Pervasive and Mobile Computing 7 (2011) 160–171

Contents lists available at ScienceDirect

Pervasive and Mobile Computing

journal homepage: www.elsevier.com/locate/pmc

Fast track article

Connectivity in time-graphs☆

Utku Günay Acer [a,]*, Petros Drineas [b], Alhussein A. Abouzeid [c]

[a] Alcatel-Lucent Bell Laboratories, Copernicuslaan 50, 2018 Antwerp, Belgium

[b] Computer Science Department, Rensselaer Polytechnic Institute, 110 8th Street, Troy, NY, 1280, USA

[c] Electrical, Computer & Systems Engineering Department, Rensselaer Polytechnic Institute, 110 8th Street, Troy, NY, 1280, USA

ARTICLE INFO

Article history:
Available online 15 December 2010

Keywords:
Time-graphs
Dynamic networks
Connectivity
Tensors

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 from that in static graphs. End-to-end connectivity is achieved opportunistically by the store–carry-forward 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. In this paper, a similar metric is investigated for time-graphs. To this end, a time-graph is represented by a 3-mode reachability tensor which indicates whether a node is reachable from another node in t steps. To evaluate connectivity, we consider the expected hitting time of a random walk, and the time it takes for epidemic routing to infect all vertices. Observations from an extensive set of simulations show that the correlation between the second singular value of the matrix obtained by unfolding the reachability tensor and these indicators is very significant.
© 2010 Elsevier B.V. All rights reserved.

1. Introduction

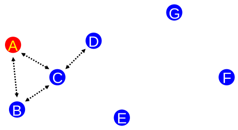
In wireless mobile networks, end-to-end connectivity is achieved collectively without the need for an established infrastructure using self-configuring applications and protocols (i.e. routing). Because of node mobility and other forms of dynamism in the network topology, the information these protocols use changes frequently. Rather than fetching more recent information at the cost of higher overhead, the protocols may employ opportunistic methods to cope with dynamism [1]. In addition, the density of the network may be low such 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–carry-forward* 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 converge quickly 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 may be no complete paths between source–destination pairs at any given time instant, the messages are delivered relatively quickly to the destinations if the network is well mixed.

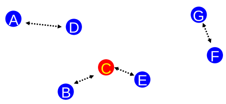
☆ An earlier version of this paper appeared in MobiOpp'10 (Acer et al. [26]).

* Corresponding author.
E-mail addresses: utku_gunay.acer@alcatel-lucent.com, utkuacer@ieee.org (U.G. Acer), drinep@cs.rpi.edu (P. Drineas), abouzeid@ecse.rpi.edu (A.A. Abouzeid).

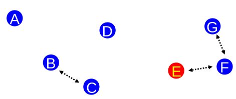
1574-1192/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.pmcj.2010.11.011



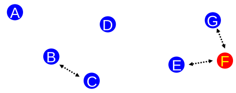
(a) Network topology at t_1.



(b) Network topology at t_2.



(c) Network topology at t_3.



(d) Network topology at t_4.

Fig. 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.

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 Fig. 1. In this example, nodes A and F are never connected at any time instant. However, node A can send 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 buffer and sends it to E at this time. At t_3, 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 is widely used in routing protocols for networks with intermittent connectivity [2]. Even though complete paths between a majority of node pairs at any given instant do not exist in this case, the links between the nodes might be formed so that a message originating at one node is delivered to another node over time.

A dynamic, variable topology network is represented by a *time-graph*, which indicates the creation and deletion of the vertices and/or edges over time. In particular, we use 3-mode adjacency tensors or three-dimensional arrays to model time-graphs and relate their structure to the network connectivity. The tensor indicates whether two nodes are connected by a common link 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.

To provide some intuition behind our work, we start with the motivating scenario of d-regular, undirected graphs with n vertices and nd edges (recall that a d-regular graph has d edges associated with each vertex). It is well known that the associated graph adjacency matrix has the first eigenvalue exactly equal to d. If the difference between this first eigenvalue and the second largest one is sufficiently large, then the graph is an expander. Intuitively, this implies that the number of edges that must be removed from the graph in order to make some large subset of its vertices disconnected from the remainder of the graph is also large. Thus, expanders can represent robust networks with good connectivity properties, while at the same time having only a small number of edges (e.g., linear in n). Our work explores whether a similar property holds for evolving graphs, by studying properties of such graphs and their connections to tensor (as opposed to matrix) eigenvalues.

