Social Networks Spread Rumors in Sublogarithmic Time

Benjamin Doerr
Max-Planck-Institut für Informatik
66123 Saarbrücken, Germany

Mahmoud Fouz
Universität des Saarlandes
Fachrichtung Informatik
66041 Saarbrücken, Germany

Tobias Friedrich
Max-Planck-Institut für Informatik
66123 Saarbrücken, Germany

ABSTRACT

With the prevalence of social networks, it has become increasingly important to understand their features and limitations. It has been observed that information spreads extremely fast in social networks. We study the performance of randomized rumor spreading protocols on graphs in the preferential attachment model. The well-known random phone call model of Karp et al. (FOCS 2000) is a push-pull strategy where in each round, each vertex chooses a random neighbor and exchanges information with it. We prove the following.

- The push-pull strategy delivers a message to all nodes within $\Theta(\log n)$ rounds with high probability. The best known bound so far was $\mathcal{O}(\log^2 n)$.
- If we slightly modify the protocol so that contacts are chosen uniformly from all neighbors but the one contacted in the previous round, then this time reduces to $\Theta(\log n / \log\log n)$, which is the diameter of the graph. This is the first time that a sublogarithmic broadcast time is proven for a natural setting. Also, this is the first time that avoiding double-contacts reduces the run-time to a smaller order of magnitude.

Categories and Subject Descriptors

F.2 [**Theory of Computation**]:
Analysis of Algorithms and Problem Complexity

General Terms

Algorithms, Measurement, Theory

1. INTRODUCTION

The first picture [23] of US Airways Flight 1549's crash landing on the Hudson River became known to a broad audience through *Twitter* so quickly that people were wondering if *Twitter* will replace traditional news media in the near future. Even with traditional means of communication, all of us have witnessed how news spread remarkably fast among our friends, colleagues and other social networks.

Not least because of the emergence of huge online social networks like *Facebook* with currently 600 Million members [7] is the subject of information dissemination on social networks relevant from an algorithmic point of view.

In this paper, we study a natural rumor spreading protocol and prove that it spreads a rumor in sublogarithmic time in a classic graph theoretic model for social networks.

As graph model we use the *preferential attachment* (PA) model originally introduced by Barabási and Albert [1]. It builds on the paradigm that new vertices attach to already present vertices with a probability proportional to their degree. Rigorous studies [2–6, 11, 17] show that this model indeed enjoys many properties observed in social networks, e.g., a power law distribution of the vertex degrees, a small diameter and a small average degree. The precise definition of the PA-model can be found in Section 2.

To model the rumor spreading process, we always assume a discrete time line. The rumor first appears at an arbitrary vertex in round 0. We are interested in the number of rounds necessary until all vertices are informed.

A simple way to model the rumor spreading process is to assume that in each round, each vertex that knows the rumor, forwards it to a randomly chosen neighbor. This is known as the *push strategy*. For many network topologies, this strategy is a very efficient way to spread a rumor. Let n denote the number of vertices of a graph. Then the push model with high probability (i.e., with probability $1 - o(1)$) sends the rumor to all vertices in time $\Theta(\log n)$, if the graph is a complete graph [19, 26], a hypercube [16], an Erdős-Rényi random graph $G_{n,p}$ with $p \geq (1+\varepsilon)\log(n)/n$ [16, 20], a random regular graph [18], or an expander graph [18, 27]. In contrast to this, Chierichetti, Lattanzi, and Panconesi [8] showed that the push model with non-vanishing probability needs $\Omega(n^\alpha)$ rounds on PA-graphs for some $\alpha > 0$.

Opposite to the push strategy is the *pull strategy*: each vertex in each round contacts a random neighbor and learns the rumor if its contact knows the rumor already. There is a symmetry between the two models. This was observed for a quasirandom version of the two models in [13], but similar arguments also hold for the two random models discussed so far. Thus, the above results also hold for the pull model.

Karp, Schindelhauer, Shenker, and Vöcking [22] pointed out that for complete graphs, the pull strategy is inferior to the push strategy until roughly $n/2$ vertices are informed, and then the pull strategy becomes more effective. This motivates to combine both approaches. In this so-called *push-*

STOC'11, June 6–8, 2011, San Jose, California, USA.

pull strategy each vertex contacts another vertex chosen uniformly at random among its neighbors. It *pushes* the rumor in case it has the rumor, and *pulls* the rumor in case the neighbor has the rumor. For complete graphs this protocol also needs $\Theta(\log n)$ rounds, though with better implicit constants [12, 14, 22]. Elsässer [14] also proved a lower bound of $\Omega(\log n)$ rounds for Erdős-Rényi random graphs $G_{n,p}$ with $p \geq \text{polylog}(n)/n$. For preferential attachment graphs, however, the push-pull strategy is much better than push or pull alone. Chierichetti et al. [8] showed that with this strategy, $\mathcal{O}(\log^2 n)$ rounds suffice with high probability.

So far it has been open how sharp this bound is. The recent works on graphs with high conductance only show that for graphs with conductance Φ the broadcast time is bounded by $\mathcal{O}(\Phi^{-1} \log^2(\Phi^{-1}) \log n)$ [9]. Unfortunately, the conductance of the preferential attachment model seems not known. Several power law graphs have a conductance of $\Phi = \Omega(\log^{-1} n)$ [10, 21] and this has also been observed empirically for real social networks [24]. Mihail, Papadimitriou, and Saberi [25] showed that certain graphs that are similar to PA-graphs have constant conductance. If this was true also for the PA-model, a bound of $\mathcal{O}(\log n)$ would follow.

Our results: We prove that the push-pull protocol indeed with high probability spreads the rumor to all nodes in a PA-graph in time $\Theta(\log n)$. If we assume a slightly more clever process, namely that contacts are chosen uniformly at random among all neighbors except the one that was chosen just in the round before, then $\mathcal{O}(\log n/\log\log n)$ rounds suffice (cf. Theorem 3.1). This is asymptotically optimal as the diameter of a PA-graph is $\Theta(\log n/\log\log n)$ [6]. This result can be seen as an explanation why rumor spreading in actual social networks is extremely fast.

We should note that the idea of excluding previously contacted nodes is not new. Elsässer and Sauerwald [15] used the exclusion of the previous three contacts to design protocols that reduce the number of messages sent, an aspect important when using such protocols to disseminate information in networks, e.g., to maintain distributed databases [12]. However, excluding previous contacts so far did not yield a faster rumor spreading. In fact, Elsässer and Sauerwald [15] have shown that the $\Omega(\log n)$ lower bound for rumor spreading in Erdős-Rényi random graphs $G_{n,p}$, $p > \text{polylog}(n)/n$, remains true if arbitrary exclusion schemes are used.

