ABSTRACT: We study randomized gossip-based processes in dynamic networks that are motivated by information discovery in large-scale distributed networks such as peer-to-peer and social networks. A well-studied problem in peer-to-peer networks is resource discovery, where the goal for nodes (hosts with IP addresses) is to discover the IP addresses of all other hosts. Also, some of the recent work on self-stabilization algorithms for P2P/overlay networks proceed via discovery of the complete network. In social networks, nodes (people) discover new nodes through exchanging contacts with their neighbors (friends). In both cases the discovery of new nodes changes the underlying network — new edges are added to the network — and the process continues in the changed network. Rigorously analyzing such dynamic (stochastic) processes in a continuously changing topology remains a challenging problem with obvious applications.
This paper studies and analyzes two natural gossip-based discovery processes. In the push discovery or triangulation process, each node repeatedly chooses two random neighbors and connects them (i.e., “pushes” their mutual information to each other). In the pull discovery process or the two-hop walk, each node repeatedly requests or “pulls” a random contact from a random neighbor and connects itself to this two-hop neighbor. Both processes are lightweight in the sense that the amortized work done per node is constant per round, local, and naturally robust due to the inherent randomized nature of gossip.

Our main result is an almost-tight analysis of the time taken for these two randomized processes to converge. We show that in any undirected $n$-node graph both processes take $O(n \log^2 n)$ rounds to connect every node to all other nodes with high probability, whereas $\Omega(n \log n)$ is a lower bound. We also study the two-hop walk in directed graphs, and show that it takes $O(n^3 \log n)$ time with high probability, and that the worst-case bound is tight for arbitrary directed graphs, whereas $\Omega(n^2)$ is a lower bound for strongly connected directed graphs. A key technical challenge that we overcome in our work is the analysis of a randomized process that itself results in a constantly changing network leading to complicated dependencies in every round. We discuss implications of our results and their analysis to discovery problems in P2P networks as well as to evolution in social networks.

1. INTRODUCTION

Many large-scale, real-world networks such as peer-to-peer networks, the Web, and social networks are highly dynamic with continuously changing topologies. The evolution of the network as a whole is typically determined by the decentralized behavior of nodes, i.e., the local topological changes made by the individual nodes (e.g., adding edges between neighbors). Understanding the dynamics of such local processes is critical for both analyzing the underlying stochastic phenomena, e.g., in the emergence of structures in social networks, the Web and other real-world networks [10,36,38], and for designing practical algorithms for associated algorithmic problems, e.g., in resource discovery in distributed networks [25,33] or in the analysis of algorithms for the Web [12,18]. In this paper, we study the dynamics of network evolution that result from local gossip-style processes. Gossip-based processes have recently received significant attention because of their simplicity of implementation, scalability to large network size, and robustness to frequent network topology changes; see, e.g., [8,11,13,19,27–31,35] and the references therein. In particular, gossip-based protocols have been used to efficiently and robustly construct various overlay topologies dynamically in a fully decentralized manner [27]. In a local gossip-based algorithm (e.g., [13]), each node exchanges information with a small number of randomly chosen neighbors in each round.1 The randomness inherent in the gossip-based protocols naturally provides robustness, simplicity, and scalability. While many of the theoretical gossip-based work (including those on rumor spreading), especially, the push-pull type algorithms [13,14,22,

---

1Gossip, in some contexts (see e.g., [28, 29]), has been used to denote communication with a random node in the network, as opposed to only a directly connected neighbor. The former model essentially assumes that the underlying graph is complete, whereas the latter (as assumed here) is more general and applies even to arbitrary graphs. The local gossip process is typically more difficult to analyze due to the dependences that arise as the network evolves.
23, 28, 29] focus on analyzing various gossip-based tasks (e.g., computing aggregates or spreading a rumor) on static graphs, a key feature of this work is rigorously analyzing a gossip-based process in a dynamically changing graph, where the graph topology changes in a probabilistic fashion due to the behavior of the gossip-style process itself. We also note that there have been recent works (see e.g., [4–7, 15–17, 24, 37] and the references therein) that have analyzed gossip-style processes, random walks, and flooding processes in dynamic networks where the topological changes are determined by an adversary (usually, but not always, independent of the processes that occur over these networks).

We present two illustrative application domains for our study. First, consider a P2P network, where nodes (computers or end-hosts with IDs/IP addresses) can communicate only with nodes whose IP address are known to them. A basic building block of such a dynamic distributed network is to efficiently discover the IP addresses of all nodes that currently exist in the network. This task, called resource discovery [25], is a vital mechanism in a dynamic distributed network with many applications [1, 25]: when many nodes in the system want to interact and cooperate they need a mechanism to discover the existence of one another. Resource discovery is typically done using a local mechanism [25]; in each round nodes discover other nodes and this changes the resulting network—new edges are added between the nodes that discovered each other. As the process proceeds, the graph becomes denser and denser and will finally result in a complete graph. Such a process was first studied in [25] which showed that a simple randomized process is enough to guarantee almost-optimal time bounds for the time taken for the entire graph to become complete (i.e., for all nodes to discover all other nodes). Their randomized Name Dropper algorithm operates as follows: in each round, each node chooses a random neighbor and sends all the IP addresses it knows. Note that while this process is also gossip-based the information sent by a node to its neighbor can be extremely large (i.e., of size $\Omega(n)$).

More recently, self-stabilization protocols have been designed for constructing and maintaining P2P overlay networks e.g, [9, 26]. These protocols guarantee convergence to a desired overlay topology (e.g., the SKIP+ graph) starting from any arbitrary topology via local checking and repair. For example, the self-stabilizing protocol of [9] proceeds by continuously discovering new neighbors (via transitive closure) till a complete graph is formed. Then the repair process is initiated. This can also be considered as a local gossip-based process in an underlying virtual graph with changing (added) edges. In both the above examples, the assumption is that the starting graph is arbitrary but (at least) weakly connected. The gossip-based processes that we study also have the same goal—starting from an arbitrary connected graph, each node discovers all nodes as quickly as possible—in a setting where individual message sizes are small ($O(\log n)$ bits).

Second, in social networks, nodes (people) discover new nodes through exchanging contacts with their neighbors (friends). Discovery of new nodes changes the underlying network—new edges are added to the network—and the process continues in the changed network. For example, consider the LinkedIn network, a large social network of professionals on the Web. The nodes of the network represent people and edges are added between people who directly know each other (i.e., between direct contacts). Edges are generally undirected, but LinkedIn also allows directed edges, where only one node is in the contact list of another node. LinkedIn allows two mechanisms to discover new contacts. The first can be thought of as a triangulation process (see Fig. 1a): A person can introduce two of

2http://www.linkedin.com.
his friends that could benefit from knowing each other—he can mutually introduce them by giving their contacts. The second can be thought of as a two-hop process (see Fig. 1b): If you want to acquire a new contact then you can use a shared (mutual) neighbor to introduce yourself to this contact; i.e., the new contact has to be a two-hop neighbor of yours. Both the processes can be modeled via gossip in a natural way (as we do shortly below) and the resulting evolution of the network can be studied: e.g., how and when do clusters emerge? how does the diameter change with time? In the social network context, our study focuses on the following question: how long does it take for all the nodes in a connected induced subgraph of the network to discover all the nodes in the subgraph? This is useful in scenarios where members of a social group, e.g., alumni of a school, members of a club, discover all members of the group through local gossip operations.

**Gossip-based discovery.** Motivated directly by the above applications, we analyze two lightweight, randomized gossip-based discovery processes. We assume that we start with an arbitrary undirected connected graph and the process proceeds in **synchronous rounds**. Communication among nodes occurs only through edges in the network. We further assume that the size of each message sent by a node in a round is at most $O(\log n)$ bits, i.e., the size of an ID.

1. **Push discovery (triangulation):** In each round, each node chooses two random neighbors and connects them by “pushing” their mutual information to each other. In other words, each node adds an undirected edge between two of its random neighbors; if the two neighbors are already connected, then this does not create any new edge. Note that this process, which is illustrated in Fig. 1a, is completely local. To execute the process, a node only needs to know its neighbors; in particular, no two-hop information is needed. (Note that the selection of edges to add for each node is simultaneous at each round.)

2. **Pull discovery (two-hop walk):** In each round, each node connects itself to a random neighbor of a neighbor chosen uniformly at random, by “pulling” a random neighboring ID from a random neighbor. Alternatively, one can think of each node doing a two-hop random walk and connecting to its destination. This process, illustrated in Fig. 1b, can also be executed locally: a node (say $u$) simply asks one of its neighbors $v$ for an ID of one of $v$’s neighbors and then adds an undirected edge to the received contact (say $w$). (This can be thought of as a two-step process: first, $u$ learns of $w$’s address from $v$; second, $u$ also notifies $w$ of its address.)

*Random Structures and Algorithms* DOI 10.1002/rsa

---

**Fig. 1.** (a) Push discovery or triangulation process. (b) Pull discovery or two-hop walk process. (c) Non-monotonicity of the triangulation (push) process—the expected convergence time for the 4-edge graph exceeds that for the 3-edge subgraph. The left graph makes progress towards the complete graph faster than the right graph. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]
Both the above processes are local in the sense that each node only communicates with its neighbors (or its two-hop neighbors, as in the case of pull discovery) in any round, and lightweight in the sense that the average work done per node is only a constant per round. Both processes are also easy to implement and generally oblivious to the current topology structure, changes, or failures. It is interesting also to consider variants of the above processes in directed graphs. In particular, we study the two-hop walk process which naturally generalizes in directed graphs: each node does a two-hop directed random walk and adds a directed edge to its destination. We are mainly interested in the time taken by the process to converge to the transitive closure of the initial graph, i.e., till no more new edges can be added.

Our results. Our main contribution is an analysis of the above gossip-based discovery processes in both undirected and directed graphs. In particular, we show the following results (the precise theorem statements are in the respective sections.)

- **Undirected graphs:** In Sections 3 and 4, we show that for any undirected $n$-node graph, both the push and the pull discovery processes converge to the transitive closure of the graph in $O(n \log^2 n)$ rounds with high probability. We also show that $\Omega(n \log n)$ is a lower bound on the number of rounds needed, with high probability, for any $n$-node connected graph that is missing $\Omega(n)$ edges. Hence our analysis is tight within a logarithmic factor.

