, i.e. $\mathcal{B} \in \Re^{n \times n \times \tau}$. Define the matrix $S \in \Re^{\tau, n \times n}$, where the columns of the matrix consist of varying the 3rd mode of $\mathcal{B}$, i.e. the mode or the dimension that describes the time. Note that this mode the *distinguished* dimension, i.e. it is qualitatively different than the other modes, which both model the vertices in the network. Each column of this matrix consists of a slab of the tensor. We refer to the construction S as *matricizing* or *unfolding* $\mathcal{B}$ along mode 3. Before matricizing $\mathcal{B}$, all the rows in each of its slabs are normalized. In the next section, we experimentally show that there is a high correlation between the first and second singular values of S, ρ_1 and ρ_2 respectively and the expected hitting time in the dynamic network.

4. EVALUATION

In this section, we perform the expected hitting time analysis on data gathered by simulations with a variety of mobility models and extract the structural properties of reachability tensor.

4.1 Simulation Setup: Gathering Data

Our observations are based on data generated by a custom simulator. Each node moves independently according to a common mobility model. Unless otherwise noted, we used 50 nodes in the simulations. The nodes have 250m communication range and two nodes can communicate directly if they are in the communication range of each other. The nodes move in a $X \times X$ m^2 region. In order to capture the different levels of population densities, X varies from 1000 m to 3000 m in the simulations. This way, we obtain node densities that are very low so that every node has at most one neighbor most of the time as well as high densities where nodes almost always have multiple neighbors. Regardless, the network topology continuously changes. We use a large range of speed values to model different levels of dynamism. The snapshots are taken at every 0.05 seconds. The simulation time is 1000 seconds. For each scenario, we have performed 5 runs. The results in the graphs are the averaged values. The correlation values on the other hand are calculated using the raw data as the correlation between the averaged values are much higher.

Each run starts with random node displacement and initial warm-up duration to reach stationary node distribution of the mobility model. We calculate the expected hitting time for each node with random walk equally likely to start every node other than the destination to obtain the expected hitting time for the particular time-graph. All results are averaged over 5 different instances. For evaluation, we used the following mobility models[1]:

- Random Walk Mobility Model: In this mobility model, mobile nodes have fixed journey durations, t. At the beginning

---

[1]A more detailed discussion of these models can be found in [4].

(a) Dynamism   (b) Expected Hitting Time   (c) First Singular Value of S (ρ_1)

