How to Explore a Fast-Changing World

(Extended abstract)

Chen Avin[1], Michal Koucký[2], and Zvi Lotker[1]

[1] Communication Systems Engineering, Ben Gurion University of the Negev. Israel.
`{avin,zvilo}@cse.bgu.ac.il`
[2] Academy of Sciences of Czech Republic
`koucky@math.cas.cz`

Abstract. Motivated by real world networks and use of algorithms based on random walks on these networks we study the simple random walks on dynamic undirected graphs, i.e., graphs which are modified by inserting or deleting edges at every step of the walk. We are interested in the expected time needed to visit all the vertices of such a dynamic graph, the *cover time*, under the assumption that the graph is being modified by an oblivious adversary. It is well known that on static undirected graphs the cover time is polynomial in the size of the graph, on the contrary and somewhat counter-intuitively, we show that there are adversary strategies which force the expected cover time of dynamic graphs to be exponential, and relate this result to the cover time of static directed graphs. In addition we provide a simple strategy, the *lazy* random walk, that guarantees polynomial cover time regardless of the changes made by the adversary.

1 Introduction

A random walk on a graph is a simple process of visiting the nodes of the graph in some random sequential order. The walk starts at some fixed node, and at each step it moves to a neighbor of the current node chosen at random. The random walk is called *simple* when the next node is chosen uniformly at random from the set of neighbors. In the context of communication networks (e.g., Internet, wireless ad-hoc networks and sensor networks) and information networks (e.g., peer-to-peer file sharing networks and distributed databases), a random walk on a network (graph) will result when messages are sent at random from device to device.

Since this process presents locality, simplicity, low memory-overhead and robustness to changes in the network structure applications based on random-walk techniques are becoming more and more popular in the networking community. In recent years, different authors have proposed the use of random walk for a large variety of tasks and networks; to name but a few: querying in sensor and ad-hoc networks [**?**,**?**,**?**], searching in peer-to-peer networks [**?**], gossiping [**?**], PageRank and search engines on the web [**?**].

One of the main reasons that random walk techniques are so appealing for networking application is their robustness to dynamics. Many communication networks are subject to dramatic structural changes created by mobility, sleep modes, channel fluctuations, device failures, nodes joining or leaving the system and other factors. Topology-driven algorithms are at a disadvantage for such networks, since they incur high overhead to maintain up-to-date topology and routing information such as routing tables, clusters and spanning trees. In contrast, algorithms that require no knowledge of network topology, such as the random walk, are at an advantage.

While at first glance, the process of a token wandering randomly in the network may seem overly simplistic and highly inefficient, many encouraging results prove that it is comparable to other approaches that have been used over the years. One important property of random walks on graphs that needs to be evaluated to study the efficiency of the approach is the *cover time* [**?**]. The *cover time* C_G of a graph G is the expected time (measured by number of steps or in our case by the number of messages) taken by a simple random walk to visit all the nodes in G. Methods on bounding the cover time of graphs have been thoroughly investigated with the major result being that cover time is always at most polynomial for undirected graphs. More precisely, it has been shown by Aleliunas *et al.* in their seminal work [**?**] that C_G is always $O(mn)$, where m is the number of edges in the graph and n is the number of nodes. Tighter bounds for many classes of graphs have been established and they can be found in the extensive literature on the subject.

Since real-world networks change over time researches have recently started to study random walks on such dynamic graphs. Motivated by robotic exploration of the Web, Cooper and Frieze [**?**] studied the question of covering a graph that grows over time. They considered a particular model of so-called *web graphs* and showed that a simple random walk on the graph fails to visit a constant fraction of nodes if a new node appears and is connected to the graph after every constant number of steps of the walk.

Motivated by sensor networks we consider a similar question on a different model of dynamic graphs. We consider dynamic graphs with fixed number of nodes where connections between the nodes appear and disappear over the time. The question that we study is the cover time of such graphs.

1.1 Overview of Our Results

We show that somewhat counter-intuitively, there are dynamic graphs of this type that have exponential cover time when explored by a simple random walk. (For the sake of clarity let us say that our examples are deterministic but oblivious to the actual random walk.) Moreover, we show that a random walk on any directed graph G can be simulated (in a way we define later) by a random walk on an undirected dynamic graph that we construct from G; this gives yet another justification to our previous claim. Our examples are also valid when we allow the random walk to make more than a single step between each graph change. Indeed, we can allow up-to $n^{1-\epsilon}$ steps before making each change and still obtain an exponential cover time. Although one could question whether our graphs could appear in a real-word scenario we do not consider these graphs to be far-fetched: for example a particular implementation of sensor networks with links (network interfaces) going to sleep periodically or nodes switching communication frequencies could exhibit such behavior.

In addition to these negative results we also show several positive results. Most importantly we show that a *lazy* random walk (also known as a *max-degree* random walk in literature [**?**]) does not suffer from these issues. We define as a lazy random walk a walk that picks each adjacent edge with probability $1/d_{\max}$, where $d_{\max}$ is the maximum degree of the graph, and with the remaining probability it stays at the current vertex. We show that a lazy random walk covers any connected dynamic graph in time polynomial in the size of the graph. Furthermore, we also show that when the dynamic graph itself is obtained by sampling from a certain probability distribution, a simple random walk will also cover such a graph in expected polynomial time.

2 Models and Preliminaries

2.1 Random Walks on Graphs

Let $G(V, E)$ be an undirected graph, with V the set of nodes and E the set of edges. Let $n = |V|$ and $m = |E|$. For $v \in V$, let $N(v) = \{u \in V \mid (vu) \in E\}$ be the set of neighbors of v and $d(v) = |N(v)|$ the degree of v. A d-regular graph is a graph in which the degree of all the nodes is d.

The *simple random walk* is a walk where the next node is chosen uniformly at random from the set of neighbors of the current node, i.e., when the walk is at node v, the probability to move in the next step to u is $P(v, u) = \frac{1}{d(v)}$ for $(v, u) \in E$ and 0 otherwise.

The *hitting time*, H_{uv}, is the expected time for a random walk starting at u to arrive at v for the first time, and the *commute time*, C_{uv}, is the expected time for a random walk starting at u to first arrive at v and then return to u. Let $H_{\max}$ be the maximum hitting time over all the pairs of nodes in G.