  Our results also apply when we require only a subset of nodes to converge. In particular, consider a subset of $k$ nodes that induce a connected subgraph and run the gossip-based process restricted to this subgraph. Then by just applying our results to this subgraph, we immediately obtain that it will take $O(k \log^2 k)$ rounds, with high probability (in terms of $k$), for all the nodes in the subset to converge to a complete subgraph. As discussed above, such a result is applicable in social network scenarios where all nodes in a subset of network nodes discover one another through gossip-based processes.

- **Directed graphs:** In Section 5, we show that the pull process takes $O(n^2 \log n)$ time for any $n$-node directed graph, with high probability. We show a matching lower bound for weakly connected graphs, and an $\Omega(n^2)$ lower bound for strongly connected directed graphs. Our analysis indicates that the directionality of edges can greatly impede the resource discovery process.

Applications. The gossip-based discovery processes we study are directly motivated by the two scenarios outlined above, namely algorithms for resource discovery in distributed networks and analyzing how discovery process affects the evolution of social networks. Since our processes are simple, lightweight, and easy to implement, they can be used for resource discovery in distributed networks. The Name Dropper discovery algorithm has been applied to content delivery systems [25]. As mentioned earlier, Name Dropper and other prior algorithms for the discovery problem [1, 25, 32, 33] complete in polylogarithmic number of rounds ($O(\log^2 n)$ or $O(\log n)$), but may transfer $\Theta(n)$ bits per edge per round. As a

3One can also study the push process in directed graphs: each node chooses two random outgoing neighbours and adds a directed edge in one or both directions. It will be interesting to analyze this process and compare with the corresponding results of the push process on undirected graphs.

Random Structures and Algorithms DOI 10.1002/rsa
result, they may not be scalable for bandwidth and resource-constrained networks (e.g., peer-to-peer, mobile, or sensor networks). One approach to use these algorithms in a bandwidth-limited setting ($O(\log n)$-bits per message) is to spread the transfer of long messages over a linear number of rounds, but this requires coordination and maintaining state. In contrast, the “stateless” nature of the gossip processes we study and the fact that the results apply to any initial graph make the process attractive in unpredictable environments. Our analyses can also give insight into the growth of social networks that grow in a decentralized way by the local actions of the individual nodes. Although, convergence to the complete graph as a whole is unrealistic from the point of large social networks, as mentioned earlier, it might be more relevant to discovering members in a (smaller) subgraph. In addition to the application of discovering all members of a group, analyses of the processes such as the ones we study can help analyze both short-term and long-term evolution of social networks. In particular, it can help in predicting the sizes of the immediate neighbors as well as the sizes of the second and third-degree neighbors (these are listed for every node in LinkedIn). An estimate of these can help in designing efficient algorithms and data structures to search and navigate the social network.

**Technical contributions.** Our main technical contribution is a probabilistic analysis of localized gossip-based discovery in arbitrary networks. While our processes can be viewed as graph-based coupon collection processes, one significant distinction with past work in this area [2, 3, 20] is that the graphs in our processes are constantly changing. The dynamics and locality inherent in our process introduces nontrivial dependencies, which makes it difficult to characterize the network as it evolves.

A further challenge is posed by the fact that the expected convergence time for the two processes is not monotonic; that is, the processes may take longer to converge starting from a graph $G$ than starting from a subgraph $H$ of $G$. Figure 1c presents a small example illustrating this phenomenon. This seemingly counterintuitive phenomenon is, however, not surprising considering the fact that the cover time of random walks also share a similar property. One consequence of these hurdles is that analyzing the convergence time for even highly specialized or regular graphs is challenging since the probability distributions of the intermediate graphs are hard to specify. Our lower bound analysis for a specific strongly connected directed graph in Theorem 15 illustrates some of the challenges. In our main upper bound results (Theorems 3 and 12), we overcome these technical difficulties by presenting a uniform analysis for all graphs, in which we study different local neighborhood structures and show how each leads to rapid growth in the minimum degree of the graph.

2. **PRELIMINARIES**

In this section, we define the notations used in our proofs, and prove some common lemmas for Section 3 and Section 4. Let $G$ denote a connected graph, $d(u)$ denote the degree of node $u$, and $N^i(u)$ denote the set of nodes that are at distance exactly $i$ from $u$. Let $\delta$ denote the minimum degree of $G$. We note that $G$, $d(u)$, and $N^i(u)$ all change with time, and are, in fact, random variables. For any nonnegative integer $t$, we use subscript $t$ to denote the random variable at the start of round $t$; for example $G_t$ refers to the graph at the start of round $t$. For convenience, we list the notations in Table 1.

We state two lemmas that are used in the proofs in Section 3 and Section 4. Lemma 1 gives a lower bound on the number of neighbors within distance 4 for any node $u$ in $G_t$.
TABLE 1. Notation table

| Notation     | Description                                      |
|--------------|--------------------------------------------------|
| \( \delta_t \) | Minimum degree of graph \( G_t \)               |
| \( N_t^i (u) \) | Set of nodes that are at distance exactly \( i \) from \( u \) in \( G_t \) |
| \(| N_t^i (u) | Number of nodes in \( N_t^i (u) \)             |
| \( d_t (u) \) | Degree of node \( u \) in \( G_t \)            |
| \( d_t (u, S) \) | Number of edges from \( u \) to nodes in \( S \), i.e., degree induced on \( S \) |

while Lemma 2 is a standard analysis of a sequence of Bernoulli experiments and can be proved by a direct coupon collector argument or using a Chernoff bound.

**Lemma 1.** \[ \left| \bigcup_{i=1}^{4} N_t^i (u) \right| \geq \min \{2\delta_t, n-1\} \] for all \( u \) in \( G_t \).

**Proof.** If \( N_t^3 (u) \) is not an empty set, consider a node \( v \in N_t^3 (u) \). Since \( d_t (v) \geq \delta_t \), we have that \[ \left| \bigcup_{i=2}^{4} N_t^i (u) \right| \geq \delta_t \]. It furthermore holds that \[ |N_t^1 (u) | \geq \delta_t \] because \( d_t (u) \geq \delta_t \). We also know \( N_t^1 (u) \) and \( \bigcup_{i=2}^{4} N_t^i (u) \) are disjoint. Thus, \[ \left| \bigcup_{i=1}^{4} N_t^i (u) \right| \geq 2\delta_t \].

If, on the other hand, \( N_t^3 (u) \) is an empty set, then \( N_t^1 (u) \cup N_t^2 (u) = n-1 \) because \( G_t \) is connected. Thus \( \left| \bigcup_{i=1}^{4} N_t^i (u) \right| = n-1 \). Combining the the above two cases completes the proof.

**Lemma 2.** Consider \( k \) Bernoulli experiments, in which the success probability of the \( i \)th experiment is at least \( i/m \) where \( m \geq k \). If \( X_i \) denotes the number of trials needed for experiment \( i \) to output a success and \( X = \sum_{i=1}^{k} X_i \), then \( \Pr \{ X > (c+1)m \ln m \} \) is less than \( 1/m^c \).

**Proof.** Since \( X \) only increases with \( k \), without loss of generality assume that \( k = m \). Now we can view this as a coupon collector problem [34] where \( X_{m+1:m} \) is the number of steps to collect the \( i \)th coupon. Consider the probability of not obtaining the \( i \)th coupon after \( (c+1)m \ln m \) steps. This probability is

\[
\left( 1 - \frac{1}{m} \right)^{(c+1)m \ln m} < e^{-(c+1) \ln m} = \frac{1}{m^{c+1}}
\]

By union bound, the probability that some coupon has not been collected after \( (c+1)m \ln m \) steps is less than \( 1/m^c \). And this completes the proof of this lemma.

**3. PROOFS FOR THE TRIANGULATION PROCESS**

In this section, we analyze the triangulation process on undirected connected graphs, which is described by the following simple iteration: In each round, for each node \( u \), we add edge \((v, w)\) where \( v \) and \( w \) are drawn uniformly at random from \( N_t^1 (u) \). (If \( v = w \), then no new edge is added.) The triangulation process yields the following push-based resource discovery protocol. In each round, each node \( u \) introduces two random neighbors \( v \) and \( w \) to one another. The main result of this section is the following.

*Random Structures and Algorithms* DOI 10.1002/rsa
Theorem 3 (Upper bound for triangulation process). For any connected undirected graph, the triangulation process converges to a complete graph in \(O(n \log^2 n)\) rounds with high probability.

We also establish an \(\Omega(n \log n)\) lower bound on the triangulation process for almost all \(n\)-node graphs.

Theorem 4 (Lower bound for triangulation process). For any connected undirected graph \(G\) that has \(k \geq 1\) edges less than \(\delta_0/2\) edges to \(S\) (i.e., \(d_i(v, S) < \delta_0/2\)), and strongly tied to \(S\) if \(v\) has at least \(\delta_0/2\) edges to \(S\) (i.e., \(d_i(v, S) \geq \delta_0/2\)). (Recall that \(\delta_0\) is the minimum degree at start of round 0.)

### 3.1. Upper Bound

We obtain the \(O(n \log^2 n)\) upper bound of Theorem 3 by proving that the minimum degree of the graph increases by a constant factor (or equals \(n - 1\)) in \(O(n \log n)\) steps. Towards this objective, we study how the neighbors of a given node connect to the two-hop neighbors of the node. We say that a node \(v\) is weakly tied to a set of nodes \(S\) if \(v\) has less than \(\delta_0/2\) edges to \(S\) (i.e., \(d_i(v, S) < \delta_0/2\)), and strongly tied to \(S\) if \(v\) has at least \(\delta_0/2\) edges to \(S\) (i.e., \(d_i(v, S) \geq \delta_0/2\)).

Lemma 5. If \(d_i(u) < \min\{n - 1, (1 + 1/4)\delta_0\}\) and \(w \in N_i^1(u)\) has at least \(\delta_0/4\) edges to \(N_i^2(u)\), then the probability that \(u\) connects to a node in \(N_i^2(u)\) through \(w\) in round \(t\) is at least \(1/(6n)\).