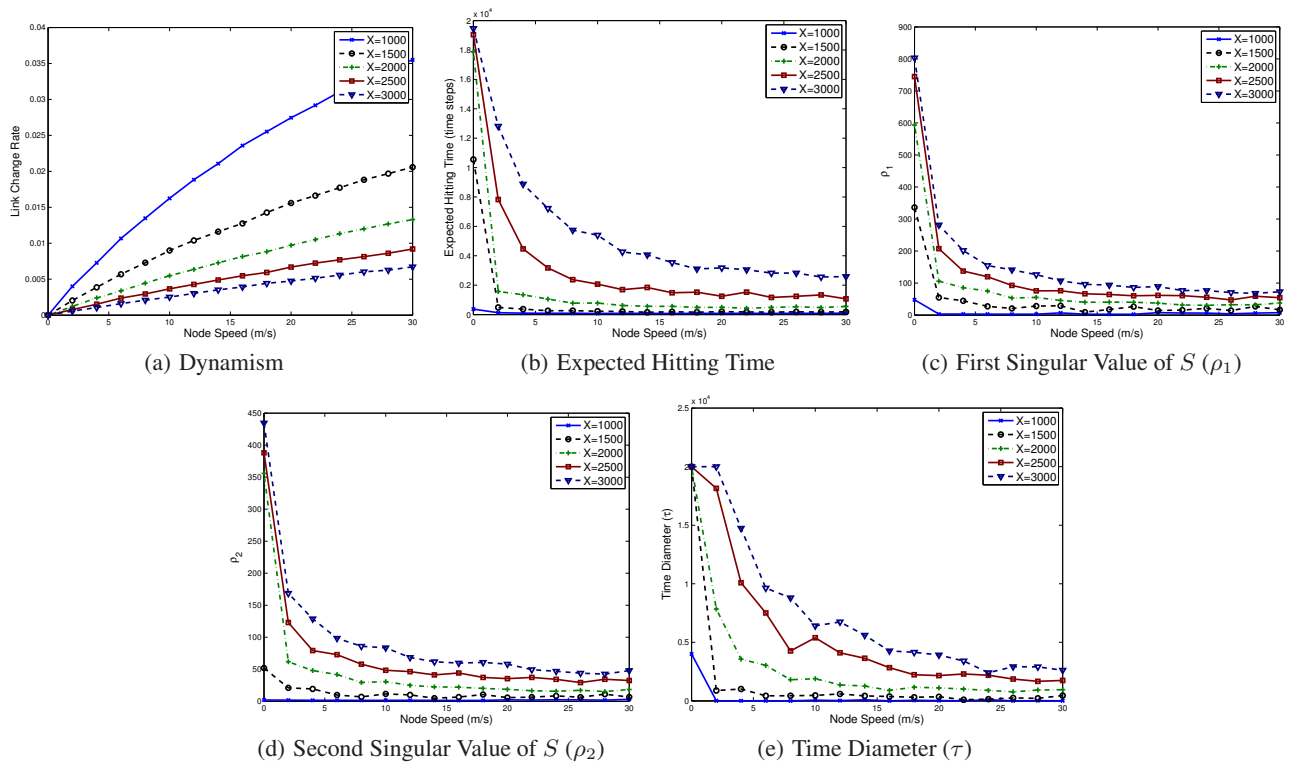(d) Second Singular Value of S (ρ_2)   (e) Time Diameter (τ)

Figure 3: The results for Random Waypoint Mobility Model. The curves that represent the expected hitting time and the first two singular values of S are very similar. Experiments show that the correlation between the expected hitting time and the each of the singular values is very high, above 0.9.

of each journey, the node randomly selects a speed value and a direction value. The speed is a uniform random variable within the interval $[v_{min}, v_{max}]$. Similarly, the direction is also uniformly distributed within the interval $[0, 2\pi]$. If a node reaches a simulation boundary, it bounces off the boundary with an angle that depends on its original direction. In the simulations, $t = 10$ seconds. In each run, the node speed varies within the interval $[v_{min}, v_{max}]$. We used $v_{max} = 2v_{min}$ and varied v_{min} from 0 m/s to 30 m/s.

- Random Waypoint Mobility Model: In this model, the nodes pick destination points and speed values at the beginning of their journeys and move towards their destinations with the selected speed. The duration of a journey is as long as it takes the node to arrive at the destination point. Once the node arrives at its destination, it can pause for some duration of time before selecting a new destination point and a new speed value. The speed values in this model are the same as the ones used in the Random Walk Mobility Model.

- Boundless Simulation Area Mobility Model: Each node contains a speed value v and a direction θ correlated in time. Each node updates its speed and direction every Δt seconds according to

$$v(t + \Delta t) = \min\left(\max\left(v(t) + \Delta v, v_{min}\right), v_{max}\right)$$
$$\theta(t + \Delta t) = \theta(t) + \Delta\theta$$

where Δv is uniformly distributed between $[-A_{max}\Delta t, A_{max}\Delta t]$ and $\Delta\theta$ is uniformly distributed in $[-\alpha, \alpha]$, where A_{max} is the maximum acceleration and α is the maximum change in

the direction of node mobility. In this simulation model, the simulation area is a 2-D torus instead of a rectangular area. If a node reaches a boundary, they continue their movement and reappear on the other side of the area. Contrary to previous models, the node mobility is not memoryless in this mobility model. In the simulations we used, $\Delta t = 0.2$ seconds, $A_{max} = v_{max}\Delta t$ and $\alpha = \pi/10$. v_{min} and v_{max} vary in the same way as they do in the previous models.