2. PRECISE MODEL AND PRELIMINARIES

Preferential attachment graphs were first introduced by Barabási and Albert [1]. In this work, we follow the formal definition of Bollobás et al. [5, 6]. Let G be an undirected graph. We denote by $\deg_G(v)$ the degree of a vertex v in G.

Definition 2.1 (Preferential attachment graph). *Let $m \geq 2$ be a fixed parameter. The random graph G_m^n is an undirected graph on the vertex set $V := \{1, \dots, n\}$ inductively defined as follows.*

- *G_m^1 consists of a single vertex with m self-loops.*
- *For all $n > 1$, G_m^n is built from G_m^{n-1} by adding the new node n together with m edges $e_n^1 = \{n, v_1\}, \dots, e_n^m = \{n, v_m\}$ inserted one after the other in this order. Let $G_{m,i-1}^n$ denote the graph right before the edge e_n^i is added. Let $M_i = \sum_{v \in V} \deg_{G_{m,i-1}^n}(v)$ be the sum of the degrees of all the nodes in $G_{m,i-1}^n$. The endpoint v_i is selected randomly such that $v_i = u$ with probability $\deg_{G_{m,i-1}^n}(u)/(M_i + 1)$, except for n that is selected with probability $(\deg_{G_{m,i-1}^n}(n) + 1)/(M_i + 1)$.*

This definition implies that when e_n^i is inserted, the vertex v_i is chosen with probability proportional to its degree (except for $v_i = n$). Since many real-world social networks are conjectured to evolve using similar principles, the PA-model can serve as a model for social networks. Another property observed in many real-world networks has been formally proven for preferential attachment graphs, namely that the degree distribution follows a power-law [6].

It can be easily seen that for $m = 1$ the graph is disconnected with high probability; so we focus on the case $m \geq 2$. Under this assumption, Bollobás and Riordan [5] showed that the diameter is only $\Theta(\log(n)/\log\log n)$ with high probability.

With a slight abuse of notation we write $(u, v) \in E$ or $(v, u) \in E$ both to denote $\{u, v\} \in E$. Note that the definition of G_m^n can lead to multiple edges and self-loops, though they typically make up only a vanishing fraction of the edges.

We examine the following broadcasting protocol.

Definition 2.2 (Push-pull strategy with memory). *Let $M \geq 0$ be a fixed parameter. Assume that every vertex can store M vertices. The protocol runs as follows:*

- *In each round $t \geq 1$, every vertex u chooses uniformly at random a neighbor v which it has not contacted in the last $\min\{\deg(u) - 1, M\}$ rounds. If u knows the rumor, it sends the rumor to v ("push"). If v knows the rumor, it sends the rumor to u ("pull").*

Note that for $M = 0$, this is the classic push-pull strategy.

We denote by $\log n$ the natural logarithm to the base e.

3. STATEMENT OF RESULTS

Our main result is that preferential attachment graphs allow sublogarithmic time rumor spreading.

Theorem 3.1. *With probability $1 - o(1)$, the push-pull protocol with memory $M \geq 1$ broadcasts a rumor from any node of G_m^n to all other nodes in $\mathcal{O}(\log n/\log\log n)$ rounds.*

Our proof uses several arguments of Bollobás and Riordan [5] who showed that preferential attachment graphs have a diameter of $\Theta(\log(n)/\log\log n)$. In particular, we heavily use the equivalent non-recursive definition of preferential attachment graphs (see Section 5.1 for details). Of course, some additional work is needed to show that the process indeed only needs a time of order of the diameter. Recall that the diameter is only a lower bound for the rumor spreading process. As the complete graph with diameter one and rumor spreading time $\Omega(\log n)$ demonstrates, there can be a substantial gap between the two quantities.

The proof of Theorem 3.1 consists of three main steps. In Section 5.3, we analyze the time needed until the rumor reaches a so-called *useful node*. Roughly speaking, a node is useful if its degree is at least polylogarithmic (see Section 5.2 for details). We give a simple proof that for $M \geq 2$, a useful node is reached in only $\mathcal{O}(\log\log n)$ rounds. The more involved proof that for $M = 1$, a time of $\mathcal{O}(\log^{3/4}(n) \log\log n)$ rounds suffices, will appear in the full version of the paper.

The core of the proof (see Section 5.4) consists of showing that once a useful node u has been informed, within $\mathcal{O}(\log(n)/\log\log n)$ time steps the rumor is propagated to node 1. To this aim, we show that there is a path from u to 1 such that every second node (i) has degree exactly m and (ii) has the property that once one of its neighbors becomes informed, it pulls the rumor from there and pushes it to all other neighbors in exactly m rounds. Thus, the nodes of constant degree seem to be a key to fast rumor spreading on social networks. This observation has a similar flavor as the structural property proven by Chierichetti et al. [8] that social networks have a connected subgraph of linear size and diameter $\mathcal{O}(\log n)$ in which every node has degree $\mathcal{O}(\log n)$. Using this property, the authors showed a running time of $\mathcal{O}(\log^2 n)$ for the push-pull protocol without memory.

Finally, in Section 5.5, we use a symmetry property of the process to show that also in $\mathcal{O}(\log(n)/\log\log n)$ time steps the rumor is sent from node 1 to all other nodes.

For the classic push-pull strategy we show in Section 5.3 that it reaches a useful node in $\mathcal{O}(\log n)$ rounds. As the second and third part of the above proof also holds in this case, this gives the following matching upper and lower bounds.

THEOREM 3.2. *With probability $1-o(1)$, the classic push-pull protocol broadcasts a rumor from any node of G_m^n to all other nodes in $\mathcal{O}(\log n)$ rounds.*

THEOREM 3.3. *With prob. $1-o(1)$, the classic push-pull protocol needs $\Omega(\log n)$ rounds to inform all nodes of G_m^n.*

To see why a single memory slot can lead to an asymptotic speed-up, it is instructional to prove the lower bound first.

4. LOWER BOUND OF CLASSIC PUSH-PULL

We use the following result by Bollobás et al. [6]. Let $\#_m^n(d)$ be the number of nodes of indegree d in G_m^n.