As a general rule, if the network is well connected, an opportunistic method such as random walk performs better. This quality stems from the fact that a 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 well connected graphs, the *hitting time*, i.e. the number of steps a random walk takes before visiting a particular subset of the graph vertices for the first time. In dynamic networks, the forwarding probabilities are derived from the adjacency tensor and continuously change over time. Therefore, a random walk follows a non-homogeneous Markov Chain. A similar opportunistic method is epidemic routing [3], where nodes hold on to the messages they receive and replicate them in the nodes they subsequently meet. In a well connected network, all nodes quickly obtain a replica of the original message. On the other hand, the structural elements of the network are 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 indicates whether a random walk starting from one node can reach another node after t time steps. We normalize the rows of the matrices of this tensor, and unfold the tensor around the "distinguished" mode or dimension [4], which in this case is the dimension that depicts 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 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 correlation. Similarly, the average time until the epidemic routing delivers a message that originates at an arbitrary node to the entire network is highly correlated with this structural metric, with a correlation coefficient more than 87%. Hence, the second singular value of the unfolded reachability tensor is a good indicator of network connectivity. Performing these experiments, we used a variety of node densities and speeds. This way, we are able to evaluate this structural metric in scenarios where the nodes are always instantaneously connected via complete paths or they are intermittently connected over time. Our experiments show that the proposed singular value can be used to evaluate the connectivity of dynamic networks. Even though tensors have been drawing a lot of interest recently, researchers still have a long way to go towards understanding the algebraic properties of the tensors. Therefore, it is not possible to support these observations with theoretical proof yet.

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 the reachability tensor via data obtained from simulations with various mobility models. Section 5 concludes the paper and discusses the open problems.

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 and Economics. Random walks have been proposed as key algorithmic ingredients in protocols for various aspects of network design and maintenance. The existing literature [5–7] reports that any task for which independent sampling would be a good algorithmic primitive, such as searching, typically benefits from random walks.

Epidemic methods can be used to quickly distribute some content (e.g. packets) to the entire network. Originally, epidemic algorithms were proposed to maintain consistency for replicated databases. In this approach, an update in a database is distributed to all the replicas [8]. In the context of routing in intermittently connected networks, this method has been proposed to deliver packets quickly to the desired destination [3]. In this method, a node always maintains a received packet in its buffers. When this node creates a link with another node, it copies all the messages in its buffer to its new neighbor. Thus, packets quickly spread to all the nodes if the network is well connected.

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 represent 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 when both of its end-vertices are alive. In [10], 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 [11] 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 [12–14]. Aforementioned line of research generally investigates 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 [4,15–17]. However, we are not aware of any attempt to connect tensors with 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 [18]. This result has inspired the design of 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 [3,19,20]. 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 reflect the structure of the time-graph.

3.1. Our time-graph model

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 instances. In this simple scenario, an edge between two nodes v_i and v_j is present at t_0 and t_1 but disappears at t_2 and t_3, and reappears 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 is connected to. Clearly, at any time instant, any number of nodes might be disconnected or isolated.

We focus on undirected, unweighted time-graphs, which provide information on whether there is a bidirectional connection between two nodes at a time instant. Let V denote 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 instances 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 three-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 instance corresponds to a *slab* of the tensor [4]. Fig. 2 depicts the tensor representation of time-graphs.

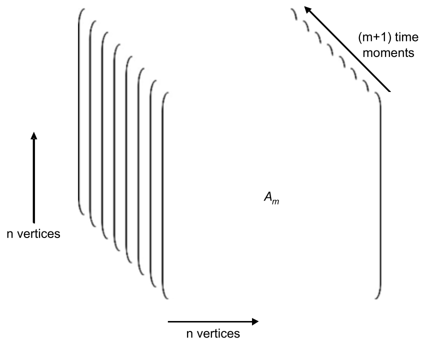


Fig. 2. A tensor representation of time-graphs. A_i represents the adjacency matrix obtained from the network snapshot obtained at time t_i. A_i is the ith slab of the adjacency tensor $\mathcal{A}$.

3.2. Expected hitting time of random walk 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 $v(t_0)$. $v(t_1)$ is chosen uniformly at random among $v(t_0)$'s neighbors. This node makes a similar decision at time t_1. The transitions of the random walk are 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 the fact 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 over 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 [21] 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

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

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

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

where ζ_k^i denotes the number of neighbors for node i 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, denoted by 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_d denote the time at which a random walk first hits node d when the random walk starts at another node $s \in V \setminus \{d\}$. Without loss of generality, assume that the random walk is initiated at time $t = 0$. The hitting time is

$$\Theta_d = \inf\{k > 0; X_k = d\}. \tag{3}$$

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 the 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

$$Q_k = \begin{bmatrix} R_k & H_k \\ 0 & 1 \end{bmatrix} \tag{4}$$

where the $(n-1) \times (n-1)$ matrix R_k represents the transition probabilities between the states in the subset $V \setminus \{d\}$ 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 $V \setminus \{d\}$ 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

$$P(\Theta_d > \eta) = P(\forall i, i \le \eta, X_i \in V \setminus \{d\}) = \alpha \left(\prod_{k=0}^{\eta-1} R_k \right) \mathbf{1}_{n-1} \tag{5}$$

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. Eq. (5) indicates the probability that the random walk still remains in $V \setminus \{d\}$ after η time steps.

Since Θ is a discrete random variable,

$$E[\Theta_d] = \sum_{\eta \ge 0} \eta P(\Theta_d = \eta) = \sum_{\eta \ge 0} P(\Theta_d > \eta) \tag{6}$$

with $P(\Theta_d > 0) = 1$ since $s \in V \setminus \{d\}$. Hence, the expected time for a random walk to hit destination d can be rewritten as

$$E[\Theta_d] = \alpha \left[I + \sum_{\eta \ge 1} \prod_{k=0}^{\eta-1} R_k \right] \mathbf{1}_{n-1} \tag{7}$$

where I is the $(n-1) \times (n-1)$ identity matrix. We take the average of $E[\Theta_d]$ values over all d in order to obtain expected hitting time $E[\Theta]$.