The *cover time* C_G of a graph G is the expected time taken by a simple random walk on G to visit all the nodes in G. Formally, for $v \in V$, let C_v be the expected number of steps needed for the simple random walk starting at v to visit all the nodes in G, and the cover time of G is $C_G = \max_v C_v$. The *cover time* of graphs and methods of bounding it have been extensively investigated [**?**]. Results for the cover time of specific graphs vary from the *optimal cover time* of $\Theta(n \log n)$ associated with the complete graph to the worst case of $\Theta(n^3)$ associated with the lollipop graph [**?**,**?**].

2.2 Evolving Graphs Model

The most general model to describe a dynamic network is called the Evolving Graph model. We will use a similar definition as in [?,?,?].

Definition 1 (Evolving Graphs) *Let $\mathcal{G} = G_1, G_2, \ldots$ be an infinite sequence of graphs on the same vertex set V. We call this sequence an* evolving graph. *We say that $\mathcal{G}$ has the graph property X if every graph G_i in the sequence has the property X.*

In simple words at time i the structure of the evolving graph $\mathcal{G}$ is G_i. For an integer $\tau \geq 1$, an evolving graph $\mathcal{G}$ is *evolving with rate* $\frac{1}{\tau}$ if for all $i \geq 1$, $G_i \neq G_{i+1}$ implies, $G_{i+1} = G_{i+1+j}$ for all $j \in \{0, \ldots, \tau - 1\}$.

A simple random walk on evolving graph $\mathcal{G}$ is defined as follows: assume that at time i the walker is at node $v \in V$, and let $N(v)$ be the set of neighbors of v in G_i, then the walker moves to one of its neighbors from $N(v)$ uniformly at random.

The strength and the weakness of the above model have the same origin, its generality. On the positive side, it captures most interesting scenarios of dynamic networks, but on the other hand, most natural problems are NP-complete such as finding strongly connected components and the equivalence of minimum spanning tree [?].

2.3 Constructive Evolving Graphs Model

Evolving graphs do not capture the underlying mechanism of how (or why) the graph evolves. In many situations the evolving graph itself is a product of some random process. (For example this is the case of web graphs considered in [?].) We will use the following definition to capture the underlying process in the case it is a Markov chain. A special case of such graphs is considered in Section **??**.

Definition 2 (Markovian Evolving Graphs) *Let the* space set $\mathbf{G}$ *be a set of graphs with the same set V of nodes and let P be a probability transition matrix. A Markov Evolving Graphs $\mathcal{M} = (\mathbf{G}, P)$ is a Markov chain with $\mathbf{G}$ and P: It is a sequence of random graphs $G_1, G_2, G_3, ...$ with the Markov property, namely that, given the present graph, the future and past graphs are independent. Formally, for $g, g_1, \ldots, g_t \in \mathbf{G}$*

$$\Pr[G_{t+1} = g \mid G_t = g_t, \ldots, G_1 = g_1] = \Pr[G_{t+1} = g \mid G_t = g_t]$$

and for $g, h \in \mathbf{G}$ the transition probability is defined by P:

$$P(h, g) = \Pr[G_{t+1} = g \mid G_t = h]$$

3 Exponential Hitting Time of Evolving Graphs

In this section we address the cover time of the simple random walk on evolving graphs by studying the maximum hitting time. Clearly the cover time must be at least as large as the maximum hitting time. First, we mention some technical issues. On static graphs the cover time is finite only for connected graphs. This is not the case for evolving graphs as we will see in Section **??**. For simplicity though we restrict our discussion mostly to evolving graphs in which every graph in the sequence $\mathcal{G}$ is connected. (In the Markovian model we require that every graph in the set $\mathbf{G}$ is connected, and call $\mathbf{G}$ *connected* if this is the case). Moreover we require that all graphs have a self-loop for each of the nodes. This is simply a technical condition to avoid pathological cases such as the walk switching forever between two nodes. In the case of static graphs this is a standard way of enforcing ergodicity. An evolving graph $\mathcal{G}$ that has the above properties we call an *explorable* evolving graph.

Now one can easily claim the following for explorable evolving graphs (a similar claim can be made for Markovian evolving graphs):

Claim 3 *Let $\mathcal{G}$ be an explorable evolving graph, then the cover time of $\mathcal{G}$ is bounded by $n^{O(n)}$.*

We outline the argument here. Fix two vertices u and v of G. For $i \geq 1$, let V_i be the set of vertices that could be visited within first i steps of a simple random walk on $\mathcal{G}$ starting from u. Since $\mathcal{G}$ is connected it must be the case that for each i, $V_i \subsetneq V_{i+1}$. In particular, V_n must contain all vertices of $\mathcal{G}$ and in particular v. Thus, the probability of reaching v starting from u is at least n^{-n}. Indeed, this is true for any two vertices u and v and starting from any time t. A standard argument now implies that the cover time is at most $n^{O(n)}$.

Requiring connectivity at each step of the evolving graph may look like a very strong condition that should imply polynomial cover time and maximum hitting time. Surprisingly we show that this is not the case.

Theorem 4 *There exists an explorable evolving graph $\mathcal{G}$, such that the maximum hitting time of the simple random walk on $\mathcal{G}$ is $\Omega(2^n)$.*

One can think of this result in the following way: consider a random walk on an evolving graph that is controlled by an oblivious *adversary* that is deciding what will be the next graph at each time step. In such a case the adversary, although unaware of the random walk location, can force the walk to step exponential number of steps before exploring the whole graph.

We give below the basic details of the proof and the main lemma behind it. Let G_1 be the star graph of size n (with the addition of a self-loop at each node) where nodes $1, 2, \ldots, n-2$ and n are always the leafs and node $n-1$ is always the center. The random walk starts at node 1 and we will bound the hitting time to node n. The adversary is the following *deterministic* process: At each time step vertices $1, \ldots, n-1$ will trade their places, i.e., the adversary changes the edges by changing the names of the nodes. The adversary uses the following renaming strategy: for $1 \leq i \leq n-1$, node i changes its name to $(i \bmod n-1) + 1$. Note that node n does not change its name, nodes $1, \ldots n-2$ increase their name by one, node $n-1$ becomes 1.