THEOREM 4.1 (BOLLOBÁS ET AL. [6]). *Let $m \geq 1$ be fixed. Let*

$$\alpha_m(d) := \frac{2m\,(m+1)}{(d+m)\,(d+m+1)\,(d+m+2)},$$

and let $\varepsilon > 0$ be fixed. Then, with probability $1-o(1)$, we have

$$(1-\varepsilon)\,\alpha_m(d) \leq \frac{\#_m^n(d)}{n} \leq (1+\varepsilon)\,\alpha_m(d),$$

for every d in the range $0 \leq d \leq n^{1/5}$.

The proof strategy is as follows. We first show that with high probability there are $\Omega(n)$ edges whose incident nodes are of constant degree. Both nodes of such an edge remain uninformed with constant probability in each round. It is then easy to show that at least for one edge the incident nodes remain uninformed after $\Omega(\log n)$ rounds.

Proof of Theorem 3.3. Let X_c denote the total degrees of all nodes of indegree at most c for some constant $c > 0$. By Theorem 4.1, we have with probability $1-o(1)$,

$$\begin{aligned} X_c &\geq (1-\varepsilon)\sum_{d=0}^{c}(d+m)\,n\,\alpha_m(d) \\ &= (1-\varepsilon)\,2m\,(m+1)n\sum_{d=m}^{c+m}\frac{1}{(d+1)(d+2)} \\ &= (1-\varepsilon)\,2m\,(m+1)n\left(\tfrac{1}{m+1}-\tfrac{1}{c+m+2}\right) \\ &= (1-\varepsilon)\,2m\,n\left(1-\tfrac{m+1}{c+m+2}\right). \end{aligned}$$

Note that each edge that connects two nodes each of degree at most c is counted twice in X_c, whereas each edge that connects one node of degree at most c with a node of degree larger than c is counted only once. Hence, $X_c - mn$ is a lower bound on the number of edges that connect two nodes of degree at most c. So with probability $1-o(1)$, we have at least $(1-\varepsilon)\,2m\,n\left(1-\frac{m+1}{c+m+2}-\frac{1}{2(1-\varepsilon)}\right)$ such edges. Thus, for sufficiently large constant c, we have $\Omega(n)$ such edges and therefore also $\Omega(n)$ such pairs of nodes each of degree at most $c+m$ that are connected to each other.

Consider such a pair nodes (u,v). Assume that both nodes are uninformed. Then, they remain uninformed after one round if in this round (i) both nodes contact each other and (ii) none of the other neighboring nodes contacts u or v. Since these two events are independent, the probability that u and v remain uninformed after one round is at least

$$\left(\tfrac{1}{m+c}\right)^2\left(1-\tfrac{1}{m}\right)^{2(m+c)} =: \delta.$$

Note that $\delta \in \Omega(1)$. The probability that (u,v) remains uninformed after $\alpha\log n$ rounds for $\alpha = 1/(2\ln\delta^{-1})$ is therefore at least $n^{-1/2}$. The probability that none of the $\Omega(n)$ pairs remains uninformed after $\alpha\ln n$ rounds is at most $(1-n^{-1/2})^{\Omega(n)} \leq e^{-\Omega(n^{1/2})}$. □

Note that this proof fails when nodes do not contact the same neighbor twice in a row. For a similar argument to work in that case, one would need to show that there exists a polynomial number of triangles that consist of small degree nodes. In Lemma 5.3, we prove that this is not the case.

5. UPPER BOUND

5.1 Alternative model

In the random process generating G_m^n the random decisions made at each step depend heavily on the previous random decisions. To circumvent this problem, Bollobás and Riordan [5] suggested an alternative way of generating G_m^n that is more accessible. We first describe the model for $m = 1$ and then generalize it to arbitrary m. We refer the reader to [5] for a proof that both models are equivalent.

Let (x_i, y_i) for $i \in [n] := \{1, 2, \ldots, n\}$ be n independently and uniformly chosen pairs from $[0,1] \times [0,1]$. With probability one, all these numbers are distinct. By reordering within each pair, we assume that $x_i < y_i$ for every $i \in [n]$. Suppose that after relabeling, $y_1 < y_2 < \cdots < y_n$. We set $W_0 := 0$ and $W_i := y_i$ for $i \in [n]$. The graph G_1^n is now defined by having an edge (i,j) if and only if $W_{j-1} < x_i < W_j$. Define $w_j := W_j - W_{j-1}$.

Similarly, for G_m^n, we sample mn pairs $(x_{i,j}, y_{i,j})$ independently and uniformly from $[0,1] \times [0,1]$ with $x_{i,j} < y_{i,j}$ for $i \in [n]$ and $j \in [m]$. We relabel the variables such that $y_{i,j}$

is increasing in lexicographic order:

$$y_{1,1} < y_{1,2} < \cdots < y_{1,m} < y_{2,1} < \cdots < y_{n,1} < \cdots < y_{n,m}.$$

We set $W_0 := 0$ and $W_i := y_{i,m}$ for $i \in [n]$. The graph is now defined by having an edge (i,j) for each $k \in [m]$ such that $W_{j-1} < x_{i,k} < W_j$. As before, define $w_j = W_j - W_{j-1}$. We write $\ell_{i,k}$ for the node j such that $W_{j-1} < x_{i,k} < W_j$.

Note that given $y_{1,1}, \ldots, y_{n,m}$, the random variables $x_{1,1}, \ldots, x_{n,m}$ are independent with x_{i_k} being chosen uniformly from $[0, y_{i,k}]$. For a better readability, we instead assume that if $y_{1,1}, \ldots, y_{n,m}$ are given, then each $x_{i,k}$ is chosen independently and uniformly from $[0, W_i]$. By this slight modification, we can work with the values of the W_i's and ignore the values of the $y_{i,j}$'s. Note that this modification only increases the probability of a loop at i. It is straightforward to check that each step of our proof remains valid if the probability of a loop is not increased. Thus, the validity of our proof is not affected.

We give a few properties of the alternative model, that hold with high probability and are useful in the analysis. Let $s = 2^a$ be the smallest power of 2 larger than $\log^7 n$, and let 2^b be the largest power of 2 smaller than $2n/3$. Let $I_t = [2^t + 1, 2^{t+1}]$.