Consider the matrices Q_k and Q_{k+1} of Eq. (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 entries of the matrices are positive,

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

$\forall i \ne d$, which indicates that the probability of transition to d increases and the probability of the random walk remaining at a state other than the destination decreases at every time step. A more detailed discussion can be found in [21].

3.3. *Diffusion time for epidemic algorithms*

Epidemic routing has been proposed to achieve end-to-end connectivity within the context of intermittently connected mobile networks. In this method, in each contact opportunity between two nodes, the nodes exchange the messages they host. The messages are replicated in neighboring nodes when corresponding links are formed. The time it takes to disseminate the message to the entire network, i.e.the diffusion time, indicates how well mixed or how well connected the network is.

In epidemic routing, a message initiated by vertex i reaches vertex j when the first path is created between i and j. Following the notation in the previous section, we define the *reachability tensor* $\mathcal{B}$ whose kth slab is

$$B_k = \min \left(\mathbf{1}, \prod_{\eta=0}^{k} A_\eta \right) \tag{9}$$

where $\mathbf{1}$ is an all 1 $n \times n$ matrix. The product $\prod_{\eta=0}^{k} A_i$ is a $n \times n$ matrix. The entry that corresponds to the ith row and the jth column of this product matrix gives the number of k-step paths that start at i and end at j. If this value is above zero, node v_j is reachable from node v_i after k steps, and $\mathcal{B}_{ijk} = 1$. In this case, a random walk starting from node i at time zero can end up at node j at time k with a probability beyond 0 or there exists at least one k-step path connecting i to j. Otherwise $\mathcal{B}_{ijk} = 0$.

$\mathcal{B}_{ijk}$ indicates whether there exists a k-step path from i to j. Let us define another tensor $\mathcal{C}$ in the following manner:

$$\mathcal{C}_{ijk} = \max_{\eta < k} (\mathcal{B}_{ij\eta}). \tag{10}$$

When $\mathcal{C}_{ijk} = 1$, there exists a least one path from i to j that takes at most k steps. In other words, epidemic routing definitely delivers a replica of the message that starts from vertex i at time 0 to vertex j by time $t = k$. The diffusion time for such a message, τ_i, is

$$\tau_i = \inf k \quad \text{such that } \mathcal{C}_{ijk} = 1 \; \forall j \ne i. \tag{11}$$

In our evaluation, we look at the overall diffusion time, $\overline{\tau}$, that is averaged over each starting node i.

3.4. Structural properties of time-graphs

We obtain the structural properties of time-graphs using the reachability tensor $\mathcal{B}$, defined in Eq. (9). Remember that $\mathcal{B}_{ijk}$ indicates the existence of at least one k-step path that starts at i and ends at j.

The number of slabs in $\mathcal{B}$ is defined by

$$\tau = \inf(k; C_k = \mathbf{1}), \tag{12}$$

where C_k is the kth slab of $\mathcal{C}$ (as defined in (10)), assuming that there exists such $k \leq m$. This is to say that every vertex in the graph is reachable from every other vertex in τ time steps. If there is no such k, then $\tau = m$. τ is the so-called *time-diameter* of the time-graph. Note that if this procedure is applied to 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. Also note that $\tau = \sup(\tau_i)$.

$\mathcal{B}$ is a $n \times n \times \tau$ tensor, i.e. $\mathcal{B} \in \Re^{n \times n \times \tau}$. We transform this tensor into a matrix, S, by *unfolding* it around the third mode. Note that this mode corresponds to time whereas the other two modes correspond to the vertices. In other words, the third mode of the tensor is qualitatively different from the other modes, i.e. it is the *distinguished* dimension [4].

S is a $\tau \times n^2$ matrix and its kth column corresponds to the kth slab of $\mathcal{B}$, i.e. B_k. Remember that $\mathcal{B}_{ijk} = 1$ indicates the existence of a k-step path from i to j. We first normalize the rows of B_k so that $\sum_j \mathcal{B}_{ijk} = 1$ for all i and k. So, the number of non-zeros items in the ith row of B_k yields the number of nodes at which a random walk can end up at k after starting at i at time 0. Then, we reshape the normalized B_k into a one-dimensional column vector of length n^2 by appending columns one at a time. By cascading these vectors, the matrix $S \in \Re^{\tau \times n^2}$ is obtained. The singular values of this matrix are used as the structural indicators for the time-graph.

4. Evaluation

In this section, we use an extensive set of simulations to gather adjacency tensor data. Using this data, we obtain expected hitting, average diffusion time, the matricized reachability tensor and its singular values by performing the analysis we described in Section 3. Our results show that there is a very high correlation between the structural properties and the metrics that indicate the connectivity of the time-graph.

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 250 m 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 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 s. The simulation time is 1000 s. 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 correlations between the averaged values are much higher.

Each run starts with random node displacement and initial warm-up duration to reach the stationary node distribution of the mobility model. We calculate the expected hitting time for each node when the random walk is equally likely to start at any node other than the destination in order to obtain the expected hitting time for the particular time-graph.