- Gauss-Markov Mobility Model: This model is similar to the previous model in how the nodes update their speed and direction values. Differently in this model, the nodes move in a rectangular area and their movement reflects off the boundary they encounter. The speed and direction values are updated according to

$$v(t + \Delta t) = \alpha v(t) + (1 - \alpha)\bar{v} + \omega_v\sqrt{(1 - \alpha^2)}$$
$$\theta(t + \Delta t) = \alpha\theta(t) + (1 - \alpha)\bar{\theta} + \omega_\theta\sqrt{(1 - \alpha^2)}$$

where α is the correlation coefficient, $\bar{v}$ and $\bar{\theta}$ mean speed and direction as $t \to \infty$ and ω_v and ω_θ are random variables from a Gaussian distribution with mean 0 and variance 1. In our simulations, we use $\alpha = 0.75$ and $\Delta t = 0.2$ seconds. We randomly initiate $\bar{\theta}$ for each node. When a node reaches a boundary, its movement is reflected. In this model, we vary $\bar{v}$ from 0 to 60 m/s.

We quantify the dynamism in the network using *normalized average link change rate* metric LCR [10], which is calculated by

$$LCR = \frac{1}{T_S} \frac{\sum_k (E_k^a + E_k^d)}{n(n-1)/2} \qquad (12)$$

where T_S is the simulation time, E_k^a is the number of edges added to the network and E_k^d is the number of edges deleted from the at the k-th time step. The total number of link changes is divided by the maximum possible number of edges in the network, $n(n-1)/2$, with n being the number of the nodes for normalization and the simulation time T_S for normalization.

We run our simulations and evaluate the data resulting from the simulations on machines with two dual-core sequence processors at 3.0 GHz and 4GB random access memory. A simulation run in which nodes move according to the random walk mobility model with $v_{min} = 5$ m/s and $v_{max} = 10$ m/s in a 3000×3000 m^2 area and the evaluation of the data resulting from this run takes about 95 minutes.

4.2   Results and Discussion

Figure 3 shows the results obtained from simulations in which nodes move according to the Random Waypoint Mobility Model. As expected, the links change with a larger rate as the node speed increases as shown in Figure 3(a). However, the nodes moving with speed values selected from the same range cause different levels of dynamism as the size of the area covered by the network changes. When the area is small, for each node, the number of nodes which are located at a distance close to the communication range is more. The movement of these nodes cause increasing number of link additions and link deletions. Figure 3(b) shows how the expected hitting time changes with the node speed in different values for the size of area. In a sparse network where source-destination pairs are likely disconnected, the node that carries the random walk is more likely to encounter with a connected portion of the network that contains the destination node at high node mobility. As a result, the expression $[R_k H_{k+1}](i)$ in (8) has higher values for all k and for each entry i. Remember that the sum of this expression over i yields the transition probability to the destination node in two steps. This in turn decreases the product $[R_k R_{k+1} \mathbf{1}_{n-1}](i)$, which is the probability that the random walk remains at a state i that is different from the destination after two steps. Consequently, the hitting time decreases and paths between disconnected nodes can be formed more quickly. This also leads to decreasing time-diameters with increasing node mobility as shown in Figure 3(e). However, more dynamism does not necessarily mean better connectivity. In a dense network, the destination nodes are usually connected with the rest of the network. Even though the links change at a very large rate, the number of neighbors that are adjacent to each node does not change dramatically over time with increasing mobility. Hence, the expected hitting time does not change much with the mobility.

Figure 3(c) shows how the first singular value of the matricized reachability tensor, ρ_1 changes. Figure 3(d) depicts the characteristics of the second singular value of the matricized reachability tensor, ρ_2. Note that these curves have very similar characteristics to the expected hitting time shown in Figure 3(b). Both ρ_1 and ρ_2 remain steady in dense networks whereas they decrease with node mobility in sparse networks. The experiments show that the correlation between ρ_1 and the expected hitting time is slightly above 0.93. When we consider ρ_2, the correlation is very close to 0.95 which is very high. Note that even though the figures show the averaged values, the correlation values are calculated using the raw data from each single run, not the averaged results. The correlation between the averaged values is even higher. In the rest of this section, we will focus on the expected hitting time and ρ_2.