LEMMA 5.1 (BOLLOBÁS AND RIORDAN [5]). *Let $m \geq 2$ be fixed. Using the definitions above, each of the following five events holds with probability $1 - o(1)$.*

- $E_1 := \{|W_i - \sqrt{i/n}| \leq \frac{1}{10}\sqrt{i/n}$ *for all* $i \in [s, n]\}$
- $E_2 := \{|\{i \in I_t \mid w_i \geq 1/(10\sqrt{in})\}| \geq 2^{t-1}$ *for all* $t \in [a, b)\}$
- $E_3 := \{w_1 \geq \frac{4}{\log(n)\sqrt{n}}\}$
- $E_4 := \{w_i \geq \log^2(n)/n$ *for all* $i < n^{1/5}\}$
- $E_5 := \{w_i < \log^2(n)/n$ *for all* $i \geq n/2\}$.

Note that the event E_5 is slightly adjusted for our purposes. In the original paper, the authors show that for $i \geq n/\log^5 n$, we have $w_i < n^{-4/5}$. It is easy to check that (essentially) the same proof holds for the above version.

Instead of working directly with the alternative model where the W_i's are random variables, we use the following *typical social network* model where we assume the W_i's to be fixed numbers that satisfy the properties $E_1, \ldots, E_5$. Since by Lemma 5.1, these properties hold with high probability, all results proven for a typical social network model carry over to G_m^n with high probability. More precisely, Let $0 < W_1 < \cdots < W_n < 1$ be distinct real numbers and let $w_i = W_i - W_{i-1}$. Assume that $W_1, \ldots, W_n$ satisfy the properties $E_1, \ldots, E_5$. A typical social network $G_m(W_1, \ldots, W_n)$ is obtained by connecting each node i with the nodes $\ell_{i,1}, \ldots, \ell_{i,m}$, where each $\ell_{i,k}$ is a node chosen randomly with $\mathbb{P}[\ell_{i,k} = j] = w_j/W_i$ for all $j \leq i$.

In the following, we always assume to have a typical social network $G_m(W_1, \ldots, W_n)$. We write $G := G_m(W_1, \ldots, W_n)$ to denote a (random) typical social network.

5.2 Useful nodes

We use the notion of a *useful* node that was introduced by Bollobás and Riordan [5]. A node i is useful if $w_i \geq \log^2(n)/n$. Note that we are slightly relaxing the original definition in [5] where the authors also assumed that $i \leq n/\log^5(n)$. For our purposes, we have by E_5 that $i < n/2$ for all useful nodes. Furthermore by E_4, every $i < n^{1/5}$ is useful. We now prove several properties of non-useful nodes.

LEMMA 5.2. *With prob. $1-o(1)$, the following event holds*

- $E_6 := \{\deg_G(v) \leq 5m \log^2 n$ *for all non-useful* $v\}$.

Proof. Let i be a fixed non-useful node. So $w_i < \log^2(n)/n$ and by E_4, $i \geq n^{1/5}$. Consider any node $j > i$. By E_1, we have $W_j \geq \frac{1}{2}\sqrt{j/n}$. Moreover, for any $k \in \{1, \ldots, m\}$, $\mathbb{P}[\ell_{j,k} = i] = w_i/W_j \leq \frac{2\log^2 n}{n\sqrt{j/n}}$. Denote by $\deg_G^+(i)$ the number of edges $(j,i) \in E$ with $j > i$. Then $\deg_G(i) \leq 2m + \deg_G^+(i)$, where the first term is due to the at most m self-loops at i. We have

$$\begin{aligned}
\mathbb{E}[\deg_G^+(i)] &= \sum_{j>i}\sum_{k=1}^{m} \mathbb{P}[\ell_{j,k} = i] \\
&\leq 2m \log^2(n) n^{-1/2} \sum_{j>i}^{n} j^{-1/2} \\
&\leq 2m \log^2(n) n^{-1/2} \int_{j>i-1}^{n} j^{-1/2}\,\mathrm{d}j \leq 4m \log^2(n).
\end{aligned}$$

By Chernoff's bound, we have $\mathbb{P}[\deg_G^+(i) \geq 4.5m \log^2 n] \leq e^{-m\log^2(n)/48} = n^{-\Omega(\log n)}$. By a union bound, we conclude that with probability $1 - n^{-\Omega(\log n)}$ all non-useful nodes have degree at most $2m + 4.5m\log^2 n \leq 5m \log^2 n$. □

We call a cycle and a path *non-useful* if they consist only of non-useful nodes.

LEMMA 5.3. *With prob. $1-o(1)$, the following event holds*

- $E_7 := \{G$ *contains* $(\log n)^{O(\log^{3/4} n)}$ *non-useful cycles of length at most* $\log^{3/4} n\}$.

Proof. Let $\ell \in [n]$. We first bound the number of non-useful cycles of length ℓ. For simplicity, we assume that ℓ is even. The case when ℓ is odd is similar. Let $i_1 < i_2 < \cdots < i_\ell$ be ℓ distinct non-useful nodes. We set $i_{\ell+1} := i_1$. For simplicity, we write $\widetilde{w}_j := w_{i_j}$ and $\widetilde{W}_j := W_{i_j}$. The probability that $i_1, \ldots, i_\ell, i_{\ell+1} = i_1$ form a cycle in G in this order is

$$\begin{aligned}
&\mathbb{P}\Big[\bigwedge_{j=1}^{\ell}(i_j, i_{j+1}) \in E\Big] \leq \prod_{j=1}^{\ell}\Big(m \max\Big\{\frac{\widetilde{w}_{j+1}}{\widetilde{W}_j}, \frac{\widetilde{w}_j}{\widetilde{W}_{j+1}}\Big\}\Big) \\
&\leq m^\ell \prod_{j=1}^{\ell}\Big(\frac{\log^2 n}{n}\max\{\widetilde{W}_j^{-1}, \widetilde{W}_{j+1}^{-1}\}\Big) \\
&\overset{(E_1)}{\leq} m^\ell \prod_{j=1}^{\ell}\Big(\frac{10\log^2 n}{9\sqrt{n}}\max\{i_j^{-1/2}, i_{j+1}^{-1/2}\}\Big) \\
&\leq m^\ell \prod_{j=1}^{\ell/2}\Big(\Big(\frac{10\log^2 n}{9\sqrt{n}}\Big)^2 \frac{1}{i_j}\Big) \leq m^\ell\Big(\frac{10\log^2 n}{9\sqrt{n}}\Big)^\ell \prod_{j=1}^{\ell/2}\frac{1}{i_j}.
\end{aligned}$$

Note that the same upper bound holds for every permutation of $i_1, i_2, \ldots, i_\ell$. Thus we can bound the expected number of cycles consisting of these nodes by $\ell!\, m^\ell\Big(\frac{10\log^2 n}{9\sqrt{n}}\Big)^\ell \prod_{j=1}^{\ell/2}\frac{1}{i_j}$.