We first present four simple scenarios to illustrate the effectiveness of the singular values of the matricized reachability in capturing end-to-end connectivity in the network. Then, we present results with a variety of mobility models with a large set of dynamism and density levels that show that our findings are not resulted from some special cases.

4.2. Illustrative example

The connectivity of time-graphs typically improves as the dynamism in the network increases. With higher dynamism, the graph *mixes* faster. However, nodes moving fast does not necessarily guarantee that the network is well connected. Instead, the structural properties that are derived from the adjacency tensor as depicted in Section 3.4 yield much better indicators of the end-to-end connectivity. We support this claim using four simple scenarios with ten nodes in a 1500×1500 m^2 area.

1. Slow scenario: Each node moves towards a random point at a speed of 1 m/s.
2. Fast scenario: Each node moves towards a random point at a speed of 5 m/s.
3. Repetitive scenario: First four nodes are located on the line $y = 250$, four other nodes are located on $y = 750$ and two nodes are located along $y = 1250$. In the initial placement, a node can be within the communication range of only one other node. All nodes move parallel to the x axis at a speed of 100 m/s. They change their direction by 180 degrees every second.

Table 1
Summary of results from example scenarios.

Scenario	$E[\Theta]$	$\overline{\tau}$	ρ_1	ρ_2
1	12 038.24	19 550.76	263.91	119.98
2	5785.32	8370.1	132.17	79.28
3	16 891.44	20 000	400	158.14
4	3294.3	4805.1	111.16	70.38

4. Ferry scenario: Nine nodes are stationary and they are placed far away from each other so that any two nodes cannot establish a link between themselves. However, there is a mobile ferry node that moves on a cyclic route at a speed of 5 m/s. Along its route, the ferry can communicate with other nine nodes.

The mobility model in the first two scenarios is known as the Random Waypoint Mobility Model (see next section). The last scenario is motivated by ferry networks [22,23]. The results from these scenarios are summarized in Table 1. Note that the quantities measuring the expected hitting time and the average diffusion time are in time steps.

In the first scenario, the nodes move very slowly and it takes a long time for a random walk to hit destinations in the network. In this scenario, it is usually not possible to reach all the nodes when the random walk starts at an arbitrary node. Hence, the overall diffusion time is also large. In accordance, we see that the first two singular values of the matricized reachability tensor, ρ_1 and ρ_2, are large. As the node speed increases in Scenario 2, all of these values decrease.

Even though the nodes move very fast in Scenario 3, the mobility pattern of the nodes prevent the end-to-end connectivity. Instead, the network is clustered into three groups. Note that the node mobility changes only x coordinates of the nodes. When there is a difference of 500 m in the y axis, two nodes cannot form a link given that the communication range is 250 m. Even though random walks hit nodes that are in the same clusters as the source nodes, end-to-end connectivity over the entire network cannot be achieved. Hence, $\overline{\tau} = m$, where m is the number of slabs in the adjacency tensor. Similarly, the ρ_1 and ρ_2 values are the largest in this scenario.

In the final scenario, there is only one mobile node. Only this ferry node is able to form links with other nodes. Even though this node moves with a speed of 5 m/s (the node speed in Scenario 2), the expected hitting time and average diffusion time values are much lower, since the trajectory of the ferry nodes enables delivery to all nodes very quickly. The corresponding singular values are the smallest in comparison to those of the other cases.

4.3. Simulation results and discussion

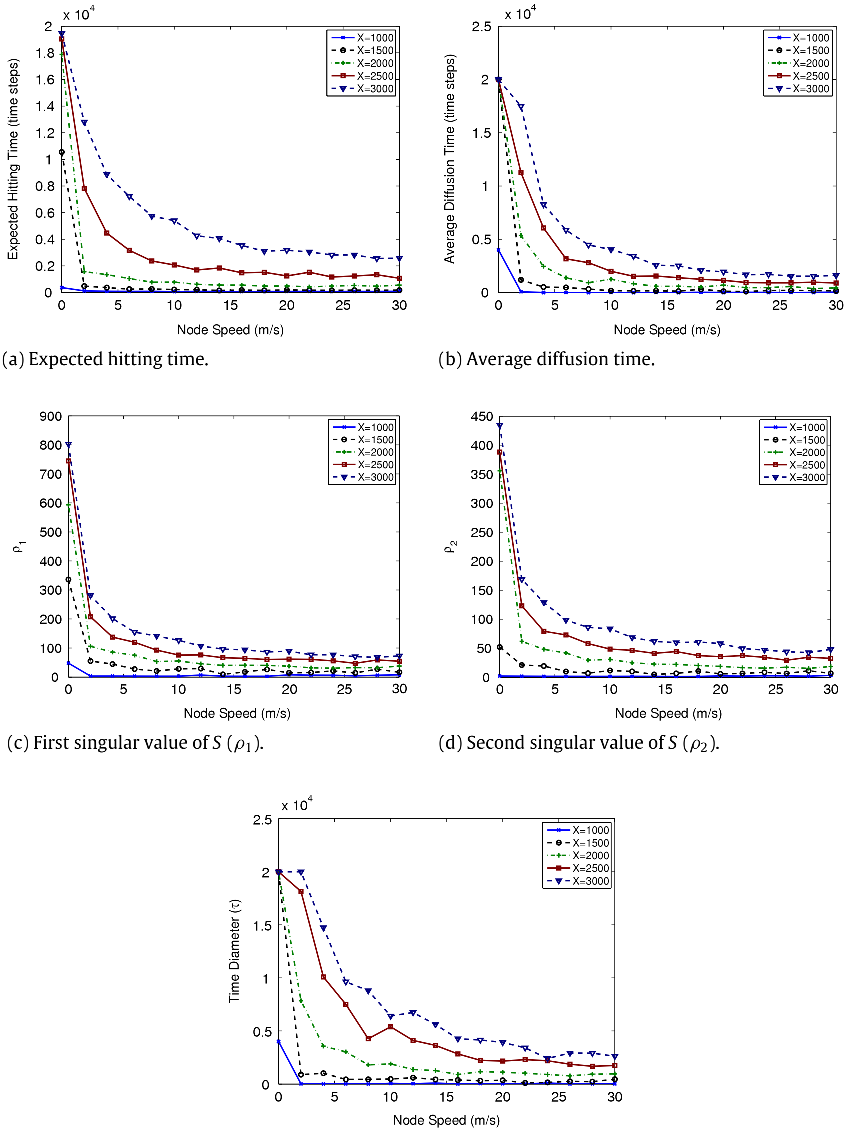
We perform a large set of simulations to show the correlation between the expected hitting time of a random walk in a time-graph and the singular values of the matricized reachability tensor. For evaluation, we used the following mobility models[1]:

- Random Walk Mobility Model: In this mobility model, mobile nodes have fixed journey durations, t. At the beginning 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, $t = 10$ s. 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 to 30 m/s.
- Random Waypoint Mobility Model: In this model, the nodes pick random 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 two-dimensional torus instead of a rectangular area. If a node reaches a boundary, it continues its movement and reappears 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, $\Delta t = 0.2$ s, $A_{max} = v_{max}\Delta t$ 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

[1] A more detailed discussion of these models can be found in [24].



(a) Expected hitting time.

(b) Average diffusion time.

(c) First singular value of S (ρ_1).

(d) Second singular value of S (ρ_2).

(e) Time-diameter (τ).

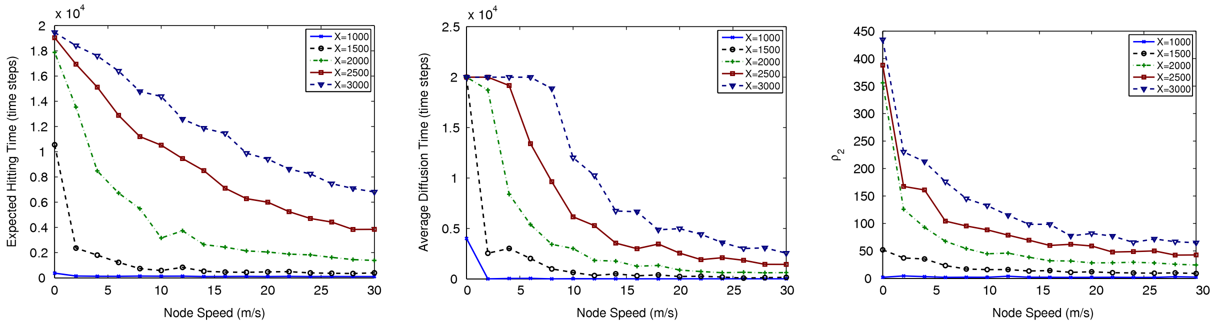
Fig. 3. The results for Random Waypoint Mobility Model. The curves that represent the expected hitting time, average diffusion time and the first two singular values of S are very similar. Experiments show that each of the singular values is highly correlated to the expected hitting time and the expected hitting time.

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 $\Delta t = 0.2$ s. 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.

Fig. 3 shows the results obtained from simulations in which nodes move according to the Random Waypoint Mobility Model. Fig. 3(a) 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 to be disconnected, the node that carries the random walk is more likely to encounter the destination or nodes that are connected to the destination at high mobility. As a result, the expression $[R_k H_{k+1}](i)$ in Eq. (8) has higher values for all k and for each entry i. Recall 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



(a) Expected hitting time.

(b) Average diffusion time.

(c) Second singular value of the reachability tensor (ρ_2).

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

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. As the mobility increases, the number of contacts between node pairs increase as well. When the messages spread in the network through an epidemic approach, the number of message passage opportunities increase. As a result, the average diffusion time decreases. This also leads to decreasing time-diameters with increasing node mobility as shown in Fig. 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 high rate, the network remains connected and the number of neighbors of each node does not change dramatically over time. Hence, the expected hitting time does not change much with mobility.