In Figure 4, we show that there is similar relationship between expected hitting time and ρ_2 when the nodes move according to the Random Walk Mobility Model. The expected hitting time de-
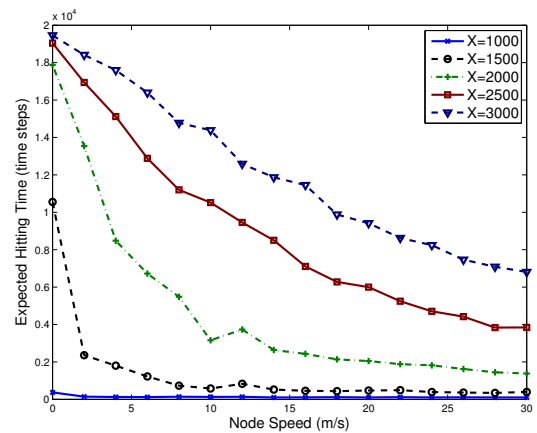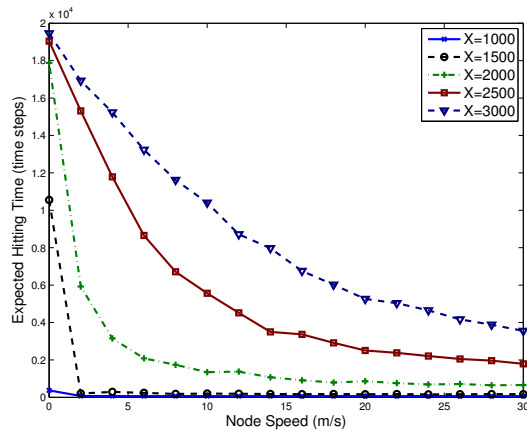

(a) Expected Hitting Time
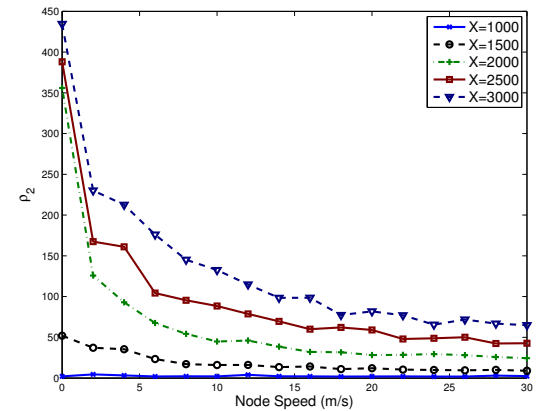
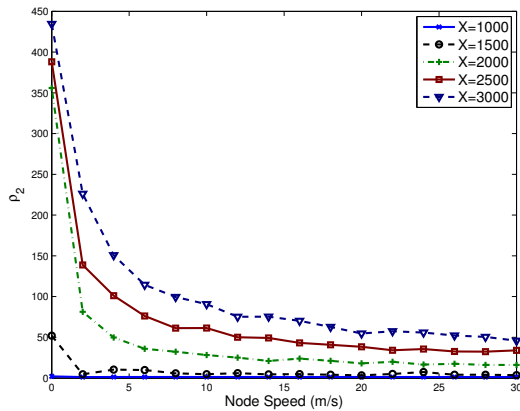(b) Second Singular Value of S (ρ_2)

Figure 4: The results for Random Walk Mobility Model. The expected hitting time and the second singular value of the unfolded reachability tensor have similar characteristics. Our experiments show that the correlation between two value is beyond 0.95 for this mobility model.

creases with node mobility in sparse networks, but does not change much in dense networks due to the same reasons explained for the Random Waypoint Mobility Model as depicted in Figure 4(a). ρ_2 also has the same characteristics as shown in Figure 4(b). The correlation between ρ_2 and the expected hitting time is around 0.96. The results on the expected hitting time and ρ_2 when the nodes move according to Boundless Simulation Area Mobility Model are given in Figure 5(a) and Figure 5(b), respectively. The characteristics of these entities are consistent with their counterparts in the previous mobility models. In this particular case, the correlation between the expected hitting time and ρ_2 is above 0.96. As depicted in Figure 6, the same arguments hold for the results with the Gauss-Markov Mobility Model. The expected hitting time and the second singular value of the matricized reachability tensor decrease with the dynamism in the network as Figures 6(a) and 6(b) show, respectively. In this mobility model, the correlation between ρ_2 and the expected hitting time is over 0.97.