Proof. The probability that \(u\) connects to a node in \(N_i^2(u)\) through \(w\) in round \(t\) equals

\[
\frac{d_i(w, N_i^2(u))}{d_i(w)} \cdot \frac{1}{n} \geq \frac{d_i(w, N_i^2(u))}{d_i(w)} \cdot \frac{1}{n} \geq \frac{|N_i^1(u)| + d_i(w, N_i^2(u))}{n} \cdot \frac{\delta_0/4}{(1 + 1/4)\delta_0 + \delta_0/2} \cdot \frac{1}{n} = \frac{1}{6n}.
\]

Lemma 6. If \(d_i(u) < \min\{n - 1, (1 + 1/4)\delta_0\}\), \(w \in N_i^1(u)\) is weakly tied to \(N_i^2(u)\), and \(v \in N_0^2(u) \cap N_i^1(w)\), then the probability that \(u\) connects to \(v\) through \(w\) in round \(t\) is at least \(1/(4\delta_0^2)\).

Proof. Since \(w\) is weakly tied to \(N_i^2(u)\) and \(d_i(w)\) is at most \(|N_i^1(u)| + d_i(w, N_i^2(u))\), we obtain that \(d_i(w)\) is at most \((1 + 1/4)\delta_0 + \delta_0/2\). Therefore, the probability that \(u\) connects to \(v\) through \(w\) in round \(t\) equals

\[
\frac{1}{d_i(w)^2} \geq \frac{1}{((1 + 1/4)\delta_0 + \delta_0/2)^2} \geq \frac{1}{(7\delta_0/4)^2} \geq \frac{1}{4\delta_0^2}.
\]

For analyzing the growth in the degree of a node \(u\), we consider two overlapping cases. The first case is when more than \(\delta_0/4\) nodes of \(N_i^1(u)\) are strongly tied to \(N_i^2(u)\), and the second is when less than \(\delta_0/3\) nodes of \(N_i^1(u)\) are strongly tied to \(N_i^2(u)\). Figure 2 illustrates the different cases and lemmas used in the proof of Theorem 3.

Random Structures and Algorithms DOI 10.1002/rsa
The analysis for the first case is relatively straightforward: when several neighbors of a node \( u \) are strongly tied to \( u \)'s two-hop neighbors, then their triangulation steps connect \( u \) to a large fraction of these two-hop neighbors. The precondition in the lemma below is actually weaker than the requirement that several neighbors are strongly tied to the two-hop neighbors; this stronger version of the lemma is helpful in the proof of Theorem 3.

**Lemma 7 (When several neighbors are strongly tied to two-hop neighbors).** There exists a \( T = O(n \log n) \) such that if more than \( \delta_0/4 \) nodes in \( N^1_t(u) \) each have at least \( \delta_0/4 \) edges to \( N^2_t(u) \) for all \( t < T \), then with probability at least \( 1 - 1/n^2 \) it holds that \( d_T(u) \geq \min\{n - 1, (1 + 1/4) \delta_0\} \).

**Proof.** If at any round \( t < T \), \( d_t(u) \geq \min\{n - 1, (1 + 1/4) \delta_0\} \), then the claim of the lemma holds. In the remainder of this proof, we assume \( d_t(u) < \min\{n - 1, (1 + 1/4) \delta_0\} \) for all \( t < T \). Let \( w \in N^1_t(u) \) be a node that has at least \( \delta_0/4 \) edges to \( N^2_t(u) \). By Lemma 5 we know that

\[
\Pr\left[u \text{ connects to a node in } N^2_t(u) \text{ through } w \text{ in round } t\right] \geq \frac{1}{6n}.
\]

We have more than \( \delta_0/4 \) such \( w \)'s in \( N^1_t(u) \), each of which independently executes a triangulation step in any given round. Consider a run of \( T_1 = 72n \ln n/\delta_0 \) rounds. This implies at least \( 18n \ln n \) chances to add an edge between \( u \) and a node in \( N^2_t(u) \). Thus,

\[
\Pr\left[u \text{ connects to a node in } N^2_t(u) \text{ after } T_1 \text{ rounds}\right] \geq 1 - \left(1 - \frac{1}{6n}\right)^{18n \ln n} \\
\geq 1 - e^{-3 \ln n} = 1 - \frac{1}{n^3}.
\]

If a node that is two hops away from \( u \) becomes a neighbor of \( u \) by round \( t \), it is no longer in \( N^2_t(u) \). Therefore, in \( T = T_1 \delta_0/4 = O(n \log n) \) rounds, \( u \) will connect to at least \( \delta_0/4 \) new nodes with probability at least \( 1 - 1/n^2 \), i.e., \( d_T(u) \geq (1 + 1/4) \delta_0 \).

*Random Structures and Algorithms* DOI 10.1002/rsa
We next consider the second case where less than $\delta_0/3$ neighbors of a given node $u$ are strongly tied to the two-hop neighborhood of $u$. This case is more challenging since the neighbors of $u$ that are weakly tied may not contribute many new edges to $u$. We break the analysis of this part into two subcases based on whether there is at least one neighbor of $u$ that is strongly tied to $N_0^2(u)$.

**Lemma 8** (When few neighbors are strongly tied to two-hop neighbors). There exists $T = O(n \log n)$ such that if less than $\delta_0/3$ nodes in $N_1^1(u)$ are strongly tied to $N_0^2(u)$ for all $t < T$, and there exists a node $v_0 \in N_0^1(u)$ that is strongly tied to $N_0^2(u)$, then $d_T(u) \geq \min\{n - 1, (1 + 1/8) \delta_0\}$ with probability at least $1 - 1/n^2$.

**Proof.** In the following, $T$, $T_1$, and $T_2$ are integers, each bounded by $O(n \log n)$, whose exact values are set at various points in the proof to satisfy desired conditions.

If at any point $t < T$, $d_t(u) \geq \min\{n - 1, (1 + 1/8) \delta_0\}$, then the claim of the lemma holds. In the remainder of this proof, we assume $d_t(u) < \min\{n - 1, (1 + 1/8) \delta_0\}$ for all $t < T$. Let $S_0^t$ denote the set of $v_0$'s neighbors in $N_0^2(u)$ which are strongly tied to $N_1^1(u)$ at round $t$. Let $W_0^t$ denote the set of $v_0$'s neighbors in $N_0^2(u)$ which are weakly tied to $N_1^1(u)$ at round $t$.

Consider any node $v$ in $S_0^t$. Less than $\delta_0/3$ nodes in $N_1^1(u)$ are strongly tied to $N_0^2(u)$, thus more than $\delta_0/2 - \delta_0/3 = \delta_0/6$ neighbors of $v$ in $N_1^1(u)$ are weakly tied to $N_0^2(u)$. Let $w$ be one such weakly tied node. By Lemma 6, the probability that $u$ connects to $v$ through $w$ in round $t$ is at least $1/(4\delta_0^3)$. We have at least $\delta_0/6$ choices for $w$, each of which executes a triangulation step each round. Consider $T_1 = 72\delta_0 \ln n$ rounds of the process. Then the probability that $u$ connects to $v$ in $T_1$ rounds is at least

$$1 - \left(1 - \frac{1}{4\delta_0^3}\right)^{12\delta_0^2 \ln n} \geq 1 - e^{-3\ln n} = 1 - \frac{1}{n^3}.$$ 

Thus, if there exists $t < T_2$ such that $|S_0^t| \geq \delta_0/8$, then in $T_1 + T_2$ rounds, $d_{T_1+T_2}(u) \geq (1 + 1/8)\delta_0$ with probability at least $1 - 1/n^2$. This implies the claim of the lemma if we set $T \geq T_1 + T_2$.

Therefore, in the remainder of the proof we consider the case where $|S_0^t| < \delta_0/8$ for all $t \leq T_2$. Define $R_0^t = R_{t-1}^0 \cup W_0^t$, $R_0^0 = W_0^0$. If at least $\delta_0/8$ nodes in $R_0^t$ are connected to $u$ at any time $t \leq T_2$, then the claim of the lemma holds. Thus, in the following we consider the case where $|R_0^t \cap N_1^1(u)| < \delta_0/8$ for all $t \leq T_2$. Consider any round $t \leq T_2$. From the definition of $R_0^t$, we can derive

$$|R_0^t| \geq |W_0^t| = d_t(v_0, N_1^1(u)) - |S_0^t| \geq d_t(v_0, N_0^2(u)) - \delta_0/8.$$ 

At round 0, $v_0$ is strongly tied to $N_0^2(u)$, i.e., $d_0(v_0, N_0^2(u)) \geq \delta_0/2$. Since $\delta_0 \leq d_t(u) < (1 + 1/8)\delta_0$, we have

$$d_t(v_0, N_0^2(u)) \geq d_t(v_0, N_0^2(u)) - \delta_0/8 \geq 3\delta_0/8.$$
Let $e_1$ denote the event \( \{ u \text{ connects to a node in } R_0^1 \setminus N_1^1(u) \text{ through } v_0 \text{ in round } t \} \).

\[
\Pr [e_1] = \frac{|R_0^1 \setminus N_1^1(u)|}{d_i(v_0)} \cdot \frac{1}{d_i(v_0)} = \frac{|R_0^0| - |R_0^0 \cap N_1^1(u)|}{d_i(v_0)} \cdot \frac{1}{d_i(v_0)} \\
\geq \frac{|R_0^1| - |R_0^0 \cap N_1^1(u)|}{d_i(v_0)} \cdot \frac{1}{d_i(v_0)} \\
\geq \frac{|R_0^1| - \delta_0/8}{|N_1^1(u)| + d_i(v_0, N_1^2(u))} \cdot \frac{1}{n} \\
\geq \frac{3\delta_0/8 - \delta_0/8 - \delta_0/8}{(1 + 1/8)\delta_0 + 3\delta_0/8} \cdot \frac{1}{n} = \frac{1}{12n}.
\]

Let $X_1$ be the number of rounds it takes for $e_1$ to occur. When $e_1$ occurs, let $v_1$ denote a witness for $e_1$; i.e., if we use $X_1$ to denote the round at which $e_1$ occurs, then let $v_1$ denote a node in $R_0^1 \setminus N_1^1(u)$ to which $u$ connects through $v_0$ in round $X_1$. Since $v_1$ is in $R_0^1$, it is also in $W_0^1$ for some $t_1 \leq X_1$; therefore, $v_1$ is weakly tied to $N_1^1(u)$ and strongly tied to $N_1^2(u) \cup N_1^3(u)$ at the start of round $t_1$. Thus, at the start of round $t_1$, $v_1$ has at least $\delta_0/2$ neighbors that are not neighbors of $u$. If $d_i(v_1, N_1^2(u)) < 3\delta_0/8$ for any round $t$ such that $X_1 \leq t \leq T_2$, then at least $\delta_0/2 - 3\delta_0/8 = \delta_0/8$ neighbors of $v_1$ that were not neighbors of $u$ at the start of round $t_1$ became neighbors of $u$ by the start of round $t$; this would imply that $d_i(u) \geq (1 + 1/8)\delta_0$, violating our assumption at the start of the proof. (Note that we will set $T_2$ large enough so that $X_1$ is at most $T_2$ with high probability.)