In consequence, the expected number of non-useful cycles of length ℓ is bounded by

$$\begin{aligned}
&\sum_{i_1<\cdots<i_\ell} \ell!\, m^\ell \Big(\frac{10\log^2 n}{9\sqrt{n}}\Big)^\ell \prod_{j=1}^{\ell/2} \frac{1}{i_j} \\
&\le \ell!\, m^\ell \Big(\frac{10\log^2 n}{9\sqrt{n}}\Big)^\ell \sum_{i_{\ell/2+1}<\cdots<i_\ell} \sum_{i_1<\cdots<i_{\ell/2}} \prod_{j=1}^{\ell/2} \frac{1}{i_j} \\
&\le \ell!\, m^\ell \Big(\frac{10\log^2 n}{9\sqrt{n}}\Big)^\ell n^{\ell/2} \left(\sum_{i=1}^{n} \tfrac{1}{i}\right)^{\ell/2} \\
&\le \ell!\, m^\ell \,(\tfrac{10}{9})^\ell \,(\log n)^{2\ell} \,(\log(n)+1)^{\ell/2} \\
&\le \ell!\, m^\ell \,(\log n)^{3\ell},
\end{aligned}$$

where the last inequality holds for sufficiently large n.

By Markov's inequality, we conclude that with probability at most $1/\log n$, there are more than $\ell!\, m^\ell (\log n)^{3\ell+1}$ non-useful cycles of length ℓ. By a simple union bound, it follows that with probability at least $1-(\log n)^{-1/4}$, there are at most $(\log^{3/4} n)^{1+\log^{3/4} n}\, m^{\log^{3/4} n}\, (\log n)^{3\log^{3/4}(n)+1} = (\log n)^{O(\log^{3/4} n)}$ non-useful cycles of length at most $\log^{3/4} n$. □

Lemma 5.4. *Assume that E_6 holds and let $K = \frac{\log n}{(\log\log n)^2}$. With probability $1-n^{-1/5+o(1)}$, the following event holds*

- *$E_8 := \{$for all non-useful v, there exists at most one cycle whose nodes are all connected to v via non-useful paths of length at most $K\}$.*

In order to prove Lemma 5.4, we show a few auxiliary results. We first bound the probability that two fixed non-useful nodes (not necessarily distinct) are neighbors.

Lemma 5.5. *Let v, v' be two fixed non-useful nodes. Then the probability that $(v,v') \in E$ is at most $n^{-3/5+o(1)}$.*

Proof. W.l.o.g. assume that $v \ge v'$. Since v' is not useful, we have $w_{v'} < \log^2(n)/n$. Using E_1 and the fact that $v \ge v' \ge n^{1/5}$, we obtain $W_v \ge W_{\lceil n^{1/5}\rceil} \ge \frac{1}{2} n^{-2/5}$ and thus

$$\mathbb{P}[(v,v') \in E] \le m w_{v'}/W_v \le n^{-3/5+o(1)}. \qquad (5.1)$$

□

Lemma 5.6 (Bollobás and Riordan [5]). *Let v be a fixed non-useful node. Then for all $k \in [m]$, the probability that $\ell_{v,k}$ is a useful node is at least $\log^{-3} n$. This event is independent from all other random decisions $\ell_{v',k'}$ with $(v',k') \ne (v',k')$.*

Note that in the original lemma, the authors only state a bound on the probability that $\ell_{v,1}$ is a useful node. However, the same proof yields the above version. Also, Lemma 5.5 and Lemma 5.6 remain valid if we condition on E_6.

For the next lemma, we need some notation. Let v be a non-useful node. Let $L^0 = \{v\}$ and for $k \ge 1$, we define

$$\begin{aligned}
L^k := \{w \in [n] \mid w \text{ is not useful} \wedge w \notin L^1, \ldots, L^{k-1} \\
\wedge \exists w' \in L^{k-1} \colon (w',w) \in E\}.
\end{aligned}$$

We define $L^{\le k} := \bigcup_{i=0}^{k} L^i$. Let $n_k = |L^k|$. We say that level L^i *causes a collision* if there exist two nodes $j, j' \in L^i$ (not necessarily distinct) that are either neighbors or share a common neighbor in L^{i+1}, or if there exists a node $j \in L^i$ that is connected to a node in L_{i+1} by two links, i.e., j is incident to a multi-edge. Note that each collision caused by L^i corresponds to a cycle that is connected to v via a path consisting only of nodes in $L^{\le i+1}$ and vice versa, every such cycle corresponds to one collision in some L^j where $j \le i+1$. Hence, the number of collisions in $L^{\le k}$ is exactly the number of cycles in $L^{\le k+1}$.

Lemma 5.7. *Assume that E_6 holds. Let $k \le \frac{\log n}{(\log\log n)^2}$ and $c > 0$ be a constant. We have*

$$\mathbb{P}\left[L^k \text{ causes } c \text{ collisions} \mid L^1, \ldots, L^k\right] \le n^{-3c/5+o(1)}.$$

Proof. In the following all probabilities are conditioned on $L_1, \ldots, L^k$. Since all L^i contain only non-useful nodes, by E_6, we have $|L^k| \le (5m\log^2 n)^k$. For any two fixed nodes $j, j' \in L^k$ (not necessarily distinct), we have

$$\begin{aligned}
&\mathbb{P}\left[j, j' \text{ cause a collision}\right] \\
&\le \mathbb{P}\left[(j,j') \in E\right] \\
&\qquad + \mathbb{P}[\exists j'' \notin L^{\le k} \colon (j,j'') \in E \wedge (j',j'') \in E] \\
&\le n^{-3/5+o(1)} + 5m\log^2(n)\, n^{-3/5+o(1)} \qquad \text{by (5.1)} \\
&= n^{-3/5+o(1)}.
\end{aligned}$$

Similarly, for any fixed node $j \in L^k$, we have

$$\begin{aligned}
&\mathbb{P}\left[\exists j' \notin L^{\le k} \colon j \text{ and } j' \text{ are connected by two edges}\right] \\
&\le 5m\log^2(n)\, n^{-3/5+o(1)} = n^{-3/5+o(1)}.
\end{aligned}$$

Thus, we have

$$\begin{aligned}
\mathbb{P}\left[L^k \text{ causes } c \text{ collisions}\right] &\le (|L^k|^2 + |L^k|)^c n^{-3c/5+o(1)} \\
&= n^{-3c/5+o(1)}.
\end{aligned}$$

□

Proof of Lemma 5.4. Let v be a fixed non-useful node. By Lemma 5.7, the probability that there exists a single level that causes two collisions is at most $Kn^{-6/5+o(1)} = n^{-6/5+o(1)}$. Similarly, the probability that there exist two levels that cause a collision each is at most $K^2 n^{-6/5+o(1)} = n^{-6/5+o(1)}$. The result follows from a simple union bound over all non-useful nodes. □

5.3 Informing the first useful node

Let $G = G_m(W_1, \ldots, W_n)$ be a typical social network. Assume that also E_6, E_7, and E_8 hold. In this section, all probabilities are taken over the product space of the random graph G and the random decisions of the rumor spreading process.