Additionally, we have performed simulations to see how the expected hitting time and ρ_2 values vary with the number of the nodes in the network and if our observation about the correlation still holds. In this case, we used the random waypoint mobility in the node movements in various network area size with fixed speed val-

(a) Expected Hitting Time

(a) Expected Hitting Time

(b) Second Singular Value of S (ρ_2)

(b) Second Singular Value of S (ρ_2)

Figure 5: The results for Boundless-area Mobility Model. The expected hitting time and ρ_2 are again very similar. The correlation in this case is again more than 0.95.
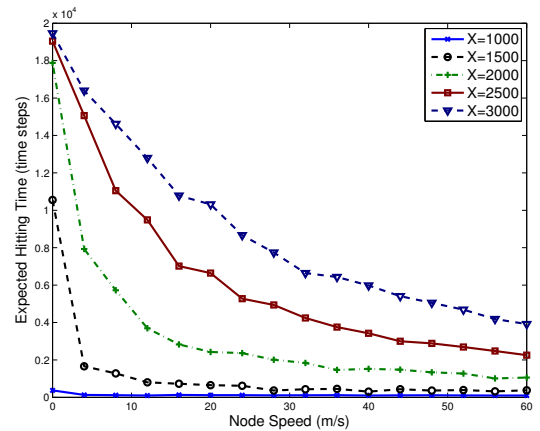
Figure 6: The results for Gauss-Markov Mobility Model. The results are consistent with the other mobility models. The correlation between the expected hitting time and ρ_2 is very large, approximately 0.97.

ues $v_{min} = 10$ m/s and $v_{max} = 20$ m/s. Fig.7 shows the characteristics of the expected hitting time and ρ_2 with respect to the number of nodes in the network. As in prior cases, the similarities between the characteristics of these entities are eminent. The correlation between these two sets of data is above 93%.

Up until now, we have compared the expected hitting time and ρ_2 values separately. When values obtained from all the simulation runs(more than 2000 runs) are considered altogether, the correlation coefficient is above 0.9.
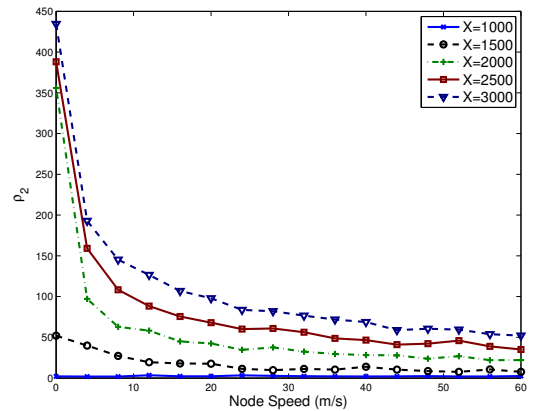
In our evaluation, we assume that the transition of the random walk between the nodes or the transition between states in the Markov Chain takes place at once, without any delay. However, the time it takes a node to transmit a packet to a neighbor can be large especially if the packet size is large. In order to capture the large packet size effect, the snapshots of the network can be taken at a lower rate, or with large intervals between two consecutive snapshots. Still, the correlation between the expected hitting time and ρ_2 remains very large.