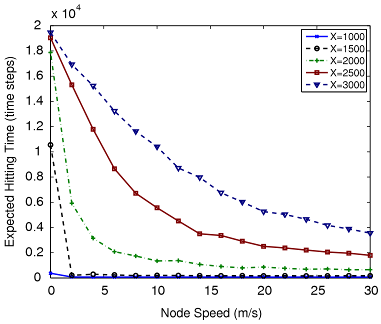
Fig. 3(c) shows how the first singular value of the matricized reachability tensor, ρ_1, changes. Fig. 3(d) depicts the characteristics of the second singular value of the matricized reachability tensor, ρ_2. Note that these curves have similar characteristics to the expected hitting time shown in Fig. 3(a) and the average diffusion time in Fig. 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. Similarly, correlation between ρ_1 and the average diffusion time is approximately 0.9. When we consider ρ_2, the correlation between the expected hitting time and the average diffusion time is very close to 0.95 and 0.9, respectively. Note that these large values indicate very strong correlation.

In the remainder of this section, we will focus on ρ_2 as the structural metric of the network. We use the second singular value because of the standard intuition from expander graphs: the second eigenvalue in static graphs predicts connectivity-related parameters, such as the expansion and conductance [25]. Besides, our results show that ρ_2 predicts end-to-end connectivity better than ρ_1.

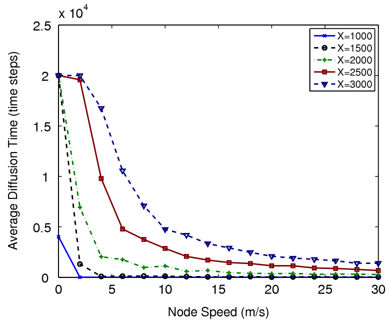
In Fig. 4, the characteristics of ρ_2 are similar when the nodes move according to the Random Walk Mobility Model. How the expected hitting time changes with speed and density is presented in Fig. 4(a); it decreases with node mobility in sparse networks, but does not change much in dense networks due to the same reasons as the ones discussed for the Random Waypoint Mobility Model. ρ_2 also has the same characteristics as shown in Fig. 4(c). The correlation between ρ_2 and the expected hitting time of a random walk is around 0.96. When the network is very sparse and the dynamism is low, the average diffusion time remains constant as shown in Fig. 4(b). This is because the number of opportunities for a message to spread with the epidemic approach is not sufficient for the message to be able to reach every node starting from any random node. Still, the correlation between the average diffusion time and the second singular value of the unfolded reachability tensor is above 0.92.

The results on the expected hitting time, the average diffusion time and ρ_2 when the nodes move according to Boundless Simulation Area Mobility Model are presented in Fig. 5. 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 and the correlation between the average diffusion time is more than 0.92. As depicted in Fig. 6, similar 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 as a function of the dynamism in the network as Fig. 6(a) and (c) show, respectively. In this mobility model, the correlation between ρ_2 and the expected hitting time is over 0.97. The correlation between ρ_2 and the average diffusion time on the other hand is approximately 0.93.

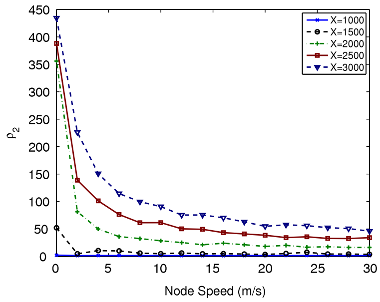
Additionally, we have performed simulations to see how the expected hitting time, the average diffusion time and ρ_2 values vary with the number of the nodes in the network and whether our observation about the correlation still holds. In this case, we used the random waypoint mobility in the node movements in various network area sizes with fixed speed values $v_{min} = 10$ m/s and $v_{max} = 20$ m/s. Fig. 7 shows the characteristics of the expected hitting time, average diffusion time and ρ_2 with respect to the number of nodes in the network. From the figures, the similarities between the characteristics of ρ_2 and other entities are eminent. ρ_2 is highly correlated both with the expected hitting time and the average diffusion time, with the respective correlation values 0.93 and 0.89.



(a) Expected hitting time.

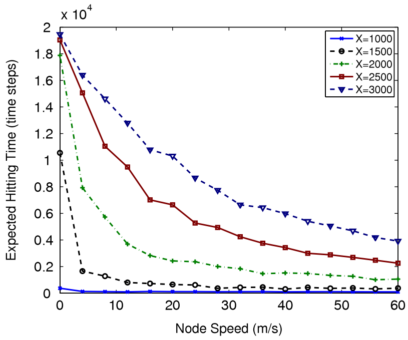


(b) Average diffusion time.

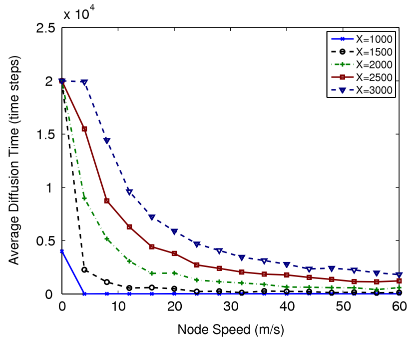


(c) Second singular value of the reachability tensor (ρ_2).

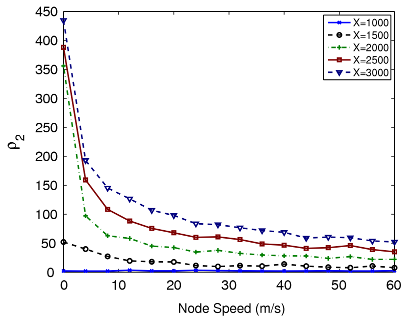
Fig. 5. The results for Boundless-area Mobility Model. The correlation between the expected hitting time and ρ_2 is again more than 0.95. The correlation between the ρ_2 and the average diffusion time is more than 0.92.