Proof of Theorem **??**. The only way to reach node n is through the center. By induction on $i = 2, \ldots, n-2$ one can see the following. Unless we have already reached the center of the star the only way how to be at the leaf named i after the adversary move is to be at the leaf named $i-1$ before the adversary renaming. That implies we must have used a self-loop at that random step. Hence, to get to the leaf named $n-2$ we must have had a sequence of $n-3$ random steps all taking a self-loop. To get to the center we have to stay at leaf $n-2$. All in all to be at the center after the adversary move the random walk must have made a sequence of $n-2$ consecutive self-loop steps. That happens with probability 2^{-n+2}. Therefore the expected time before we observe the random walk to make such a sequence of steps is $\Omega(2^n)$. □

We would like to point out that all graphs in $\mathcal{G}$ are isomorphic and rapidly mixing (the cover time of each of them is in fact $O(n \log n)$). This fact shows that common tools like spectral analysis cannot be applied naively to dynamic graphs.

3.1 Simulating Directed Graphs

One way to understand the results of the previous section is by relating random walks on explorable evolving graphs to random walks on static directed graphs. In fact we can simulate a simple random walk on a directed graph G by a careful choice of evolving graph $\mathcal{G}$. We will use the following gadget $\mathcal{H}$ to replace every directed edge of G. For $\ell > 0$, the gadget $\mathcal{H}_\ell$ is a sequence of graphs $H_\ell^0, H_\ell^1, H_\ell^2, H_\ell^0, H_\ell^1, H_\ell^2, H_\ell^0, \ldots$ with vertices L, R, s_0 and $s_{i,j}$, for $i = 1, \ldots, \ell$ and $j = 0, 1, 2$. The graph H_ℓ^k is obtained from the graph in Fig. **??** by mapping vertices $L \to L$, $R \to R$, $s_0 \to c_0$ and $s_{i,k} \to c_i$, $s_{i,k+1 \bmod 3} \to b_i$, $s_{i,k+2 \bmod 3} \to a_i$, $i = 1, \ldots, \ell$. (We deviate here from our convention of having self-loops at every node for the sake of simplicity of the analysis. As it will be clear in the next section with minor modification of bounds our claims would be true even if we would add a self-loop to every node.) The main property of a simple random walk on $\mathcal{H}$ is summarized in the following lemma.

Lemma 5 *Let $\ell > 0$, $\mathcal{H}_\ell = H_\ell^0, H_\ell^1, H_\ell^2, H_\ell^0, H_\ell^1, H_\ell^2, H_\ell^0, \ldots$ and $\epsilon = \ell(1/2)^\ell + (3/4)^\ell$. Consider a simple random walk on $\mathcal{H}_\ell$. If the walk starts at vertex L then the probability of returning to L before visiting R is at least $1 - \epsilon$. Moreover if the walk starts at vertex R then the probability of returning to R before visiting L is at most ϵ.*

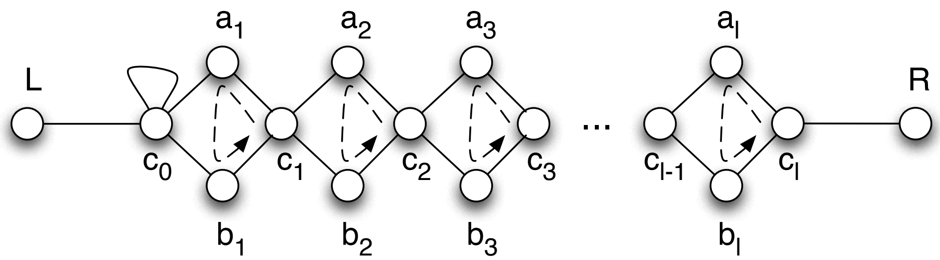

Fig. 1. The gadget $\mathcal{H}_\ell$ (dashed lines show node transformations)

We provide proof of this lemma in the appendix. Thus the gadget $\mathcal{H}_\ell$ has essentially the same effect for a simple random walk as a directed edge from R to L with a self-loop at L. Given a directed graph G with a self-loop at every vertex we can replace all its directed edges between different vertices by a copy of $\mathcal{H}_\ell$ to obtain a sequence of graphs $\mathcal{G}$ on which a simple random walk will simulate a simple random walk on G (up-to some error ϵ). Of course, replacing several edges incoming to a vertex by the gadget will introduce several self-loops to that vertex. To avoid that we can collapse the vertices c_0 from these gadgets into one thus obtaining an equivalent of one self-loop. (This collapse will affect ϵ slightly but no more than by a factor polynomial in the number of replaced edges.) We also remove the original self-loops from the graph G.

If we perform a simple random walk on $\mathcal{G}$ and we restrict ourselves to observing only visits to the vertices of the original graph G we will observe essentially the same probability distribution as of a simple random walk on G. In particular, if we choose $\ell = n^{k+1}$, for $k > 1$ and n being the size of G, then the probability of observing an edge being traversed in the opposite direction in the first 2^{n^k} steps is at most $2^{-O(n^{k+1})}$. Since for example the maximal hitting time on any strongly connected directed graph is bounded by $2^{O(n \log n)}$ this error is negligible.

4 Slowly Evolving Graphs

The previous section has shown that there are evolving graphs for which a simple random walk essentially fails as a means of exploring it. All our examples so far considered graphs that evolve at rate one. This would not really be a typical case in a real-world application. The rate at which graphs evolve is usually slower compared to unit operations such as sending a packet. So could it be the case that a simple random walk covers in polynomial time all graphs evolving at lower rate? In this section we show that this is not the case. Namely for any constant $0 < \epsilon < 1$ and an integer n large enough, we provide an example of an evolving graph on $O(n)$ vertices that evolves at rate $\frac{1}{n^{1-\epsilon}}$ so that a simple random walk needs expected time $2^{\Omega(n^\epsilon)}$ to cover the graph. Indeed the graph is essentially the gadget from the previous section with the speed of evolution slowed down.

Let K^i_ℓ be the graph H^i_ℓ from the previous section modified by adding possibly several self-loops to each vertex so that the probability of staying at the same vertex is precisely one half. (So in particular vertices of degree two will receive two self-loops and vertices of degree four will receive four of them. We remark that our claim would be true even without these self-loops but in some cases for trivial reasons. So to capture the most general situation we introduce the loops.)