Therefore, in the remainder of the proof, we consider the case where $d_i(v_1, N_1^2(u)) \geq 3\delta_0/8$, for all $X_1 \leq t \leq T_2$. Let $S_t^1$ (resp., $W_t^1$) denote the set of $v_1$’s neighbors in $N_1^2(u)$ that are strongly tied (resp., weakly tied) to $N_1^1(u)$. If $|S_t^1| \geq \delta_0/8$ for any $t \leq T_2$, then as we did for the case $|S_t^2| \geq \delta_0/8$, we argue that in at most $T_1 + T_2$ rounds, the degree of $u$ is at least $(1 + 1/8)\delta_0$ with probability at least $1 - 1/n^2$.

Thus, in the remainder, we assume that $|S_t^1| < \delta_0/8$ for all $t \leq T_2$. Consider a round $t$ such that $X_1 \leq t \leq T_2$. Define $R_t^1 = R_{t-1}^1 \cup W_t^1$, $R_t^1 = W_t^1$. Let $e_2$ denote the event

\[
\{ u \text{ connects to a node in } R_t^1 \setminus N_1^1(u) \text{ or } R_t^1 \setminus N_1^1(u) \text{ through } v_0 \text{ or } v_1 \text{ in round } t \}.
\]

By the same calculation as for $v_0$, we have $\Pr [e_2] \geq 2/(12n)$, as long as $T_2 \geq X_1 + X_2$. Similarly, we define $e_3, \ldots, e_{[\delta_0/8]}$ and $X_3, \ldots, X_{[\delta_0/8]}$, and obtain that $\Pr[e_i] \geq i/(12n)$ for $1 \leq i \leq [\delta_0/8]$, as long as $T_2 \geq \sum_{1 \leq j \leq i} X_j$. The total number of rounds for $u$ to gain at least $\delta_0/8$ nodes as neighbors is given by $\sum_{1 \leq j \leq [\delta_0/8]} X_j$, which, according to Lemma 2, is bounded by $36n \ln n$ with probability at least $1 - 1/n^2$. So, setting $T_2 = 36n \ln n$ and $T = T_1 + T_2$ completes the proof of the lemma.

**Lemma 9** (When all neighbors are weakly tied to two-hop neighbors). There exists a $T = O(n \log n)$ such that if all nodes in $N_1^1(u)$ are weakly tied to $N_2^1(u)$ for all $t < T$, then with probability at least $1 - 1/n^2$ it holds that $d_T(u) \geq \min \{(1 + 1/8)\delta_0, n - 1\}$.

**Proof.** If at any point $t < T$, $d_i(u) \geq \min \{(1 + 1/8)\delta_0, n - 1\}$, then the claim of this lemma holds. In the remainder of this proof, we assume $d_i(u) < \min \{(1 + 1/8)\delta_0, n - 1\}$ for all $t < T$. In the following, we first show that any node $v \in N_2^1(u)$ will have at least $\delta_0/4$ edges to $N_1^1(u)$, where $T_1 = O(n \log n)$. After that, $v$ will connect to $u$ in
$T_2 = O(n \log n)$ rounds. Therefore, the total number of rounds used for $v$ to connect to $u$ is $T_3 = T_1 + T_2 = O(n \log n)$.

Node $v$ at least connects to one node in $N^0_0 (u)$. Call it $w_1$. Because all nodes in $N^1_i (u)$ are weakly tied to $N^2_i (u)$, we have $d_i (w_1, N^1_i (u)) \geq \delta_0 - \delta_0/2 = \delta_0/2$. If $d_i (w_1, N^1_i (u)) \cap N^1_i (v)) < \delta_0/4$, then $v$ already has $\delta_0/4$ edges to $N^1_i (u)$.

Thus, in the following we treat the case where $d_i (w_1, N^1_i (u)) \cap N^1_i (v)) \geq \delta_0/4$. We consider the event $e_1 = \{v \text{ connects to a node in } N^1_i (u) \cap N^1_i (v) \text{ through } w_1\}$ and obtain for its probability:

$$
\Pr [e_1] = \frac{d_i (w_1, N^1_i (u)) \cap N^1_i (v))}{d_i (w_1)} \cdot \frac{1}{d_i (w_1)} \geq \frac{d_i (w_1, N^1_i (u) \cap N^1_i (v))}{N^1_i (u)} + d_i (w_1, N^2_i (u)) \cdot \frac{1}{d_i (w_1)} \geq \frac{\delta_0/4}{(1 + 1/8)\delta_0 + \delta_0/2} \cdot \frac{2}{13} \cdot \frac{1}{n} > \frac{1}{7n}.
$$

Let $X_1$ be the number of rounds needed for $e_1$ to occur. When $e_1$ occurs, let $w_2$ denote a witness for $e_1$; i.e., let $w_2$ denote a vertex in $N^1_i (u) \cap N^1_i (v)$ to which $v$ connects. Note that here the value of $t$ is the round at which the event occurs. By our choice, $w_2$ is also weakly tied to $N^2_i (u)$. By an argument similar to the one in the above paragraph, we have $d_i (w_2, N^1_i (u) \cap N^1_i (v)) \geq \delta_0/4$. Let $e_2$ denote the event $\{v \text{ connects to a node in } N^1_i (u) \text{ through } w_1 \text{ or } w_2\}$. We have $\Pr [e_2] \geq 2/(7n)$. Let $X_2$ be the number of rounds needed for $e_2$ to occur. Similarly, we can define $e_3, X_3, \ldots, e_{8\delta_0/4}, X_{8\delta_0/4}$ and show $\Pr [e_i] \geq i/(7n)$. Set $T_1 = \sum X_i$, which is the bound on the number of rounds needed for $v$ to have at least $\delta_0/4$ neighbors in $N^1_i (u)$. By Lemma 2, $T_1 \leq 28n \ln n$ with probability at least $1 - 1/n^2$. Now we show $v$ will connect to $u$ in $T_2$ rounds after this. Notice that, all $w_i$’s are still weakly tied to $N^2_i (u)$. By Lemma 6, the probability that $u$ connects to $v$ through $w_i$ in round $t$ is at least $1/(4\delta_0^2)$. We have $w_1, w_2, \ldots, w_{8\delta_0/4}$ independently executing a triangulation step each round. Consider $T_2 = 48\delta_0 \ln n$ rounds of the process. Then,

$$
\Pr [u \text{ connects to } v \text{ in } T_2 \text{ rounds}] \geq 1 - \left(1 - \frac{1}{4\delta_0^2}\right)^{12\delta_0 \ln n} \geq 1 - \frac{1}{n^3}.
$$

We have shown for any node $v \in N^2_i (u)$, it will connect to $u$ in round $T_3 = T_1 + T_2$ with probability at least $1 - 1/n^3$. This implies in round $T_3$, $u$ will connect to all nodes in $N^0_i (u)$ with probability at least $1 - [N^0_i (u)]/n^3$. Then, $N^0_i (u) \subseteq N^1_{T_3} (u)$, $N^0_i (u) \subseteq N^1_{T_3} (u) \cup N^2_{T_3} (u)$, $N^0_i (u) \subseteq N^1_{T_3} (u) \cup N^2_{T_3} (u) \cup N^3_{T_3} (u)$. We apply the above analysis twice, and obtain that in round $T = 3T_3 = O(n \log n)$, $N^0_i (u) \cup N^2_i (u) \cup N^3_i (u) \subseteq N^1_i (u)$ with probability at least $1 - [N^0_i (u) \cup N^2_i (u) \cup N^3_i (u)]/n^3 \geq 1 - 1/n^2$. By Lemma 1, $[N^0_i (u) \cup N^2_i (u) \cup N^3_i (u)] \geq \min \{2\delta_0, n - 1\}$, thus completing the proof.

We are now ready to prove Theorem 3.

**Proof of Theorem 3.** We first show that in $O(n \log n)$ rounds, either the graph becomes complete or its minimum degree increases by a factor of at least $1/8$. Then we apply this argument $O(\log n)$ times to complete the proof.

For each $u$ where $d_0 (u) < \min \{(1 + 1/8)\delta_0, n - 1\}$, we consider the following two cases. Let $S$ denote the set of nodes in $N^0_i (u)$ that are strongly tied to $N^2_i (u)$. The first case is when $|S| > \delta_0/3$; i.e., more than $\delta_0/3$ nodes in $N^0_i (u)$ are strongly tied to $N^2_i (u)$. By the definition of strongly tied, this means that each node in $S$ has at least $\delta_0/2$ edges to $N^0_i (u)$.
By Lemma 7, there exists $T = O(n \log n)$ such that if at least $\delta_0/4$ nodes in $N^t(u)$ have each at least $\delta_0/4$ edges to $N^2(u)$ for all $t < T$, then $d_T(u) \geq (1 + 1/8)\delta_0$ with probability at least $1 - 1/n^2$. We thus obtain that if every node in $S$ has at least $\delta_0/4$ edges to $N^2(u)$ for all $t < T$, then the desired claim of the lemma holds. On the other hand, if there is a node in $S$ that has fewer than $\delta_0/4$ edges to $N^2(u)$ for some $t < T$, then $u$ has gained at least $\delta_0/2 - \delta_0/4 = \delta_0/4$ new neighbors, again giving the desired claim.