For simplicity, we start with the push-pull strategy with memory $M \ge 2$ and show that with high probability the rumor reaches a useful node within $\mathcal{O}(\log\log n)$ rounds.

Lemma 5.8. *Let u be a fixed node. In the push-pull strategy with $M \ge 2$, the rumor initiated by u reaches a useful node in $\mathcal{O}(\log\log n)$ rounds with probability $1-n^{-\Omega(\log n)}$.*

Proof. We make some assumptions that simplify the analysis while only slowing down the protocol. First, we assume that all nodes perform only push operations. Second, we consider *phases* of three rounds. In the first phase only u is active. In every subsequent phase we assume that only the nodes

that were informed in the previous phase are active. In other words, we assume that nodes that are informed in one phase are *delayed* till the beginning of the next phase and remain active only for that phase.

Let D_k be the set of all nodes that are active in phase k and $n_k = |D_k|$. Let $K = 14 \log\log n$. Assume that there is no useful vertex in any D_k for $k < K$. Then by E_8, we encounter at most one cycle consisting solely of nodes in $D_1 \cup \cdots \cup D_K$ (including self-loops or cycles due to multiple edges). We first assume that there is no such cycle. Then since $M \geq 2$, a node of degree at least 3 contacts three *distinct* neighbors in three rounds. Thus, such a node in D_k informs at least two new nodes in D_{k+1} (excluding its neighbor in D_{k-1}). Since both neighbors of a node of degree 2 must have degree at least 3, we conclude that $n_{k+2} \geq 2n_k$ for all $k \leq K$. So, we have $n_K \geq \Omega(2^{K/2}) \geq \Omega(\log^7 n)$. It is easy to see that a single cycle reduces n_K by at most a factor of 2. Hence, we still have $n_K \geq \Omega(\log^5 n)$.

For each $i \in D_K$ except for at most one node, only one of the nodes $\ell_{i,1}$ and $\ell_{i,2}$ can be informed (otherwise there would be a cycle). Let ℓ_i denote any of the nodes $\ell_{i,1}$ and $\ell_{i,2}$ that is not informed. Given the sequence $D_1, \ldots, D_M$, the nodes ℓ_i, where $i \in D_K$, are mutually independent. Conditioned on $\ell_i \notin D_1, \ldots, D_K$, the probability that ℓ_i is useful can only increase since $D_1, \ldots, D_K$ only contain non-useful nodes. So by Lemma 5.6, for any $i \in D_K$, we have $\mathbb{P}[\ell_i \text{ is useful}] \geq \log^{-3} n$. Also, for any $i \in D_K$, the probability that i contacts ℓ_i in one time step is at least $1/\deg(i) \geq 1/(5m \log^2 n)$ due to E_6. Since both events are independent, the probability that no node in D_K informs a useful node in one time step is at most $(1 - 1/(5m \log^5 n))^{n_K} \leq \exp(-n_K/(5m \log^5 n)) \leq n^{-\Omega(\log n)}$. □

For the classic push-pull protocol ($M = 0$), the proof idea is similar. The main difference is that now we can not assume that a node of degree at least 3 always contacts three distinct neighbors in one phase. However, in expectation, it only takes $\mathcal{O}(1)$ rounds until it does so. To get a high probability statement, we need $\mathcal{O}(\log n)$ rounds.

LEMMA 5.9. *In the classic push-pull protocol, the rumor initiated by any node reaches a useful node in $\mathcal{O}(\log n)$ rounds with probability $1 - o(n^{-2})$.*

Proof. Again, we consider phases and active nodes. As before, D_k denotes the set of active nodes in phase k and $n_k = |D_k|$. Let $K = 210 \log\log n$. By E_7, there is at most one cycle in $D_1 \cup \cdots \cup D_K$. For simplicity, we first assume that there is no cycle in $D_1 \cup \cdots \cup D_K$.

We distinguish two stages. In the first stage, we do not fix the length of each phase: a phase will last for a constant number of rounds in expectation. Active nodes of degree 2 remain active until they contact two distinct neighbors; those of higher degree remain active until they contact three distinct neighbors. A phase lasts until all active nodes stop. Then, a new phase starts in which all nodes become active that have been informed in the previous phase. In the first phase, only the initially informed node u is active.

In the second stage, each phase will last for exactly five rounds. The second stage starts at the earliest phase K' such that $n_{K'} \geq C \log n$ nodes for some constant $C > 0$. Note that $K' = \mathcal{O}(\log\log n)$ and $N := \sum_{1 \leq k \leq K'} n_k \leq 2C \log n$.

We now bound the length of the first stage. Let X_i denote the number of rounds needed until an active node i contacts two or three distinct neighbors depending on whether it has degree 2 or more. We can then bound from above the length the first stage by $X = \sum_{1 \leq k < K'} \sum_{i \in D_k} X_i$. Note that for all i, $X_i - 1$ is stochastically dominated by the sum of two geometric random variables $Y_{i,1} + Y_{i,2}$ with parameter $p = \frac{1}{3}$. Let $Y = \sum_{1 \leq k < K'} \sum_{i \in D_k} (Y_{i,1} + Y_{i,2})$. Let $X = \sum_{1 \leq i \leq 24C \log n} X_i$ denote the sum of $24C \log n$ i.i.d. indicator random variables with $\mathbb{P}[X_i = 1] = p$. Thus, $\mathbb{E}[X] = 24pC \log n = 8C \log n \geq 4N$. By Chernoff's bound, we get

$$\begin{aligned}\mathbb{P}\big[Y > 24C \log n\big] = \mathbb{P}[X < 2N] &\leq \mathbb{P}[X \leq \tfrac{1}{2}\mathbb{E}[X]] \\ &\leq \exp\big(-\tfrac{1}{8}\mathbb{E}[X]\big) = n^{-C}.\end{aligned}$$

We conclude that the first stage lasts for at most $24C \log n$ iterations with probability $1 - O(n^{-C})$.