For $0 < \epsilon < 1$ and an integer $n \geq 2^{1/(1-\epsilon)}$, we define an evolving graph $\mathcal{G}^\epsilon_n$ to consist of repeated sequence $K^0_{2n}, K^0_{2n}, \ldots, K^0_{2n}, K^1_{2n}, K^1_{2n}, \ldots, K^1_{2n}, K^2_{2n}, \ldots, K^2_{2n}$, where each block of K^i_{2n} consists of $n^{1-\epsilon}$ copies of K^i_{2n}. Clearly $\mathcal{G}^\epsilon_n$ evolves at rate $\frac{1}{n^{1-\epsilon}}$. We claim:

Theorem 6. *The cover time of $\mathcal{G}^\epsilon_n$ is $2^{\Omega(n^\epsilon)}$.*

To prove the theorem we need tools from the next section.

4.1 A random walk on a line with a drift

In order to analyze the simple random walk on our slowly evolving graph we introduce a *random walk with a drift*. Let $k \geq 1$ be an integer. Let $p = (p_{-k}, \ldots, p_0, \ldots, p_k)$ be a probability distributions on $\{-k, -k+1, \ldots, k\}$. Let $\ell \geq 1$ be an integer. A *random walk on a line* $(\ldots, -1, 0, 1, \ldots)$ *with a drift* is a random walk that starts at origin (0) and at each step the walk makes a step to the left with probability 1/4, step to the right with probability 1/4 and stays at the current vertex with probability 1/2. Moreover, after each ℓ steps of the walk a *drift step* is taken: we pick $i \in \{-k, \ldots, k\}$ at random according to p and we make $|i|$ steps to the left if $i < 0$, and i steps to the right if $0 \leq i$. We will show that if the expectation of the drift step is negative then the probability of reaching a point at distance n on the line before visiting the point $-n$ is exponentially small in n provided that $\ell = n^{1-\epsilon}$, for some constant $0 < \epsilon < 1$.

For our purposes we will need a slight generalization of the random walk with a drift as the drift step will depend on whether we are at an even or odd distance from the origin. Thus we will consider the case of having two probability distributions p^{even} and p^{odd} on $\{-k, \ldots, k\}$, and the drift step will be taken according to p^{even} if the current vertex will be at an even distance from the origin and according to p^{odd} otherwise.

For two probability distributions p^{even} and p^{odd} on $\{-k, \ldots, k\}$ denote by $D^+ = \sum_{i>0} i \cdot (p_i^{\text{even}} + p_i^{\text{odd}})$ and $D^- = \sum_{i<0} i \cdot (p_i^{\text{even}} + p_i^{\text{odd}})$. We claim the following lemma.

Lemma 7 *Let $k \geq 1$ be an integer, $0 < \epsilon < 1$ be a constant, and p^{even} and p^{odd} be two probability distributions on $\{-k, \ldots, k\}$. If $D^+ + D^- < 0$ then for any n large enough, the probability that a random walk on a line with drifts according to p^{even} and p^{odd} starting from the origin, reaches some point $\leq -n$ before reaching a point $\geq n$ is at least $1 - 2^{-c'n^{\epsilon}}$. Here $c = -6/(D^+ + D^-)$ and $c' > 0$ is some constant that depends only on $\epsilon, k, p^{\text{even}}$, and p^{odd}.*

Proof of Lemma **??**. Consider a random walk of length $t = c \cdot n \cdot n^{1-\epsilon}$ on a line with drifts starting from the origin. We will argue that with probability exponentially close to 1 the walk will visit some point $\leq -n$ but not any point $\geq n$. During the walk of length t, there will be cn drifts. We will argue that with high probability the total drift towards the negative side will be at least $2n$ (i.e., we will move at least $2n$ steps to the left as the consequence of the drifts), at no point during the walk the total drift so far was more than $n/2$ to the right and at no point the random walk on the line without drifts would reach point $n/2$. If these three events occur simultaneously then we clearly must have reached a point $\leq -n$ during the walk but never reached any point $\geq n$.

We argue first about the random walk of length t without drifts. Consider a random walk of length t on a line which makes a step to the left with probability 1/4, to the right with probability 1/4 and stays at the current vertex with probability 1/2. The probability that the walk visits a point $n/2$ during its t steps is at most twice the probability that it is at some point $\geq n/2$ at step t. (Once we reach the point $n/2$ we have the same probability of finishing to the left of it as finishing to its right.) By the usual Chernoff bound (cf. [**?**]), the probability that we are at a position $\geq n/2$ after t steps is at most $e^{-(n/2)^2/2t} = e^{-n^{\epsilon}/8c}$. Hence the probability that the random walk without drifts reaches $n/2$ during its t steps is at most $2e^{-n^{\epsilon}/8c}$.

We will argue now about the drifts. We show for $m = n/2k, \ldots, cn$ that with probability exponentially small in m (and hence exponentially small in n) the total effect of m random drifts is a movement to the left by at least $2m/c$ positions. We need the following claim.

Claim 8 *Let $\delta = 1/(10\binom{k+1}{2}c)$. With probability at least $1 - 2^{-\Omega(m)}$, for each $i \in \{-k, \ldots, k\}$, the number m_i of drift steps by i positions within m drift steps is bounded by:*

$$\left(\frac{1}{2} - \delta\right) \cdot \left(p_i^{\text{odd}} + p_i^{\text{even}} - 2\delta\right) m \leq m_i \leq \left(\frac{1}{2} + \delta\right) \cdot \left(p_i^{\text{odd}} + p_i^{\text{even}} + 2\delta\right) m.$$

Assuming the claim let us conclude the proof of the lemma. Clearly, the total drift of m drifts is given by $\sum_{i=-k}^{k} im_i$. By the previous claim we can bound the total drift (unless some small probability event happens) as follows:

$$\begin{aligned}\sum_{i=-k}^{k} im_i &\le \sum_{i=-k}^{-1} i \cdot \left(\frac{1}{2}-\delta\right)(p_i^{\text{odd}}+p_i^{\text{even}}-2\delta)m \\ &\quad + \sum_{i=1}^{k} i \cdot \left(\frac{1}{2}+\delta\right)(p_i^{\text{odd}}+p_i^{\text{even}}+2\delta)m \\ &\le \sum_{i=-k}^{-1} i \cdot \frac{1}{2}(p_i^{\text{odd}}+p_i^{\text{even}})m \\ &\quad + \sum_{i=1}^{k} i \cdot \frac{1}{2}(p_i^{\text{odd}}+p_i^{\text{even}})m \\ &\quad + 2\binom{k+1}{2}(2\delta+\delta+2\delta^2)m \\ &\le \frac{1}{2}(D^+ + D^-)m + 10\binom{k+1}{2}\delta m \\ &= -3m/c + m/c = -2m/c.\end{aligned}$$

The first $(n/2k)-1$ drifts have always cumulative effect smaller than $n/2$. Based on the previous calculation, with probability at least $1-cn2^{-\Omega(n/2k)}$ for all $m = n/2k, \ldots, cn$, the cumulative effect of the first m drifts is negative, so in particular it is smaller than $n/2$. Moreover the total drift to the left after cn drifts is $\ge 2n$. That concludes our argument. Thus we only need to establish the claim.