5. CONCLUSION

In this paper, we investigate the relationship between the dynamism in the network and the network connectivity in mobile networks. To represent the dynamic networks, a novel combina-

torial time-graph model is proposed. Instead of a adjacency matrix, the time-graph is modeled by a 3-mode adjacency tensor, or 3-dimensional array $\mathcal{A}$. In this model, slabs of the tensor correspond to the adjacency matrices of the snapshots of the network at discrete time instants. In a time-graph, if vertice v_i is connected to vertice v_j at time t_k, the entry $\mathcal{A}_{ijk}$ in the tensor is 1. This tensor representation leads us to the expected hitting time of a random walk that proceeds in the time-graph using non-homogeneous Markov Chains and another tensor which is referred to as reachability tensor $\mathcal{B}$. The entry $\mathcal{B}_{ijk}$ equals 1 if a random walk starting from node v_i can end up in node v_j at time k; otherwise it equals 0. The rows of each slab of this tensor is normalized and then matricized or unfolded along mode-3 (time axis), which is the distinguished dimension of the time-graph in which the other two modes represent the vertices. Our observations based on an extensive set of experiments indicate that there is a significantly large correlation between the second singular value of the matricized reachability tensor of a time-graph and the expected hitting time for random walks on that time-graph, above 0.9. This a very large correlation value; therefore, it is a good indicator of end-to-end connectivity for dynamic networks.

This study provides a connection between the structure and the properties of time-graphs and the network connectivity in a dy-
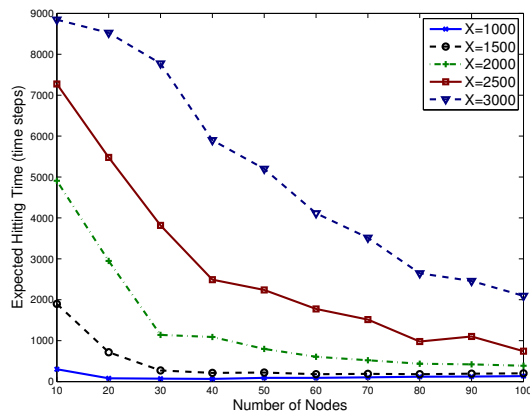
(a) Expected Hitting Time
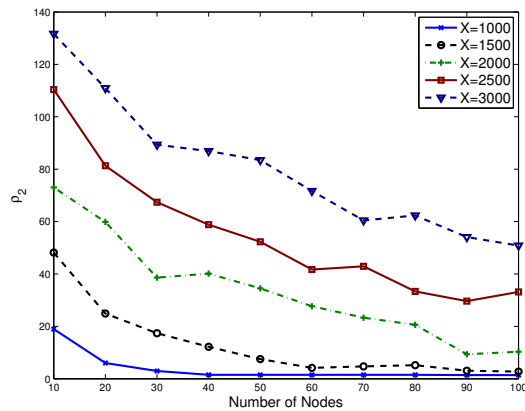
(b) Second Singular Value of S (ρ_2)

Figure 7: Expected hitting time and ρ_2 with respect to the number of nodes in the network. The prior observations are true for this case as well. The correlation between the expected hitting time and ρ_2 is 0.93.

namic network, which presents fundamental differences from the connectivity in fixed networks. To the best of our knowledge, there is no prior study that makes a similar attempt. In addition, adjacency and reachability tensors can potentially yield information that characterizes a wide range of dynamic network properties such as average number of clusters, expected cluster size, etc. In future work, we will investigate such other structural properties. We also plan to develop distributed mechanisms to calculate/estimate these metrics so that they could be used by network protocols and applications for mobile ad-hoc and delay tolerant networks. In addition, our plans involve developing the mathematical foundations behind the observations.

Acknowledgement

6. REFERENCES

[1] ACER, U. G., KALYANARAMAN, S., AND ABOUZEID, A. A. Weak State Routing for Large Scale Dynamic Networks. *MobiCom '07: Proceedings of the 13th Annual ACM International Conference on Mobile Computing and Networking* (2007), 290–301.

[2] BEN JDIDIA, M., ROBARDET, C., AND FLEURY, E. Communities detection and analysis of their dynamics in collaborative networks. *2nd International Conference on Digital Information Management, ICDIM 2* (2007), 744 – 749.