For the second stage, remember that both neighbors of a node of degree 2 must have degree at least 3. Furthermore, the probability that a node of degree 2 contacts two distinct nodes in five rounds is $1 - 2^{-4} = \frac{15}{16}$ and similarly, the probability that a node of degree at least 3 contacts three distinct nodes in five rounds is at least $1 - 3(\frac{2}{3})^5 \geq 0.6$. Assume that we are given $D_1, \ldots, D_k$ and $n_k \geq C \log n$. Let $i \in D_k$. If node i has degree at least 3, then let X_i be the indicator variable for the event that it contacts two distinct nodes in D_{k+1} which in turn contact one node in D_{k+2}, respectively. By the previous discussion, we have $\mathbb{P}[X_i = 1] \geq 0.6\big(\frac{15}{16}\big)^2 \geq 0.527$. Similarly, if node i has degree 2, let X_i be the indicator variable for the event that it contacts one node (of degree at least 3) in D_{k+1} which in turn contacts two nodes in D_{k+2}. Here, we have $\mathbb{P}[X_i = 1] \geq \frac{15}{16} 0.6 \geq 0.527$. Let $X = \sum_{i \in D_k} X_k$. Note that all X_i are independent from each other and $n_{k+2} \geq 2X$. Thus, we have $\mathbb{E}[n_{k+2} \mid n_k] \geq 1.054 n_k$. By Chernoff's bound, we further get

$$\mathbb{P}[X \leq 1.05 n_k] \leq \exp\big(-\Omega(1) n_k\big) \leq n^{-3}, \tag{5.2}$$

where the last inequality follows from $n_k \geq C \log n$ by choosing C sufficiently large, but constant. Thus, with probability $1 - o(n^{-2})$, we have $n_K \geq \Omega(1.05^K) \geq \Omega((\log n)^{10})$. Although so far we have ignored the possibility of encountering a cycle, it is clear that a single cycle does not affect this bound. Thus, by the same argument as for Lemma 5.6, once phase $K = 210 \log\log n$ is active, either a useful node was already informed or the probability that a useful node is informed in one round is $1 - o(n^{-2})$. □

When $M \geq 2$, it was easy to see that, with probability one, a significant progress is made after a constant number of rounds. In contrast, when $M = 1$, a bad situation occurs when the informed nodes form a (small) non-useful cycle. In that case, with non-negligible probability, no progress is made, since each informed node could alternatively contact one of its two neighbors on the cycle. By E_7, however, we know that such a bad situation occurs only rarely. Thus, we distinguish between the case when the starting node lies in such a non-useful cycle or not. For each case, a similar argument as before, then yields the following lemma.

LEMMA 5.10. *Let u be any node. In the push-pull protocol with $M = 1$, the probability p_u that the rumor initiated by u does not reach any useful node in $\mathcal{O}(\log^{3/4}(n) \log\log n)$ time steps, satisfies $p_u = o(1)$. Moreover, the sum $\sum_{u \in V} p_u$ of all these failure probabilities is also $o(1)$.*

5.4 Informing node 1

What ultimately makes rumor spreading by push-pull protocols in preferential attachment graphs fast, are vertices of small (constant) degree. Each of them, with constant probability, has the beautiful property that, once a neighbor becomes informed, it pulls the rumor from such a neighbor and pushes it to all other neighbors in a total number of rounds equal to its degree.

As we will see in this section, this property alone suffices to spread the rumor among all useful nodes. Since this property is satisfied by all push-pull protocols considered in this paper (i.e., memory $M \geq 0$), the results in this section are valid for all of them. In fact, they might be of independent interest for other protocols that satisfy this property.

More specifically, we show that between any two useful nodes there is a path of length $\mathcal{O}(\log n/\log\log n)$ such that every second node on the path has this property. Since these nodes (by definition with probability one) propagate the rumor in constant time, we see that the rumor is propagated along such a path in time $\mathcal{O}(\log n/\log\log n)$.

Consider a fixed graph G and a run of our rumor spreading protocol started with the rumor in some node u. Let $v \neq u$ be a node of degree exactly m. Let t be the first time that some neighbor of v is informed. We say that v is *fast* (in this run of the protocol) if the following is true. (i) In round $t+1$, v performs a pull action on the smallest neighbor that is informed at that time (such a neighbor exists by choice of t). (ii) In rounds $t+2,\dots,t+m$, v performs push actions aimed at all other neighbors without repetition.

The event that some node v is fast is independent from the random decisions of all other nodes in the protocol. For this reason, the following lemma is the key to our analysis (to be continued with Corollary 5.15).

LEMMA 5.11. *Let $W_1,\dots,W_n$ be such that properties $E_1,\dots,E_5$ are satisfied. Consider the random graph $G = G_m(W_1,\dots,W_n)$. Let $p \in [0,1]$ be a constant. Mark each node $v \in \left[\frac{2}{3}n, n\right]$ independently with probability p.*

Let $v \in [n]$, be a useful node. Then with probability $1-o(n^{-1})$ in the product space of random graph and random marks, there exists a path of length $\mathcal{O}(\log n/\log\log n)$ between v and 1 such that every second node is marked.

We start by showing that with high probability, the random graph regarded contains a linear number of marked nodes. Of course, the main ingredient for this statement is the fact that there is a linear number of nodes $i \in \left[\frac{2}{3}n, n\right]$ that have a degree equal to m. If not explicitly stated, all probabilities in this section are taken over the product space of the random graph $G_m(W_1,\dots,W_n)$ and the random marks, where $W_1,\dots,W_n$ are given numbers that satisfy properties $E_1,\dots,E_5$.