The second case is if $|S| < \delta_0/3$; i.e., less than $\delta_0/3$ nodes in $N^0(u)$ are strongly tied to $N^2(u)$. We divide into two subcases. If we remain in this case for $T = O(n \log n)$ rounds and find a node $v_0$ in $N^0(u)$ that is strongly tied to $N^2(u)$, then by Lemma 8, we have $d_T(u) \geq \min \{(1 + 1/8)\delta_0, n - 1\}$ with probability at least $1 - 1/n^2$. Otherwise, by Lemma 9, we obtain that there exists $T = O(n \log n)$ such that if we remain in this case for $T$ rounds, then $d_T(u) \geq \min \{(1 + 1/8)\delta_0, n - 1\}$ with probability at least $1 - 1/n^2$. Whenever the condition is not satisfied, i.e., more than $\delta_0/3$ nodes in $N^1(u)$ are strongly tied to $N^2(u)$, we move to the analysis in the first case, and $d_T(u) \geq (1 + 1/8)\delta_0$ in $T = O(n \log n)$ rounds with probability at least $1 - 1/n^2$.

Combining the above two cases and applying a union bound, we obtain $d_T \geq \min \{(1 + 1/8)\delta_0, n - 1\} in T = O(n \log n)$ rounds with probability at least $1 - 1/n$. We now apply the above argument $O(\log n)$ times to obtain the desired $O(n \log^2 n)$ upper bound. 

3.2. Lower Bound

Next we give an almost matching lower bound.

Proof. Proof of Theorem 4 We first observe that during the triangulation process there is a time $t$ when the number of missing edges is at least $m = \Omega(\sqrt{k})$ and the minimum degree is at least $n/3$. If $k < \frac{2}{3}n$ then this is true initially and for larger $k$ this is true at the first time $t$ the minimum degree is large enough. The second case follows since the degree of a node (and thus also the minimum degree) can at most double in each step guaranteeing that the minimum degree is not larger than $\frac{2}{3}n$ at time $t$ also implying that at least $\frac{n}{3} = \Omega(\sqrt{k})$ edges are still missing.

Given the bound on the minimum degree any missing edge $(u, v)$ is added by a fixed node $w$ with probability at most $\frac{18}{n^2}$. Since there are at most $n - 2$ such nodes the probability that a missing edge gets added is at most $\frac{18}{n}$. To analyze the time needed for all missing edges to be added we denote with $X_t$ the random variable counting the number of steps needed until the $i$th of the $m$ missing edges is added. We would like to analyze $\Pr \{X_1 \leq T, X_2 \leq T, \ldots, X_m \leq T\}$ for an appropriately chosen number of steps $T$. Note that the events $X_i < T$ and $X_j < T$ are not independent and indeed can be positively or negatively correlated. Nevertheless, independent of the conditioning onto any of the events $X_i < T$, we have that $\Pr \{X_1 \leq T\} \leq 1 - (1 - \frac{18}{n^2})T \leq 1 - \frac{1}{\sqrt{m}}$ for an appropriately chosen $T = \Omega(n \log m)$, where $m$ is again the number of missing edges at time $t$. Thus, $\Pr \{X_1 \leq T, X_2 \leq T, \ldots, X_m \leq T\} = \Pr \{X_1 \leq T|X_2 \leq T, \ldots, X_m \leq T\} \cdot \Pr \{X_2 \leq T|X_3, \ldots, X_m \leq T\} \cdots \Pr \{X_m \leq T\} \leq \left(1 - \frac{1}{\sqrt{m}}\right)^m = O\left(e^{-\sqrt{m}}\right) = O\left(e^{-k/4}\right)$

This shows that the triangulation process takes with probability at least $1 - O\left(e^{-k/4}\right)$ at least $\Omega(n \log m) = O(n \log k)$ steps to complete. 

Random Structures and Algorithms DOI 10.1002/tsa
4. THE TWO-HOP WALK: DISCOVERY THROUGH PULL

In this section, we analyze the two-hop walk process on undirected connected graphs, which is described by the following simple iteration: In each round, for each node \( u \), we add edge \((u, w)\) where \( w \) is drawn uniformly at random from \( N^1_i(v) \), where \( v \) is drawn uniformly at random from \( N^1_i(u) \); if \( w = u \) or an existing neighbor of \( u \), then no new edge is added. The two-hop walk yields the following pull-based resource discovery protocol. In each round, each node \( u \) contacts a random neighbor \( v \), receives the identity of a random neighbor \( w \) of \( v \), and sends its identity to \( w \).

While both the two-hop walk and the push-based triangulation process perform linear work per round, they differ in how this work may be distributed among the nodes. In any given round, each new edge \((u, v)\) added in the triangulation process can be associated with a distinct “mediator” node \( w \) such that \((u, w)\) and \((v, w)\) are existing edges. This cannot be always done for two-hop walks; in a given round, multiple nodes may connect to new neighbors through a single “mediator” node that is one-hop away from the nodes.

The main result of this section is that the two-hop walk process transforms an arbitrary connected \( n \)-node graph to a complete graph in \( O(n \log^2 n) \) rounds with high probability. We also establish an \( \Omega(n \log n) \) lower bound on the two-hop walk for almost all \( n \)-node graphs.

As we did for the triangulation process, we establish the \( O(n \log^2 n) \) upper bound by showing that the minimum degree of the graph increases by a constant factor (or equals \( n - 1 \)) in \( O(n \log n) \) rounds with high probability. For analyzing the growth in the degree of a node \( u \), we consider two overlapping cases. The first case is when the two-hop neighborhood of \( u \) is not too large, i.e., \( |N^2_i(u)| < \delta_0/3 \), and the second is when the two-hop neighborhood of \( u \) is not too small, i.e., \( |N^2_i(u)| \geq \delta_0/4 \).

**Lemma 10** (When the two-hop neighborhood is not too large). There exists \( T = O(n \log n) \) such that with probability at least \( 1 - 1/n^2 \), we have \( |N^2_i(u)| \geq \delta_0/3 \) or \( d_T(u) \geq \min\{2\delta_0, n - 1\} \).

**Proof.** In the following, we use integers \( T, T_1, T_2, T_3 \), each bounded by \( O(n \log n) \), whose precise values will be set at various points of the proof to satisfy desired conditions. We assume that \( d_t(u) < 2\delta_0 \) for all \( t < T \); otherwise, the desired claim of the lemma holds.

By the definition of \( \delta_0, d_0(w) \geq \delta_0 \) for all \( w \) in \( N^0_i(u) \). Let \( X \) be the first round at which \( |N^2_{X+1}(u)| \geq \delta_0/3 \). We consider two cases. If \( X \) is at most \( cn \log n \) for a constant \( c \) to be specified later, then the claim of the lemma holds. In the remainder of this proof we consider the case where \( X \) is greater than \( cn \log n \); thus, for \( 0 \leq t \leq cn \log n, \ |N^2_{t}(u)| < \delta_0/3 \).

Fix a node \( v \) in \( N^2_0(u) \). In the following, we first show that in \( O(n \log n) \) rounds, \( v \) is strongly tied to the neighbors of \( u \) with probability at least \( 1 - 1/n^2 \). Let \( T_1 \) denote the first round at which \( v \) is strongly tied to \( N^1_{T_1}(u) \), i.e., when \( |N^1_{T_1}(v) \cap N^1_{T_1}(u)| \geq \delta_0/2 \). We know that \( v \) has at least one neighbor, say \( w_1 \), in \( N^0_{T_1}(u) \). Since \( d_0(w_1) \leq \delta_0 \) and \( |N^2_{T_1}(u)| < \delta_0/3, \ w_1 \) has at least \( 2\delta_0/3 \) edges to nodes in \( N^1_{T_1}(u) \). Consider any \( t < T_1 \). By our definition of \( T_1 \), \( v \) is weakly tied to \( N^1_{T_1}(u) \), and hence also to \( N^1_{T_1}(u) \) since the latter is a subset of the former. Therefore, at time \( t \), \( w_1 \) has at least \( 2\delta_0/3 - \delta_0/2 = \delta_0/6 \) neighbors in \( N^1_{T_1}(u) \) which do not have an edge to \( v \) at time \( t \). Since \( d_t(w_1) \leq 5\delta_0/2 \), we obtain

\[
\Pr[v \text{ connects to a node in } N^1_{T_1}(u) \setminus N^1_{T_1}(u) \text{ through } w_1 \text{ in round } t] \geq \frac{1}{n} \cdot \frac{\delta_0}{6} \cdot \frac{2}{5\delta_0} = \frac{1}{15n}
\]
Let $e_1$ denote the event $\{v \text{ connects to a new node in } N^1_0(u)\}$, and $X_1$ be the number of rounds for $e_1$ to occur. When $e_1$ occurs, let $w_1$ denote a witness for $e_1$; i.e., $w_1$ is a node in $N^1_0(u)$ to which $v$ connects in round $X_1$. We note that $w_1, w_2 \in N^1_0(u) \subseteq N^1_{x_1}(u)$ if $v$ is weakly tied to $N^1_{x_1}(u)$, both $w_1$ and $w_2$ have at least $\delta_0/6$ neighbors in $N^1_{x_1}(u)$ that do not have an edge to $v$ yet. Let $e_2$ denote the event $\{v \text{ connects to a new node in } N^1_{x_1}(u)\}$, and $X_2$ be the number of rounds for $e_2$ to occur. In the two-hop walk, the event that a two-hop walk from $v$ connects to a node through $w_1$ is mutually exclusive from the event that a two-hop walk from $v$ connects to a node through $w_2$. Since the lower bound of Eq. (1) applies to both $w_1$ and $w_2$, we obtain that $\Pr[e_2] \geq 2/(15n)$. Similarly, we define $e_3, X_3, \ldots, e_{6_0/2}, X_{6_0/2}$ and obtain $\Pr[e_i] \geq i/(15n)$. We now apply Lemma 2 to obtain that $X_1 + X_2 + \ldots + X_{6_0/2}$ is at most $75n \ln n$ with probability at least $1 - 1/n^3$. Thus, with probability at least $1 - 1/n^4$, $T_1 \leq 75n \ln n$. This is true for each $v$ in $N^0_0(u)$. Thus, with probability at least $1 - 1/n^3$, every node in $N^3_0(u)$ is strongly tied to $N^1_0(u)$ in at most $75n \ln n$ rounds. Once this event happens, we obtain that for any $v \in N^2_0(u)$,

$$\Pr[u \text{ connects to } v \text{ in a single round}] \geq \frac{\delta_0/2}{2n} \cdot \frac{1}{n} = \frac{1}{4n},$$

which implies that with probability at least $1 - 1/n^3$, $u$ has an edge to every node in $N^2_0(u)$ in another $T_2 \leq 12n \ln n$ rounds.