(a) Expected hitting time.



(b) Average diffusion time.



(c) Second singular value of S (ρ_2).

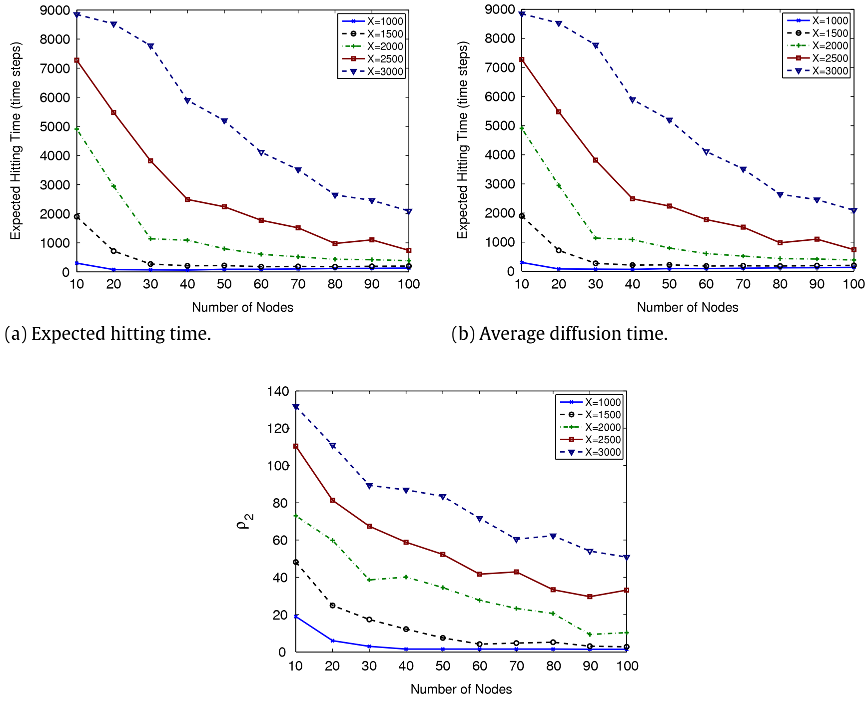
Fig. 6. The results for Gauss–Markov Mobility Model. The results are consistent with the other mobility models. The correlation between the expected hitting time and ρ_2 is very large, approximately 0.97. Considering the average diffusing time, the correlation is approximately 0.93.

Up until now, we have compared the expected hitting time and the average diffusion time with the ρ_2 values separately. When values obtained from all the simulation runs (more than 2000 runs) are considered all together, the correlation coefficient for ρ_2 and the expected time becomes slightly more than 0.9. For ρ_2 and the average diffusion time, the correlation is 0.87.

In our evaluation, we assume that the transition of the random walk between the nodes or the transition between states in the Markov Chain or the diffusion of messages in the epidemic approach takes place at once, without any delay. However, the time needed by 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 values remain very large.

5. Conclusion and open problems

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 combinatorial time-graph model is proposed. Instead of an



(a) Expected hitting time.

(b) Average diffusion time.

(c) Second singular value of S (ρ_2).

Fig. 7. Expected hitting time, average diffusion 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 whereas the correlation between the average diffusion time and ρ_2 is 0.91.

adjacency matrix, the time-graph is modeled by a 3-mode adjacency tensor, or three-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 vertex v_i is connected to vertex v_j at time t_k, the entry $\mathcal{A}_{ijk}$ in the tensor is 1. This tensor representation allowed us to obtain the expected hitting time of a random walk that proceeds in the time-graph using non-homogeneous Markov Chains and the average time for a message that diffuses in epidemic approach to spread to the entire network, i.e. average diffusion time. Starting from $\mathcal{A}$, we obtain another tensor which we refer to as the reachability tensor $\mathcal{B}$. The entry $\mathcal{B}_{ijk}$ is set to 1 if a random walk starting from node v_i can end up in node v_j at time k; otherwise it is set to 0. Each row of each slab of this tensor is normalized and then matricized or unfolded along the third mode (time axis), which is the distinguished dimension of the time-graph in which the other two modes represent the vertices. Our observations are based on an extensive set of experiments and indicate that there is a significant 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. For the epidemic approach, the correlation between this singular value and the average diffusion time is close to 0.87. These are very high correlation values; therefore, the second singular value of the matricized reachability tensor is a good indicator of connectivity for dynamic networks.

This study provides a connection between the structure and the properties of time-graphs and the network connectivity in a dynamic 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. As the eigenvalues of the tensors are understood properly, other structural elements might also emerge that predict the connectivity of the networks that are defined by tensors. These issues are among the many open problems that we plan to address in the future.