Proof of the Claim: Let m^{even} and m^{odd} be the number of drifts according to p^{even} and p^{odd}, resp., among the first m drifts during a walk on a line with drifts. We will establish first that with probability exponentially close to 1, $\left(\frac{1}{2}-\delta\right)m \le m^{\text{even}}, m^{\text{odd}} \le \left(\frac{1}{2}+\delta\right)m$. By Binomial Theorem

$$0 = (1-1)^{n'} = \sum_{k \text{ is even}} \binom{n'}{k} - \sum_{k \text{ is odd}} \binom{n'}{k}$$

Thus regardless of the starting point of a random walk, the probability that the walk of length $n^{1-\epsilon}$ on a line ends at an even distance from the origin is precisely the same as that the walk ends at an odd distance. Hence, the expectation of m^{even} and m^{odd} is precisely $m/2$. Applying Chernoff bound on these two random variables we obtain that with probability at least $1-2^{-\Omega(m)}$, both m^{even} and m^{odd} are in the range $((1/2-\delta)m, (1/2+\delta)m)$.

Conditioned on that m^{even} and m^{odd} are in the above range, again using Chernoff bound, we conclude that with probability at least $1-2^{-\Omega(m)}$, the number m_i of drifts by i is between $(p_i^{\text{even}}-\delta)m^{\text{even}}+(p_i^{\text{odd}}-\delta)m^{\text{odd}}$ and $(p_i^{\text{even}}+\delta)m^{\text{even}}+(p_i^{\text{odd}}+\delta)m^{\text{odd}}$. That concludes the proof of the claim. □

4.2 Proof of Theorem ??

In this section we argue that the cover time of $\mathcal{G}_n^\epsilon$ from Theorem **??** is $2^{\Omega(n^\epsilon)}$. Our goal is to argue that the expected hitting time of vertex R is $2^{\Omega(n^\epsilon)}$ when the simple random walk starts from vertex L. In order to prove this we will establish the following claim.

Claim 9 *Consider a simple random walk on $\mathcal{G}_n^\epsilon$ and let us assume that at some time t it is at a vertex corresponding to either one of a_n, b_n, c_n in Fig.* **??**. *Then the probability that the walk after time t will visit a vertex corresponding to some vertex to the left of $c_{3n/4}$ before visiting R is at least $1-2^{-\Omega(n^\epsilon)}$.*

We establish this claim below by relating the walk on $\mathcal{G}_n^\epsilon$ to a random walk on a line with drift. We conclude the proof of the theorem first assuming the claim. A random walk starting from L must pass through a vertex that corresponds to either one of a_n, b_n, c_n in Fig. **??**. We will focus on times when it passes through such a vertex. Let t be the first time it passes through one of these vertices. By the above claim the probability that it would reach R without visiting some vertex to the left of $c_{3n/4}$ is at most $2^{-\Omega(n^\epsilon)}$. But once it visits some vertex to the left of $c_{3n/4}$ it again has to pass through one of the vertices corresponding to either one of a_n, b_n, c_n. We stop at that time and repeat the process again. It is clear that expected number of repetitions of the process is $2^{\Omega(n^\epsilon)}$. Hence the expected hitting time of R starting from L is at least $2^{\Omega(n^\epsilon)}$.

It remains to prove the above claim. We prove it by relating the walk on $\mathcal{G}_n^\epsilon$ to a random walk on a line with a drift. Consider the case when t is a multiple of $n^{1-\epsilon}$ and that we are at the vertex corresponding to c_n in G_t. For each graph $G_{t'}$ in the sequence $\mathcal{G}_n^\epsilon$, $t' \geq t$, we map the vertex corresponding to c_{n+j} to a point $2j$ on a line and both vertices corresponding to a_{n+j} and b_{n+j} to the same point $2j-1$ on a line. Here, $-3n/4 \leq j \leq 3n/4$. Clearly a random walk on $\mathcal{G}_n^\epsilon$ from time t induces a walk on a line $-3n/2, \ldots, 0, \ldots, 3n/2$ with a drift every $n^{1-\epsilon}$ steps (until we would leave the vertices for which we defined the mapping.) The drift depends on whether we are at an odd position or even position on the line. If we are at the odd position $2j-1$, then in $\mathcal{G}_n^\epsilon$ we are with probability $1/2$ at the vertex corresponding to a_{n+j} and with the same probability at the vertex corresponding to b_{n+j}. Hence the drift at odd position has the following distribution: $(0, 1/2, 1/2)$. If we are at the even position $2j$ then we are at the vertex corresponding to c_{n+j} and the drift will have the distribution $(1, 0, 0)$. Since the expected drift is negative, Lemma **??** implies that with probability $\geq 1 - 2^{-\Omega(n^\epsilon)}$ we will reach a position $\leq -n$ before reaching a position $\geq n$. (Our stopping condition guarantees that we will not leave the region on the line where we defined the mapping so Lemma **??** can indeed be applied to our walk.) This translates into reaching a vertex to the left of $c_{3n/4}$ before visiting R.

The case where in the graph G_t the walk is at a vertex corresponding either to a_n or b_n is analyzed similarly. The assumption that t is a multiple of $n^{1-\epsilon}$ can be removed by letting the walk run till we reach the closest multiple of $n^{1-\epsilon}$ and doing similar analysis. (Within these less than $n^{1-\epsilon}$ steps we can reach only a vertex corresponding to $a_{n'}$, $b_{n'}$ or $c_{n'}$ for some $|n'-n| \leq n^{1-\epsilon}$. We map this vertex to the origin of the line and continue with the argument as before.) This concludes the proof of the claim and hence of Theorem **??**.