[3] BISNIK, N., AND ABOUZEID, A. A. Optimizing random walk search algorithms in p2p networks. *Computer Networks 51*, 6 (2007), 1499 – 1514.

[4] CAMP, T., BOLENG, J., AND DAVIES, V. A survey of mobility models for ad hoc network research. *Wireless Communications and Mobile Computing 2*, 5 (2002), 483 – 502.

[5] FERREIRA, A. Building a reference combinatorial model for manets. *IEEE Network 18*, 5 (2004), 24 – 29.

[6] GKANTSIDIS, C., MIHAIL, M., AND SABERI, A. Random walks in peer-to-peer networks. *Proceedings - IEEE INFOCOM 1* (2004), 120 – 130.

[7] GROSSGLAUSER, M., AND TSE, D. N. Mobility increases the capacity of ad hoc wireless networks. *IEEE/ACM Transactions on Networking 10*, 4 (2002), 477 – 486.

[8] KOLDA, T. G. Orthogonal tensor decompositions. *SIAM Journal on Matrix Analysis and Applications 23* (2001), 243–255.

[9] KUMAR, R., NOVAK, J., RAGHAVAN, P., AND TOMKINS, A. On the bursty evolution of blogspace. *World Wide Web 8*, 2 (2005), 159 – 78.

[10] KWAK, B.-J., SONG, N.-O., AND MILLER, L. E. A standard measure of mobility for evaluating mobile ad hoc network performance. *IEICE Transactions on Communications E86-B*, 11 (2003), 3236 – 3243.

[11] LEIBOVICI, D., AND SABATIER, R. A singular value decomposition of a k-way array for a principal component analysis of multiway data, pta-k. *Linear Algebra and Its Applications 269*, 1-3 (1998), 307 – 329.

[12] LESKOVEC, J., KLEINBERG, J., AND FALOUTSOS, C. Graphs over time: Densification laws, shrinking diameters and possible explanations. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2005), 177 – 187.

[13] LINIAL, N., AND WIGDERSON, A. Expander graphs and their applications. Lecture Notes, The Hebrew University, Israel, 2003.

[14] LV, Q., CAO, P., COHEN, E., LI, K., AND SHENKER, S. Search and replication in unstructured peer-to-peer networks. *Proceedings of the International Conference on Supercomputing* (2002), 84 – 95.

[15] MAHONEY, M. W., MAGGIONI, M., AND DRINEAS, P. Tensor-cur decompositions for tensor-based data. In *KDD '06: Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining* (2006), pp. 327–336.

[16] PALLA, G., BARABASI, A., AND VICSEK, T. Quantifying social group evolution. *Nature 446* (April 2007), 664–667.

[17] PLATIS, A., LIMNIOS, N., AND LE DU, M. Hitting time in a finite non-homogeneous markov chain with applications. *Applied Stochastic Models and Data Analysis 14*, 3 (1998), 241 – 253.

[18] SCHERRER, A., BORGNAT, P., FLEURY, E., GUILLAUME, J.-L., AND ROBARDET, C. Description and simulation of dynamic mobility networks. *Computer Networks 52*, 15 (2008), 2842 – 2858.

[19] SPYROPOULOS, T., PSOUNIS, K., AND RAGHAVENDRA, C. S. Spray and wait: An efficient routing scheme for intermittently connected mobile networks. *Proceedings of ACM SIGCOMM 2005 Workshops: Conference on Computer Communications* (2005), 252 – 259.

[20] VAHDAT, A., AND BECKER, D. Epidemic Routing for Partially Connected Ad Hoc Networks. Tech. rep., Department of Computer Science, Duke University, Durham, NC, 2000.

[21] ZHANG, T., AND GOLUB, G. H. Rank-one approximation to high order tensors. *SIAM J. Matrix Anal. Appl. 23*, 2 (2001), 534–550.

[22] ZHANG, Z. Routing in intermittently connected mobile ad hoc networks and delay tolerant networks: overview and challenges. *Communications Surveys & Tutorials, IEEE 8*, 1 (2006), 24–37.