In this paper, we have showed that end-to-end connectivity in time-graphs can be predicted by the second singular value of the matricized reachability tensor. In order to do that, we have used an extensive set of simulations to gather adjacency information. An analysis that proves these claims theoretically rather than using empirical observations is an important open problem. We plan to work on this problem in future work.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant Nos. 0546402 and 0627039, and has been partially supported by the European Commission within the framework of the Euro-NF network of excellence, ICT-NoE-FP7-216366.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

[1] U.G. Acer, S. Kalyanaraman, A.A. Abouzeid, Weak state routing for large scale dynamic networks, in: MobiCom'07: Proceedings of the 13th Annual ACM International Conference on Mobile Computing and Networking, 2007 pp. 290–301.
[2] Z. Zhang, Routing in intermittently connected mobile ad hoc networks and delay tolerant networks: overview and challenges, Communications Surveys & Tutorials, IEEE 8 (1) (2006) 24–37.
[3] A. Vahdat, D. Becker, Epidemic routing for partially connected ad hoc networks, Tech. rep., Department of Computer Science, Duke University, Durham, NC, 2000. URL: http://citeseer.ist.psu.edu/vahdat00epidemic.html.
[4] M.W. Mahoney, M. Maggioni, P. Drineas, 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.
[5] N. Bisnik, A.A. Abouzeid, Optimizing random walk search algorithms in p2p networks, Computer Networks 51 (6) (2007) 1499–1514.
[6] C. Gkantsidis, M. Mihail, A. Saberi, Random walks in peer-to-peer networks, Proceedings — IEEE INFOCOM 1 (2004) 120–130.
[7] Q. Lv, P. Cao, E. Cohen, K. Li, S. Shenker, Search and replication in unstructured peer-to-peer networks, in: Proceedings of the International Conference on Supercomputing, 2002 pp. 84–95.
[8] A. Demers, D. Greene, C. Hauser, W. Irish, J. Larson, S. Shenker, H. Sturgis, D. Swinehart, D. Terry, Epidemic algorithms for replicated database maintenance, in: PODC'87: Proceedings of the Sixth Annual ACM Symposium on Principles of Distributed Computing, 1987, pp. 1–12.
[9] R. Kumar, J. Novak, P. Raghavan, A. Tomkins, On the bursty evolution of blogspace, World Wide Web 8 (2) (2005) 159–178.
[10] A. Ferreira, Building a reference combinatorial model for manets, IEEE Network 18 (5) (2004) 24–29.
[11] A. Scherrer, P. Borgnat, E. Fleury, J.-L. Guillaume, C. Robardet, Description and simulation of dynamic mobility networks, Computer Networks 52 (15) (2008) 2842–2858.
[12] M. Ben Jdidia, C. Robardet, E. Fleury, Communities detection and analysis of their dynamics in collaborative networks, in: 2nd International Conference on Digital Information Management, ICDIM 2, 2007 pp. 744–749.
[13] J. Leskovec, J. Kleinberg, C. Faloutsos, Graphs over time: densification laws, shrinking diameters and possible explanations, in: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2005 pp. 177–187.
[14] G. Palla, A. Barabasi, T. Vicsek, Quantifying social group evolution, Nature 446 (2007) 664–667.
[15] T.G. Kolda, Orthogonal tensor decompositions, SIAM Journal on Matrix Analysis and Applications 23 (2001) 243–255.
[16] D. Leibovici, R. Sabatier, 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.
[17] T. Zhang, G.H. Golub, Rank-one approximation to high order tensors, SIAM Journal on Matrix Analysis and Applications 23 (2) (2001) 534–550.
[18] M. Grossglauser, D.N. Tse, Mobility increases the capacity of ad hoc wireless networks, IEEE/ACM Transactions on Networking 10 (4) (2002) 477–486.
[19] T. Spyropoulos, K. Psounis, C.S. Raghavendra, Spray and wait: an efficient routing scheme for intermittently connected mobile networks, in: Proceedings of ACM SIGCOMM 2005 Workshops: Conference on Computer Communications, 2005, pp. 252–259.
[20] J. Burgess, B. Gallagher, D. Jensen, B.N. Levine, Maxprop: routing for vehicle-based disruption-tolerant networks, in: INFOCOM 2006, 25th IEEE International Conference on Computer Communications, Proceedings, 2006, pp. 1–11.
[21] A. Platis, N. Limnios, M. Le Du, Hitting time in a finite non-homogeneous markov chain with applications, Applied Stochastic Models and Data Analysis 14 (3) (1998) 241–253.
[22] W. Zhao, M. Ammar, E. Zegura, A message ferrying approach for data delivery in sparse mobile ad hoc networks, in: Proceedings of the International Symposium on Mobile Ad Hoc Networking and Computing, MobiHoc, 2004 pp. 187–198.
[23] R.C. Shah, S. Roy, S. Jain, W. Brunette, Data mules: modeling and analysis of a three-tier architecture for sparse sensor networks, Ad Hoc Networks 1 (2–3) (2003) 215–233.
[24] T. Camp, J. Boleng, V. Davies, A survey of mobility models for ad hoc network research, Wireless Communications and Mobile Computing 2 (5) (2002) 483–502.
[25] N. Linial, A. Wigderson, Expander graphs and their applications, in: Lecture Notes, The Hebrew University, Israel, 2003.
[26] U.G. Acer, P. Drineas, A.A. Abouzeid, Random walks in time-graphs, in: ACM/Sigmobile MobiOpp'10: The Second International Workshop on Mobile Opportunistic Networking.