5 Polynomial Cover Time of Dynamic Graphs

We turn our attention to cases were the cover time of evolving graphs is "good", i.e., polynomial. Our first example is of a simple Markovian case.

Definition 10 (Bernoulli evolving graph) *Let* $\mathbf{G}$ *be a set of graphs with the same set* V *of nodes and* $\bar{P}$ *a probability distribution over* $\mathbf{G}$*. A Bernoulli evolving graph* $\mathcal{B} = (\mathbf{G}, \bar{P})$ *is a Markovian evolving graph in which the rows of the transition matrix* P *are identical and equal to* $\bar{P}$*, i.e., the random graphs* G_i*, are i.i.d.*

We show that the bound for the cover time of the simple random walk on Bernoulli evolving graphs is very similar to the bound of static graphs; essentially when the process is time invariant and the graph is always connected then the bound of Aleliunas *et al.* [**?**] can be extended to dynamic graphs.

Theorem 11 *For any explorable Bernoulli evolving graph,* $\mathcal{B} = (\mathbf{G}, \bar{P})$*, the cover time of the simple random walk on* $\mathcal{B}$ *is* $O(n^3 \log n)$ *and the maximum hitting time is* $O(n^3)$*.*

Proof. (**outline**): We will give the proof for the case where $\bar{P}$ is the uniform distribution (the case of general $\bar{P}$ is a simple extension). We first show that the simple random walk on the Bernoulli evolving graph is isomorphic to a reversible Markov chain M. This chain is a simple random walk on the multi-graph G', which is the union of all graphs in $\mathbf{G}$. This chain is, in turn, isomorphic to a random walk (not simple) on a weighted graph G''. We use the theory of electrical networks and random walks [**?**,**?**,**?**] to bound the maximum hitting time between any two nodes in V. In particular, we use the equality for commute time in a reversible Markov chain, which states that: $C_{uv} = WR_{uv}$, where C_{uv} is the commute time, W is the total sum of the edge weights (in both directions) and R_{uv} is the effective electrical resistance between u and v. Since W is equal to the total number of edges in G', we have $W \leq jn^2$, where j is the cardinality of $\mathbf{G}$. Next, we bound R_{uv}. Since every graph in $\mathbf{G}$ is connected, there are at least j edge-disjoint paths between u and v in G', each of length at most n. From this, we can conclude that $R_{uv} \leq n/j$. Therefore we have $H_{uv} < C_{uv} \leq n^3$ for any pair of nodes in V and $H_{\max} < n^3$. To bound the cover time, we use Matthews' bound [**?**] which bounds the cover time of reversible Markov chains using the maximum hitting time as $C \leq O(H_{\max} \log n)$.

The property that $\mathbf{G}$ is *connected* is not necessary to obtain a polynomial bound on the cover time as the following lemma shows (we omit the proof due to space requirements, but it uses similar arguments to the ones in the previous theorem):

Claim 12 *Let $\mathcal{B} = (\mathbf{G}, \bar{P})$ be a Bernoulli evolving graph, $\mathbf{G}$ be the set of all maximum matching of the complete graph (any such graph is disconnected) and $\bar{P}$ is the uniform distribution over $\mathbf{G}$. The cover time of the simple random walk on $\mathcal{B}$ is the same as the cover time of the complete graph, $n \log n(1 + o(1))$.*

5.1 *d*-regular Dynamic Graphs

It is known that simple random walks on regular, connected, non-bipartite static graph have cover time of $O(n^2)$ [**?**]. Interestingly, it turns out that a similar result holds true for regular, connected, non-bipartite evolving graphs.

Theorem 13. *For any d-regular connected non-bipartite evolving graph $\mathcal{G}$ the cover time of the simple random walk on $\mathcal{G}$ is $O(d^2 n^3 \ln^2 n)$.*

We will use the following lemma from Lovász:

Lemma 14 ([?], ex. 11.26) *Let G be an undirected connected d-regular (multi)graph on n vertices. Let A_G be the adjacency matrix of G normalized by $1/d$. If $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ are the eigenvalues of A_G, then $\lambda_1 = 1$ and for $i \geq 2$, $\lambda_i \leq 1 - \frac{1}{dn^2}$. Furthermore if G is non-bipartite, then for $i \geq 2$, $\lambda_i^2 \leq 1 - \frac{1}{d^2n^2}$.*

For completeness we provide a proof in the appendix. We will need the following lemma:

Lemma 15 *Let G be an undirected d-regular graph on n vertices and $p = (p_1, \ldots, p_n)$ be a probability distribution on its vertices. Let A_G be the transition matrix of a simple random walk on G. Then:*

1. $$\left\| pA_G - \frac{\mathbb{I}}{n} \right\|_2^2 \leq \left\| p - \frac{\mathbb{I}}{n} \right\|_2^2 .$$

2. *if G is connected*
$$\left\| pA_G - \frac{\mathbb{I}}{n} \right\|_2^2 \leq \left(1 - \frac{1}{d^2n^2}\right) \left\| p - \frac{\mathbb{I}}{n} \right\|_2^2 .$$

Proof. We first prove the second claim by a standard argument. Let $\alpha_1, \alpha_2, \ldots, \alpha_n$ be an orthonormal set of eigenvectors of A_G with corresponding eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Since A_G is a symmetric stochastic matrix, $\lambda_1 = 1$ and $\alpha_1 = \frac{\mathbb{I}}{\sqrt{n}}$ and all eigenvectors and eigenvalues are real. Clearly

$$\begin{aligned} \left\| pA_G - \frac{\mathbb{I}}{n} \right\|_2^2 &= \left\| pA_G - \frac{\mathbb{I}}{n} A_G \right\|_2^2 \\ &= \left\| (p - \frac{\mathbb{I}}{n}) A_G \right\|_2^2 . \end{aligned}$$

Since $(p - \frac{\mathbb{I}}{n})$ is a vector orthogonal to $\alpha_1 = \frac{\mathbb{I}}{n}$,

$$p - \frac{\mathbb{I}}{n} = \sum_{i=2}^{n} \beta_i \alpha_i,$$

for some $\beta_2, \ldots, \beta_n \in \mathcal{R}$. As $\alpha_2, \ldots, \alpha_n$ are orthonormal, a standard calculation reveals that

$$\begin{aligned} \left\| p - \frac{\mathbb{I}}{n} \right\|_2^2 &= \left\langle p - \frac{\mathbb{I}}{n}, p - \frac{\mathbb{I}}{n} \right\rangle \\ &= \left\langle \sum_{i=2}^{n} \beta_i \alpha_i, \sum_{i=2}^{n} \beta_i \alpha_i \right\rangle \\ &= \sum_{i=2}^{n} \beta_i^2. \end{aligned}$$