Let $T_3$ equal $T_1 + T_2$; we set $c$ to be a constant such that $X \geq 3T_3$. We thus have $N^2_0(u) \supseteq N^1_{x_2}(u), N^3_0(u) \subseteq N^1_{x_1} \cup N^2_{x_3}(u)$, and $N^4_0(u) \subseteq N^1_{x_1} \cup N^2_{x_3} \cup N^3_{x_3}$. We now repeat the above analysis again twice and obtain that at time $T = 3T_3, N^2_0(u) \cup N^3_0(u) \subseteq N^1_{x_3}(u)$ with probability at least $1 - \left|N^2_0(u) \cup N^3_0(u) \cup N^4_0(u)\right|/n^3 \geq 1 - 1/n^2$. By Lemma 1, we have $|N^3_{x_1}(u)| \geq \min\{2\delta_0, n - 1\}$, thus completing the proof of the lemma.

**Lemma 11** (When the two-hop neighborhood is not too small). There exists $T = O(n \log n)$ such that with probability at least $1 - 1/n^2$ we have $|N^2_0(u)| < \delta_0/4$ or $d^r(u) \geq \min\{(1 + 1/8)\delta_0, n - 1\}$.

**Proof**. Let $X$ be the first round at which $N^2_0(u) < \delta_0/4$. We consider two cases. If $X$ is at most $cn \log n$ for a constant $c$ to be specified later, then the claim of the lemma holds. In the remainder of this proof we consider the case where $X$ is greater than $cn \log n$; thus, for $0 \leq t \leq cn \log n$, $|N^2_t(u)| \geq \delta_0/4$. If $v \in N^2_0(u)$ is strongly tied to $N^1_0(u)$, then

$$\Pr[u \text{ connects to } v \text{ in a single round}] \geq \frac{d^t(v, N^1_0(u))}{|N^1_0(u)|} \cdot \frac{1}{n} \geq \frac{\delta_0/4}{(1 + 1/8)\delta_0} \cdot \frac{1}{n} = \frac{2}{9n}.$$ 

Thus, in $T = 13.5n \ln n$ rounds, $u$ will add an edge to $v$ with probability at least $1 - 1/n^3$. If there are at least $\delta_0/8$ nodes in $N^2_0(u)$ that are strongly tied to $N^1_0(u)$ at the start of round 0 is less than $\delta_0/8$. In this case, because $|N^2_0(u)| \geq \delta_0/4$, more than $\delta_0/8$ nodes in $N^2_0(u)$ are weakly tied to $N^1_0(u)$, and, thus, have at least $\delta_0/4$ edges to nodes in $N^3_0(u) \cup N^4_0(u)$.

In the following we show $u$ will connect to $\delta_0/8$ nodes in $O(n \log n)$ rounds with probability at least $1 - 1/n^2$. For any round $t$, let $W_t$ denote the set of nodes in $N^2_t(u)$ that are weakly tied to $N^1_t(u)$. We refer to a length-2 path from $u$ to a node two hops away
as an out-path. Let $P_0$ denote the set of out-paths to $W_0$. Since we have at least $\delta_0/8$ nodes in $N^2_0 (u)$ that are weakly tied to $N^1_0 (u)$, $|P_0|$ is at least $\delta_0/8$ at time $t = 0$. Define $e_i = \{ u \text{ connects to a new node in } N^2_0 (u) \text{ through an out-path in } P_0 \}$, and $X_1$ to be the number of rounds for $e_i$ to occur. When $0 \leq t \leq X_1$, for each $w_i \in N^1_0 (u)$, let $p_i$ be the number of edges from $w_i$ to nodes in $N^2_0 (u)$ that are weakly tied to $N^1_0 (u)$.

$$\Pr [e_i] \geq \frac{1}{|S|} \sum_i \frac{d_i (u)}{d_i (w_i)} \geq \frac{1}{|S|} \sum_i \frac{d_i (u)}{d_i (w_i)} \geq \frac{\sum_i p_i}{n} \geq \frac{\sum_i p_i}{n} \geq \frac{1}{n} \cdot \frac{n}{4} \cdot \frac{n}{8} \cdot \frac{1}{n} \cdot \frac{1}{9n}.$$ 

After $X_1$ rounds, $u$ will pick an out-path in $P_0$. Let the node $u$ connects to through this out-path be $v_1$. Define $P_i$ to be the set of out-paths from $u$ to $W_{X_i}$. In general, for $i \geq 1$, we define $e_i$, $P_i$, $X_i$, $v_i$ and $Y_i$ as follows. Let $e_i$ denote the event \{ $u$ connects to a new node in $N^2_0 (u)$ through an out-path in $P_{i-1}$ \}. Let $X_i$ be the number of rounds after event $e_i$ that it takes for $u$ to pick an out-path in $P_{i-1}$, $v_i$ to be a node to which $u$ connects through this out-path, and $P_i$ to be the set of out paths from $u$ to $W_{Y_i}$, where $Y_i = \sum_{i \leq j \leq X_i} X_j$.

We show by induction that $\Pr [e_i] \geq i/(9n)$ and $|P_i \setminus P_{i-1}| \geq \delta_0/8$. We have established the induction base $i = 1$ above. We now establish the induction step. Since $v_i \in N^1_0 (u)$ is added to $N^1_1 (u)$, those out-paths in $P_{i-1}$ consisting of edges from $v_i$ to nodes in $N^1_0 (u)$ are not in $P_i$. The number of out-paths we lose because of this is at most $\delta_0/4$. But $v_i$ also has at least $3\delta_0/4$ edges to $N^2_0 (u) \cup N^2_0 (u)$. The end points of these edges are in $N^1_1 (u) \cup N^1_2 (u)$. If more than $\delta_0/8$ of them are in $N^1_1 (u)$, then $d_{y_i} (u) \geq (1 + 1/8)\delta_0$. Now let’s consider the case that less than $\delta_0/8$ such end points are in $N^1_1 (u)$. This means the number of edges from $v_i$ to $N^2_1 (u)$ is at most $3\delta_0/4 - \delta_0/4 - \delta_0/8 = 3\delta_0/8$. Among the end points of these edges, if more than $\delta_0/8$ of them are strongly tied to $N^1_1 (u)$, then the degree of $u$ will become at least $(1 + 1/8)\delta_0$ in $O(n \log n)$ rounds with probability $1 - 1/n^2$ by our earlier argument. If not, we know that more than $\delta_0/4$ newly added edges are pointing to nodes that are weakly tied to $N^1_1 (u)$. Thus, $|P_i \setminus P_{i-1}|$ is at least $\delta_0/4$, and we obtain from the induction hypothesis $|P_i| \geq (i - 1)\delta_0/8 + \delta_0/8 = i\delta_0/8$. Thus, during round $t$, $Y_{t-1} \leq t \leq Y_t$, $\Pr [e_i]$ is at least $\frac{2}{\delta_0}$. Let $T$ equal the random variable $X_1 + X_2 + \cdots + X_{\delta_0/8}$, so that $d_T (u) \geq (1 + 1/8)\delta_0$. By Lemma 2, $T$ is at most $(2 + 1/9)n \ln n = 27n \ln n$, with probability at least $1 - 1/n^2$, completing the proof of this lemma. 

**Theorem 12 (Upper bound for two-hop walk process).** For connected undirected graphs, the two-hop walk process completes in $O(n \log^2 n)$ rounds with high probability.

**Proof.** We first show that in time $T = O(n \log n)$ time, the minimum degree of the graph increases by a factor of $1/8$, i.e., $\delta_T \geq \min \{(1 + 1/8)\delta_0, n - 1\}$. Then we can apply this argument $O(\log n)$ times, and thus, complete the proof of this theorem.

For each $u$ where $d_{\delta_0} (u) < \min \{(1 + 1/8)\delta_0, n - 1\}$, we analyze by the following 2 cases. First, if $|N^1_2 (u)| \geq \delta_0/2$, by Lemma 11 we know as long as $|N^1_2 (u)| \geq \delta_0/4$ for all $t \geq 0$, $d_T (u) \geq \min \{(1 + 1/8)\delta_0, n - 1\}$ with probability $1 - 1/n^2$ where $T = O(n \log n)$. Whenever the condition is not satisfied, we know at least $\delta_0/4$ nodes in $N^1_0 (u)$ have been moved to $N^1_1 (u)$, which means $d_T (u) \geq \min \{(1 + 1/4)\delta_0, n - 1\}$.

Second, if $|N^1_0 (u)| < \delta_0/2$, by Lemma 10 we know as long as $|N^1_2 (u)| < \delta_0/2$ for all $t \geq 0$, $d_T (u) \geq \min \{(1 + 1/8)\delta_0, n - 1\}$ with probability $1 - 1/n^2$ where $T = O(n \log n)$. 

*Random Structures and Algorithms* DOI 10.1002/rsa
Whenever the condition is not satisfied, we are back to the analysis in the first case, and the minimum degree will become \( \min \{ (1 + 1/8)\delta_0, n - 1 \} \) with probability \( 1 - 1/n^2 \).

Combining the the above two cases we get that with probability \( 1 - 1/n \) the minimum degree of \( G \) will become at least \( \min \{ (1 + 1/8)\delta_0, n - 1 \} \) in \( O(n \log n) \) rounds, since there are at most \( n \) nodes whose degree is between \( \delta_0 \) and \( \min \{ (1 + 1/8)\delta_0, n - 1 \} \).

Now we can apply the above argument \( O(\log n) \) times, and have shown the two-hop walk process completes in \( O(n \log^2 n) \) with high probability.

The proof of Theorem 13 is essentially the same as that for Theorem 4, and is omitted.

**Theorem 13** (Lower bound for two-hop walk process). For any connected undirected graph \( G \) that has \( k \geq 1 \) edges less than the complete graph the two-hop process takes \( \Omega(n \log k) \) steps to complete with probability at least \( 1 - O(e^{-k^{1/4}}) \).

5. TWO-HOP WALK IN DIRECTED GRAPHS

In this section, we analyze the two-hop walk process in directed graphs. We say that the process terminates at time \( t \) if for every node \( u \) and every node \( v \), \( G_t \) contains the edge \( (u, v) \) whenever \( u \) has a path to \( v \) in \( G_0 \).