Furthermore,

$$\begin{aligned} \left\| (p - \frac{\mathbb{I}}{n}) A_G \right\|_2^2 &= \left\| \sum_{i=2}^{n} \beta_i \alpha_i A_G \right\|_2^2 \\ &= \left\| \sum_{i=2}^{n} \lambda_i \beta_i \alpha_i \right\|_2^2 = \sum_{i=2}^{n} \lambda_i^2 \beta_i^2. \end{aligned}$$

By Lemma **??**, $\lambda_2^2, \ldots, \lambda_n^2 \leq \left(1 - \frac{1}{d^2 n^2}\right)$. Thus

$$\begin{aligned} \left\| (p - \frac{\mathbb{I}}{n}) A_G \right\|_2^2 &\leq \left(1 - \frac{1}{d^2 n^2}\right) \sum_{i=2}^{n} \beta_i^2 \\ &\leq \left(1 - \frac{1}{d^2 n^2}\right) \left\| p - \frac{\mathbb{I}}{n} \right\|_2^2. \end{aligned}$$

This finishes the proof of the second claim. The first claim can be proven in a similar way. One uses the bound $\lambda_1^2, \ldots, \lambda_n^2 \leq 1$ instead of $\lambda_2^2, \ldots, \lambda_n^2 \leq \left(1 - \frac{1}{d^2 n^2}\right)$.

As an immediate corollary to the previous lemma we obtain:

Corollary 16 *Let $\mathcal{G} = g_1, g_2, \ldots$ be a sequence of d-regular graphs on the same vertex set $V = \{1, \ldots, n\}$. For integers $0 \leq \ell \leq t$ let at least ℓ of the graphs $g_1, \ldots, g_t$ be non-bipartite connected. If p_0 is the initial probability distribution on V and we perform a simple random walk on $\mathcal{G}$ starting from p_0, then the probability distribution p_t of the walk after t steps satisfies:*

$$\left\| p_t - \frac{\mathbb{I}}{n} \right\|_2^2 \leq \left(1 - \frac{1}{d^2 n^2}\right)^{\ell} \left\| p_0 - \frac{\mathbb{I}}{n} \right\|_2^2.$$

A technique similar to [**?**] gives the following lemma.

Lemma 17 *Let $Y_0, Y_1, Y_2, \ldots$ be a sequence of random variables with range $V = \{1, \ldots, n\}$ satisfying for all $u, v \in V$ and $i > 0$, $\Pr[Y_i = u | Y_{i-1} = v] \geq 1/2n$. If $t = \min\{i;\ \{Y_0, Y_1, \ldots, Y_i\} = V\}$ then the expectation $E[t] \leq 3n \ln n + O(\sqrt{n} \ln n)$.*

Proof. For every $\ell > 0$ and every $v \in V$, $\Pr[v \notin \{Y_{\ell+1}, \ldots, Y_{\ell+3n \ln n}\}] < (1 - 1/2n)^{3n \ln n} < e^{-(3/2) \ln n} = n^{-3/2}$. Thus, $\Pr[\exists v \in V;\ v \notin \{Y_{\ell+1}, \ldots, Y_{\ell+3n \ln n}\}] < n \cdot n^{-3/2} = 1/\sqrt{n}$. For each integer $k \geq 0$, if we set $\ell = k \cdot 3n \ln n$ then the probability that $Y_{\ell+1}, \ldots, Y_{\ell+3n \ln n}$ does not cover whole V is at most $1/\sqrt{n}$. Thus the expected k before V is covered is at most $1/(1 - 1/\sqrt{n}) = 1 + O(1/\sqrt{n})$. Hence the expected cover time of V is bounded by $E[t] \leq 3n \ln n + O(\sqrt{n} \ln n)$.

Proof of Theorem **??**. Let $X_0, X_1, \dots$ be a random walk on $\mathcal{G}$. For an integer $i \geq 0$, define $Y_i = X_{i \cdot 4d^2n^2 \ln n}$. Pick $u, v \in V$. For $i > 1$, let p_i be the probability distribution of Y_i conditioned on $Y_{i-1} = v$. By Corollary **??**, $\left\|p_i - \frac{\mathbb{I}}{n}\right\|_2^2 \leq \left(1 - \frac{1}{d^2n^2}\right)^{4d^2n^2 \ln n} < n^{-4}$. Hence, all coordinates of the vector $(p_i - \frac{\mathbb{I}}{n})$ are in absolute value smaller than $1/n^2$. Thus $\Pr[Y_i = u | Y_{i-1} = v] \geq \frac{1}{n} - \frac{1}{n^2} \geq 1/2n$, provided that $n \geq 2$. Applying Lemma **??** yields the result. □

6 Random Walk Strategy

Consequently to the previous section the following simple strategy for the random walk guarantees that an evolving graph will be covered in expected polynomial time:

Definition 18 (Lazy Random Walk) *At each step of the walk pick a vertex v from $V(G)$ uniformly at random and if there is an edge from the current vertex to the vertex v then we move to v otherwise we stay at the current vertex.*

In effect what this strategy does is that it makes the graph n-regular; every edge adjacent to the current vertex is picked with the probability $1/n$ and with the remaining probability we use one of many self-loops. If we have an apriori upper bound $d_{\max}$ on the maximum degree of the dynamic graph we can achieve a slightly faster cover time. In that case we can reformulate the strategy as follows:

> At each step of the walk with probability $1 - (d(u)/(d_{\max} + 1))$ stay at the current vertex u and with the remaining probability pick uniformly at random one of the neighbors v of the current vertex and move to v.

We call this strategy $d_{\max}$-*lazy random walk*. If the only upper bound on the maximum degree that we have is n then this strategy becomes the previous one. We claim the following as an immediate corollary of Theorem **??**:

Theorem 19. *For any connected evolving graph $\mathcal{G}$ with maximum degree $d_{\max}$ the cover time of the $d_{\max}$-lazy random walk on $\mathcal{G}$ is $O(d_{\max}^2 n^3 \ln^2 n)$.*

Indeed these strategies do not even require the dynamic graph to be connected at each step. By Corollary **??** and Lemma **??** as long as the dynamic graph is connected for polynomial fraction of the time, the cover time of a random walk using our strategy will still be polynomial. In that case we can obtain the following generalization of Theorem **??**.

Theorem 20. *Let $\mathcal{G} = G_1, G_2, \dots$ be an evolving graph with maximum degree $d_{\max}$. Let $\epsilon > 0$ be such that for every integer ℓ, at least $\epsilon\ell$ graphs among $G_1, G_2, \dots, G_\ell$ are connected. Then the cover time of the $d_{\max}$-lazy random walk on $\mathcal{G}$ is $O(\epsilon^{-1} d_{\max}^2 n^3 \ln^2 n)$.*

The constant in the big-O is a universal constant that does not depend on $\mathcal{G}$.

7 Conclusions

In this paper we demonstrate that the cover time of the simple random walk on dynamic graphs is significantly different from the case of static graphs. While the latter was well known to be polynomial, the former is shown here to be exponential on some evolving graphs. Moreover, we show that even if the random walk takes many steps before the graph evolves the cover time can still be exponential.

We prove that in order to accelerate the cover time one can use a *lazy* random walk and reduce the cover time to polynomial. This approach has been used previously on static graphs in order to sample nodes uniformly at random, but contrary to our situation, it can be shown that it cannot accelerate the cover time for static graphs.

To summarize, the main results presented here provide theoretical justification to the wide use of random-walk-techniques in dynamic networks. Nevertheless, careful attention to the network dynamics is required when choosing the implementation of the random walk.

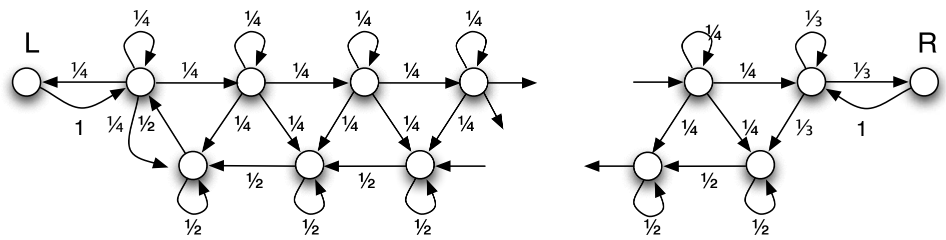

Fig. 2. Transition probabilities for the simple random walk on $\mathcal{H}_\ell$

Appendix - Proofs

Proof of Lemma **??**. We make several observations about the simple random walk on $\mathcal{H}_\ell$. If we happen to be at a vertex corresponding to a_i in Fig. **??** after step t of the walk, then with probability $1/2$ we will be at the vertex corresponding to a_{i-1} after the step $t+1$ and with the remaining probability we will again be at the (different) vertex corresponding to a_i after step $t+1$. Similarly for b_i and c_i. We summarize these transition probabilities below:

from:\ to:	a_{i-1}	b_{i-1}	c_{i-1}	a_i	b_i	c_i	a_{i+1}	b_{i+1}	c_{i+1}
a_i	1/2			1/2					
b_i	1/2			1/2					
c_i					1/4	1/4		1/4	1/4

where for the a_i and b_i in the left columns, $i = 2, \ldots, \ell$ and for c_i, $i = 1, \ldots, \ell - 1$. Furthermore:

from:\ to:	L	c_0	a_1	b_1	c_1
c_0	1/4	1/4		1/4	1/4
a_1		1/2	1/2		
b_1		1/2	1/2		

and

from:\ to:	R	c_ℓ	b_ℓ	b_1	c_1
c_ℓ	1/3	1/3	1/3		

Hence the walk on $\mathcal{H}_\ell$ is equivalent to a walk on Markov chain in Fig. **??**. It is straightforward that starting from c_0 the probability of visiting c_ℓ before returning to c_0 is at most $(1/2)^\ell$ as if we step on a_i or b_i during the first ℓ steps we will return to c_0 without visiting c_ℓ. Hence the probability of visiting R during the first ℓ returns to c_0 is at most $\ell(1/2)^\ell$. Also, the probability of not visiting L during the first ℓ returns to c_0 is at most $(3/4)^\ell$. Thus Pr[visiting R before visiting L when starting from c_0] $\leq$ Pr[visiting R during first ℓ returns to c_0] + Pr[L is not visited within first ℓ returns to c_0] $\leq \ell(1/2)^\ell + (3/4)^\ell$. □

Proof of Lemma **??**. Clearly, the eigenvector corresponding to $\lambda_1 = 1$ is $\mathbb{I}$. Consider any normalized real eigenvector $x \perp \mathbb{I}$ with its eigenvalue λ. Hence, $\sum_{i=1}^n x_i^2 = 1$ and $\sum_{i=1}^n x_i = 0$. Let x_t be the largest (in absolute value) coordinate of x. Clearly $|x_t| \geq 1/\sqrt{n}$. W.L.O.G., x_t is positive. Let $x_s < 0$ be the smallest value of x, and $s = v_1 \to v_2 \to \cdots \to v_k = t$ be the vertices on a path from s to t in G. Consider $x(\mathbb{I} - A)x^T$.

$$1 - \lambda = x(\mathbb{I} - A)x^T = \frac{1}{d} \cdot \sum_{\{i,j\} \in E(G)} (x_i - x_j)^2$$

$$\geq \frac{1}{d} \cdot \sum_{i=1}^{k-1} (x_{v_i} - x_{v_{i+1}})^2$$

$$\geq \frac{1}{d(k-1)} \cdot \left(\sum_{i=1}^{k-1} x_{v_i} - x_{v_{i+1}} \right)^2$$

$$
\begin{aligned}
&= \frac{1}{d(k-1)} \cdot (x_s - x_t)^2 \\
&\geq \frac{1}{dn^2},
\end{aligned}
$$

where the second inequality follows from Cauchy-Schwarz bound. If G is connected and non-bipartite then A_G^2 corresponds to a connected d^2-regular graph on n vertices. Since A_G^2 has eigenvalues $\lambda_1^2, \lambda_2^2, \ldots, \lambda_n^2$, the second part of the lemma follows. □