**Theorem 14.** On any \( n \)-node directed graph, the two-hop walk terminates in \( O(n^2 \log n) \) rounds with high probability. Furthermore, there exists a (weakly connected) directed graph for which the process takes \( \Omega(n^2 \log n) \) rounds to terminate.

**Proof.** Consider any pair of nodes, \( u \) and \( v \) that do not have an edge between them. Consider a shortest path from \( u \) to \( v \) \( (v_0, v_1, v_2, \ldots, v_m) \), where \( v_0 = u, v_m = v \) and \( 2 \leq m \leq n \). Fix a time step \( t \). Let \( e_i \) denote the event an edge is added from \( v_j \) to \( v_{j+2} \) in step \( t \). The probability of occurrence of \( e_i \) is \( \Pr[e_i] \geq 1/n^2 \). All the \( e_i \)'s are independent from one another.

\[
\Pr[\bigcup e_i] \geq \sum_i \Pr[e_i] - \sum_i \sum_j \Pr[e_i \cap e_j] \\
= \sum_i \Pr[e_i] - \sum_i \sum_j \Pr[e_i] \Pr[e_j] \\
\geq m \frac{1}{n^2} - m(m - 1) \frac{1}{n^4} \\
\geq \frac{m}{n^2} - \frac{1}{n^2},
\]

since \( 2 \leq m \leq n \). Let \( X_1 \) denote the number of steps it takes for the length of the above path to decrease by 1. It is clear that \( E[X_1] \leq n^2/(m - 1) \). In general, let \( X_i \) denote the number of steps it takes for the length of the above path to decrease by \( i \). By Lemma 2, the number of steps it takes for the above path to shrink to an edge is at most \( 8n \ln n \) with probability \( 1/n^3 \). Taking a union bound over all the edges yields the desired upper bound.

For the lower bound, consider a graph \( G_0 \) with the node set \( \{1, 2, \ldots, n\} \) and the edge set \( \{(3i,j), (3i+1,j) : 0 \leq i < n/4, 3n/4 \leq j < n\} \cup \{(3i, 3i+1), (3i+1, 3i+2) : 0 \leq i < n/4\} \).
The only edges that need to be added by the two-hop process are the edges $(3i, 3i + 2)$ for $0 \leq i < n/4$. The probability that node $3i$ adds the edge $(3i, 3i + 2)$ in any round is at most $16/n^2$. The probability that edge $(3i, 3i + 2)$ is not added in $(n^2 \ln n)/32$ rounds is at least $1/\sqrt{n}$. Since the events associated with adding each of these edges are independent, the probability that all the $n/3$ edges are added in $(n^2 \ln n)/32$ rounds is at most $(1 - 1/\sqrt{n})^{n/3} \leq e^{-\sqrt{n}/3}$, completing the lower bound proof.

The lower bound in the above theorem takes advantage of the fact that the initial graph is not strongly connected. Extending the above analysis for strongly connected graphs appears to be much more difficult since the events corresponding to the addition of new edges interact in significant ways. We present an $\Omega(n^2)$ lower bound for a strongly connected graph by a careful analysis that tracks the event probabilities with time and takes dependencies into account. The graph $G_0$, depicted in Fig. 3, is similar to the example in [25] used to establish an $\Omega(n)$ lower bound on the Random Pointer Jump algorithm, in which each node gets to know all the neighbors of a random neighbor in each step. Since the graphs are constantly changing over time in both the processes, the dynamic edge distributions differ significantly in the two cases, and we need a substantially different analysis.

**Theorem 15.** There exists a strongly connected directed graph for which the expected number of rounds taken by the two-hop process is $\Omega(n^2)$.

**Proof.** The graph $G_0 = (V, E)$ is depicted in Fig. 3 and formally defined as $G_0 = (V, E)$ where $V = \{1, 2, \ldots, n\}$ with $n$ being even, and

$$E = \{(i, j) : 1 \leq i, j \leq n/2\} \cup \{(i, i + 1) : n/2 \leq i < n\} \cup \{(i, j) : i > j, i > n/2, i, j \in V\}.$$

We first establish an upper bound on the probability that edge $(i, i + h)$ is added by the start of round $t$, for given $i$, $1 \leq i \leq n - h$. Let $p_{h,t}$ denote this probability. The following base cases are immediate: $p_{0,0}$ is 1 for $h = 1$ and $h < 0$, and 0 otherwise. Next, the edge $(i, i + h)$ is in $G_{r+1}$ if and only if $(i, i + h)$ is either in $G_{r-1}$ or added in round $t$. In the latter case, $(i, i + h)$ is added by a two-hop walk $i \rightarrow i + k \rightarrow i + h$, where $-i < k \leq n - i$. Since the out-degree of every node is at least $n/2$, for any $k$ the probability that $i$ takes such a walk is at most $4/n^2$.

$$p_{h,t+1} \leq p_{h,t} + \frac{4}{n^2} \sum_{k=-i}^{n-i} p_{k,t} p_{h-k,t}$$

$$= p_{h,t} + \frac{4}{n^2} \left( \sum_{k=1}^{h-1} p_{h+k,t} + \sum_{k=1}^{h-i} p_{k,t} p_{h-k,t} + \sum_{k=h+1}^{n-i} p_{k,t} \right). \tag{2}$$
We show by induction on $t$ that

$$p_{h,t} \leq \left(\frac{\alpha t}{n^2}\right)^{h-1}, \text{ for all } t \leq \epsilon n^2. \tag{3}$$

where $\alpha$ and $\epsilon$ are positive constants that are specified later.

The induction base is immediate. For the induction step, we use the induction hypothesis for $t$ and Eq. (2) and bound $p_{h,t+1}$ as follows.

$$p_{h,t+1} \leq \left(\frac{\alpha t}{n^2}\right)^{h-1} + \frac{4}{n^2} \sum_{k=1}^{h-1} \left(\frac{\alpha t}{n^2}\right)^{h+k-1} + \sum_{k=1}^{h-1} \left(\frac{\alpha t}{n^2}\right)^{h-k-1} + \sum_{k=h+1}^{n-i} \left(\frac{\alpha t}{n^2}\right)^{k-1}$$

In the second inequality above, we combine the first and third summations and bound them by their infinite sums. In the third inequality, we use $t \leq \epsilon n^2$ and set $\alpha, \epsilon$ such that $\alpha \epsilon \leq 1/2$, implying that $2(\alpha t/n^2)^2 \leq 1 \leq h - 1$. For the fourth inequality, we set $\alpha$ sufficiently large so that $\alpha \geq 4 + 4/(1 - \alpha \epsilon)$. The two preceding conditions for $\alpha$ and $\epsilon$ can be satisfied by setting $\alpha = 9$ and setting $\epsilon$ sufficiently small. The final inequality follows from Taylor series expansion.

For an integer $x$, let $C_x$ denote the cut $\{(u \mid u \leq x), \{v \mid v > x\}\}$. We say that a cut $C_x$ is untouched at the start of round $t$ if the only edge in $G_t$ crossing the cut $C_x$ is the edge $(x, x + 1)$; otherwise, we say $C_x$ is touched. Let $X$ denote the smallest integer such that $C_x$ is untouched. We note that $X$ is a random variable that also varies with time. Initially, $X = n/2$.

We divide the analysis into several phases, numbered from 0. A phase ends when $X$ changes. Let $X_i$ denote the value of $X$ at the start of phase $i$; thus $X_0 = n/2$. Let $T_i$ denote the number of rounds in phase $i$. A new edge is added to the cut $C_{X_i}$ only if either $X_i$ selects edge $(X_i, X_i + 1)$ as its first hop or a node $u < X_i$ selects $u \rightarrow X_i \rightarrow X_i + 1$. Since the degree of every node is at least $n/2$, the probability that a new edge is added to the cut $C_i$ is at most $2/n + n(4/n^2) = 6/n$, implying that $E[T_i] \geq n/6$.

We now place a bound on $X_{t+1}$. Fix a round $t \leq \epsilon n^2$, and let $E_t$ denote the event that $C_x$ is touched by round $t$. We first place an upper bound on the probability of $E_t$ for arbitrary $x$ using Eq. (3).

$$\Pr[E_t] \leq \sum_{h \geq 2} h \left(\frac{\alpha t}{n^2}\right)^{h-1} \leq \frac{\alpha t(4 - 3(\alpha t)/n^2 + (\alpha t)^2/n^4)}{n^2(1 - (\alpha t)/n^2)^3},$$

for $t \leq \epsilon n^2$, where we use the inequality $\sum_{h \geq 2} h^2 \delta^h = \delta(4 - 3\delta + \delta^2)/(1 - \delta)^3$ for $0 < \delta < 1$. The expression $\delta(4 - 3\delta + \delta^2)/(1 - \delta)^3$ is a decreasing function of $\delta$ for $\delta < 1$. Since $t \leq \epsilon n^2$.
we have \( at/n^2 \leq \alpha \epsilon \). Therefore, we have

\[
\frac{at(4 - 3(at)/n^2 + (at)^2/n^4)}{n^2(1 - (at)/n^2)^3} \leq \alpha \epsilon \left(\frac{4 - 3\alpha \epsilon - (\alpha \epsilon)^2}{(1 - \alpha \epsilon)^3}\right) \leq 45\epsilon,
\]

where we use the setting \( \alpha = 9 \) and set \( \epsilon \) sufficiently small so that \((4 - 3\alpha \epsilon - (\alpha \epsilon)^2)/(1 - \alpha \epsilon)^3 \leq 5\).

If \( E_x \) were independent from \( E_y \) for \( x \neq y \), then we can invoke a straightforward analysis using a geometric probability distribution to argue that \( E[X_{i+1} - X_i] \) is at most \( 1/(1 - 45\epsilon) = O(1) \); to see this, observe that \( X_{i+1} - X_i \) is stochastically dominated by the number of tosses of a biased coin needed to get one head, where the probability of tail is \( 45\epsilon \). The preceding independence does not hold, however; in fact, for \( y > x \), \( \Pr\{E_y \mod E_x \} > \Pr\{E_y\} \). We show that the impact of this correlation is very small when \( x \) and \( y \) are sufficiently far apart.

Consider a sequence of cuts \( C_{x_1}, C_{x_2}, \ldots, C_{x_t} \) where \( x_i = x_{i-1} + c\ell \), for a constant \( c \) chosen sufficiently large, and we set \( x_0 = X_1 + 2 \). We bound the conditional probability of \( E_{x_{i+1}} \) given \( E_{x_{i+1}} \cap E_{x_{i+2}} \cdots \cap E_{x_t} \) as follows. In the following, we use \( e(C) \) to denote the set of edges crossing the cut \( C \).

\[
\begin{align*}
\Pr\{E_{x_{i+1}} | E_{x_{i+1}} \cap E_{x_{i+2}} \cdots \cap E_{x_t}\} &= \frac{\Pr\{E_{x_{i+1}} \cap E_{x_{i+2}} \cdots \cap E_{x_t}\}}{\Pr\{E_{x_{i+1}} \cap E_{x_{i+2}} \cdots \cap E_{x_t}\}} \\
&\leq \frac{\Pr\{E_{x_{i+1}} \cap E_{x_{i+2}} \cdots \cap E_{x_t}\} \cap \left( e(C_{x_{i+1}}) \cap (e(C_{x_{i+1}}) \cup \cdots \cup e(C_{x_t})) = \emptyset \right) }{\Pr\{E_{x_{i+1}} \cap E_{x_{i+2}} \cdots \cap E_{x_t}\}} \\
&\leq \frac{\Pr\{E_{x_{i+1}} \cap E_{x_{i+2}} \cdots \cap E_{x_t}\} \cap \left( e(C_{x_{i+1}}) \cap (e(C_{x_{i+1}}) \cup \cdots \cup e(C_{x_t})) \neq \emptyset \right) }{\Pr\{E_{x_{i+1}} \cap E_{x_{i+2}} \cdots \cap E_{x_t}\}} \\
&\leq \Pr\{E_{x_{i+1}}\} + \frac{(\alpha \epsilon)/n^2)^c \ell - 1}{(1 - \alpha \epsilon/n^2)^c \ell - 1} \\
&\leq 45\epsilon + \epsilon = 46\epsilon,
\end{align*}
\]

where we set \( c \) sufficiently large in the last step. Since \( X_{i+1} \) is at most the smallest \( x_\ell \) such that \( C_{x_\ell} \) is untouched, we obtain that

\[
E[X_{i+1} - X_i] \leq 2 + \sum_{\ell \geq 2} (46\epsilon)\epsilon \ell^2 \leq c',
\]

for a constant \( c' \) chosen sufficiently large. We thus obtain that after \( \epsilon' n \) phases, \( E[X] \) is at most \( \epsilon' c' n \), where \( \epsilon' \) is chosen sufficiently small so that \( n - E[X] \) is \( \Omega(n) \). Since the expected length of each phase is \( \Omega(n) \), it follows that the expected number of rounds it takes for the two-hop process to complete is \( \Omega(n^2) \) rounds.

Random Structures and Algorithms DOI 10.1002/rsa
6. CONCLUSION

We have analyzed two natural gossip-based discovery processes in networks and showed almost-tight bounds on their convergence in arbitrary networks. Our processes are motivated by the resource discovery problem in distributed networks as well as by the evolution of social networks. While we have addressed the convergence problem to a complete graph in this paper, it will be interesting to study other relevant questions, especially in the context of social networks: (1) How and when do well-connected clusters (communities) emerge? (2) When does the clustering coefficient become large? In particular, can we generalize our results to obtain bounds on how long it takes for evolving graphs to achieve a particular graph density or clustering coefficient? (3) How does the diameter change with time? (4) The triangulation and two-hop walks are both “memoryless” processes. It has been shown recently that the inclusion of memory in rumor-spreading can be beneficial in certain classes of graphs [21]. Can adding memory to the two processes we study have a similar effect? These questions can lead to a better understanding of how discovery processes influence the evolution of social networks.

We would like to study variants of the processes that take into account failures associated with forming connections, the joining and leaving of nodes, or having only a subset of nodes to participate in forming connections. We believe our techniques can be extended to analyze such situations as well. From a technical standpoint, the main problem left open by our work is to resolve the logarithmic factor gap between the upper and lower bounds. It is not hard to show that from the perspective of increasing the minimum degree by a constant factor, our analysis is tight up to constant factors. It is conceivable, however, that a sharper upper bound can be obtained by an alternative analysis that uses a “smoother” measure of progress.

REFERENCES

[1] I. Abraham and D. Dolev, Asynchronous resource discovery, Comput Netw 50 (2006), 1616–1629.
[2] M. Adler, E. Halperin, R. M. Karp, and V. V. Vazirani, A stochastic process on the hypercube with applications to peer-to-peer networks, In STOC, 2003, pp. 575–584.
[3] N. Alon, Problems and results in extremal combinatorics – II, Discrete Mathe 308 (2008), 4460–4472.
[4] J. Augustine, A. Rahman Molla, E. Morsy, G. Pandurangan, P. Robinson, and E. Upfal, Storage and search in dynamic peer-to-peer networks, In SPAA, 2013, pp. 53–62.
[5] J. Augustine, G. Pandurangan, and P. Robinson, Fast byzantine agreement in dynamic networks, In PODC, 2013, pp. 74–83.
[6] J. Augustine, G. Pandurangan, P. Robinson, and E. Upfal, Towards robust and efficient computation in dynamic peer-to-peer networks, In SODA, 2012, pp. 551–569.
[7] C. Avin, M. Koucký, and Z. Lotker, How to explore a fast-changing world (cover time of a simple random walk on evolving graphs), In ICALP (1), 2008, pp. 121–132.
[8] O. Babaoglu and M. Jelasity, Self-* properties through gossiping, Philos Trans R Soc A 366 (2008), 3747–3757.
[9] A. Berns, S. Ghosh, and S. Pemmaraju, A framework for building self-stabilizing overlay networks, In PODC, 2010, pp. 398–399. Brief Announcement.
[10] S. Bornholdt and H. Georg Schuster, editors Handbook of graphs and networks, Wiley-VCH, 2003.
[11] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, Randomized gossip algorithms, IEEE Trans. on Infor. Theory 52 (2006), 2508–2530.
[12] S. Chakrabarti, A. Frieze, and J. Vera, The influence of search engines on preferential attachment, In SODA, 2005, pp. 293–300.
[13] J.-Y. Chen and G. Pandurangan, Almost-optimal gossip-based aggregate computation, SIAM J Comput 41 (2012), 455–483.
[14] F. Chierichetti, S. Lattanzi, and A. Panconesi, Almost tight bounds on rumor spreading and conductance, In STOC, 2010, pp. 399–408.
[15] A. E. F. Clementi, P. Crescenzi, C. Doerr, P. Fraigniaud, M. Isopi, A. Panconesi, F. Pasquale, and R. Silvestri, Rumor spreading in random evolving graphs, In ESA, 2013, pp. 325–336.
[16] A. E. F. Clementi, C. Macci, A. Monti, F. Pasquale, and R. Silvestri, Flooding time in edge-markovian dynamic graphs, In PODC, 2008, pp. 213–222.
[17] A. E. F. Clementi, R. Silvestri, and L. Trevisan, Information spreading in dynamic graphs, In PODC, 2012, pp. 37–46.
[18] C. Cooper and A. Frieze, Crawling on web graphs, In STOC, 2002, pp. 419–427.
[19] A. Demers, D. Greene, C. Hauser, W. Irish, J. Larson, S. Shenker, H. Sturgis, D. Swinehart, and D. Terry, Epidemic algorithms for replicated database maintenance, In PODC, 1987, pp. 1–12.
[20] N. B. Dimitrov and C. G. Plaxton, Optimal cover time for a graph-based coupon collector process, In ICALP, 2005, pp. 702–716.
[21] B. Doerr, M. Fouz, and T. Friedrich, Social networks spread rumors in sublogarithmic time, In STOC, 2011, pp. 21–30.
[22] B. Doerr, T. Friedrich, and T. Sauerwald, Quasi-random rumor spreading, In SODA, 2008, pp. 773–781.
[23] G. Giakkoupis, Tight bounds for rumor spreading in graphs of a given conductance, In STACS, 2011, pp. 57–68.
[24] G. Giakkoupis, T. Sauerwald, and A. Stauffer, Randomized rumor spreading in dynamic graphs, In ICALP(2), 2014, pp. 495–507.
[25] M. Harchol-Balter, F. Thomson Leighton, and D. Lewin, Resource discovery in distributed networks, In PODC, 1999, pp. 229–237.
[26] R. Jacob, A. Richa, C. Scheideler, S. Schmid, and H. Taubig, A distributed polylogarithmic time algorithm for self-stabilizing skip graphs, In PODC, 2009, pp. 131–140.
[27] M. Jelasity, A. Montresor, and O. Babaoglu, T-man: Gossip-based fast overlay topology construction, Comput Netw 53 (2009), 2321–2339.
[28] R. M. Karp, C. Schindelhauer, S. Shenker, and B. Vöcking, Randomized rumor spreading, In FOCS, 2000, pp. 565–574.
[29] D. Kempe, A. Dobra, and J. Gehrke, Gossip-based computation of aggregate information, In FOCS, 2003, pp. 482–491.
[30] D. Kempe and J. Kleinberg, Protocols and impossibility results for gossip-based communication mechanisms., In FOCS, 2002, pp. 471–480.
[31] D. Kempe, J. Kleinberg, and A. Demers, Spatial gossip and resource location protocols, In STOC, 2001, pp. 163–172.
[32] S. Kutten, D. Peleg, and U. Vishkin, Deterministic resource discovery in distributed networks, In SPA, 2001, pp. 77–83.
[33] C. Law and K.-Y. Siu, An O(log n) randomized resource discovery algorithm, In DISC, 2000, pp. 5–8. Brief Announcement.
[34] M. Mitzenmacher and E. Upfal, Probability and computing: Randomized algorithms and probabilistic analysis, Cambridge University Press, 2005.
[35] D. Mosk-Aoyama and D. Shah, Computing separable functions via gossip, In PODC, 2006, pp. 113–122.

[36] M. J. Newman, A.-L. Barabasi, and D. J. Watts, Structure and dynamics of networks, Princeton University Press, 2006.

[37] A. D. Sarma, A. R. Molla, and G. Pandurangan, Fast distributed computation in dynamic networks via random walks, In DISC, 2012, pp. 136–150.

[38] F. Vega-Redondo, Complex social networks, Cambridge University Press, 2007.