LEMMA 5.12. *Let $\varepsilon_m := \frac{1}{8}pe^{-3m}$. With probability $1-e^{-\Omega(n)}$, there are at least $\varepsilon_m n$ marked nodes.*

Proof. Since $\sum_{i=1}^n w_i = 1$, at least half of the $i \in \left[\frac{2}{3}n, n\right] := C$ have $w_i \leq 6/n$. Let $i \in C$ be such that $w_i \leq 6/n$. Note that i has degree equal to m if and only if no node $j > i$ is a neighbor of i. Even conditioning arbitrarily on the degrees of all nodes in $C \setminus \{i\}$, we have for all $k \in \{1,\dots,m\}$, $\mathbb{P}[\ell_{j,k} = i] \leq w_i/(W_j - \sum_{r=\frac{2}{3}n}^{j} w_r) \leq (6/n)/W_{\frac{2}{3}n} \leq (6/n)/(0.9\cdot\sqrt{2/3}) \leq 9/n$, using the lower bound on $W_{\frac{2}{3}n}$ from property E_1. Thus, the degree of i equals m with probability at least $(1-9/n)^{(1/3)nm} \geq (1-o(1))\exp(-\frac{9nm}{3n}) = (1-o(1))e^{-3m}$.

Thus, the expected number of nodes in C having degree m is at least $(1-o(1))\frac{1}{6}e^{-3m}n$. Since we allowed arbitrary conditioning on other degrees in C, we may apply Chernoff bounds and see that with probability $1-e^{-\Omega(n)}$, at least $\frac{1}{7}e^{-3m}n$ of the nodes in C have degree equal to m.

Each of these nodes was marked independently with probability p. The expected number of marked nodes is at least $\frac{1}{7}e^{-3m}pn$, and with probability $1-e^{-\Omega(n)}$, at least $\frac{1}{8}e^{-3m}pn$ of the nodes in C have degree m and are marked. □

We construct a path from a useful node u to node 1 that has each second node marked. We say a node i is *good* if

$$i \in [s+1, 2^b] \text{ and } w_i \geq 1/(10\sqrt{in}), \tag{5.3}$$

where, as before, $s = 2^a$ is the smallest power of 2 larger than $\log^7 n$ and 2^b is the largest power of 2 smaller than $\frac{2}{3}n$. We consider sets Γ_k and Γ'_k defined recursively as follows. We set $\Gamma_0 = \{u\}$. Given Γ_k, Γ'_k is defined to be the set of all *marked* nodes $i \geq \frac{2}{3}n$ that have a neighbor in Γ_k and have not been included in any Γ'_ℓ with $\ell \leq k-1$. Similarly, Γ_k is defined as the set of all *good* nodes that have a neighbor in Γ'_{k-1} and have not been included in any Γ_ℓ with $\ell \leq k-1$. Note that for all $k \geq 0$, Γ_k only contains nodes $i < \frac{2}{3}n$, while Γ'_k only contains nodes $i \geq \frac{2}{3}n$. This is true for Γ_0 since u is useful and by E_5, all useful nodes are smaller than $n/2$. We define the *weight* of a set Γ_k by

$$f_k := \begin{cases} w_u & \text{if } k = 0 \\ \sum_{i\in\Gamma_k} \frac{1}{\sqrt{in}} & \text{if } k \geq 1. \end{cases} \tag{5.4}$$

Since for $k \geq 1$, Γ_k only contains good nodes, and by definition, $w_u = f_0$, we have for $k \geq 0$,

$$\sum_{i\in\Gamma_k} w_i \geq f_k/10. \tag{5.5}$$

We denote by $N_k = \Gamma_0 \cup \Gamma_1 \cup \dots \cup \Gamma_k$ (note that the Γ'_i are not included). Let $C_0 \subseteq \left[\frac{2}{3}n, n\right]$ be the set of marked nodes and for $k \geq 1$, let $C_k = C_0 \setminus \{\Gamma'_0,\dots,\Gamma'_{k-1}\}$ be the set of marked nodes excluding nodes in $\Gamma'_0, \Gamma'_1, \dots, \Gamma'_{k-1}$. By Lemma 5.12, we have $C_0 \geq \varepsilon_m n$ with probability $1-e^{-\Omega(n)}$. We also need the following technical lemma.

LEMMA 5.13 (BOLLOBÁS AND RIORDAN [5]). *Let $\varepsilon > 0$, and $K := (1/2+\varepsilon)(\log(n)/\log\log(n)) - 1$. Let $f_0, f_1, \dots$ be a sequence of real numbers with $f_0 \geq \log^2(n)/n$ and*

$$f_{k+1} \geq \min\{2\log_2(\varepsilon_m f_k n/\log n) - 29, b-a\}\varepsilon_m f_k/3564 \tag{5.6}$$

for all $k \geq 0$. Then, for n sufficiently large, $\ell = \min\{k\colon f_k \geq \log^3(n)/\sqrt{n}\}$ exists and is at most K.

Note that in the original paper the authors assume $f_{k+1} \geq \min\{2\log_2(f_k n/\log n) - 32, b-a\}f_k/1000$ and obtain that $\ell = \min\{k\colon f_k \geq \log^2(n)/\sqrt{n}\}$ is at most K. It is easy to check that the same proof holds for the above version. Remember that $I_t := [2^t+1, 2^{t+1}]$ for $t \in [a,b)$.

LEMMA 5.14. *Let $k \geq 0$ be such that $f_k \geq \log^2(n)/n$ and $|C_k| \geq \varepsilon_m n/2$. Then given C_k and $\Gamma_0, \Gamma'_0, \Gamma_1, \Gamma'_1, \dots, \Gamma_k$, with prob. $1-O(n^{-6/5})$, one of the following is satisfied:*

- *$|N_{k+1} \cap I_t| \geq 2^{t-2}$, for some $t \in [a,b)$, or*

- $f_{k+1} \geq \min\{2\log_2(\varepsilon_m \frac{f_k n}{\log n}) - 29, b-a\}\varepsilon_m f_k/3564$.

Proof. All probabilities are conditioned on the assumptions in the lemma. We first show that $|\Gamma'_k| = \Omega(nf_k)$ holds with high probability. Let $j \in C_k$. We have

$$\mathbb{P}[\ell_{j,1} \in \Gamma_k] = \Big(\sum_{i\in\Gamma_k} w_i\Big)/\Big(W_j - \sum_{i\in\Gamma_0\cup\Gamma'_0\cup\Gamma_1\cup\cdots\cup\Gamma_{k-1}} w_i - \sum_{i:\ i\in C_k,\, i\leq j} w_i\Big) \geq \Big(\sum_{i\in\Gamma_k} w_i\Big)/W_j \geq f_k/10,$$

where the last inequality follows from (5.5). Hence, $\mathbb{E}[|\Gamma'_k|] \geq |C_k|f_k/10 \geq \varepsilon_m n f_k/20$. By Chernoff's bound, we obtain

$$\mathbb{P}[|\Gamma'_k| \leq \varepsilon_m n f_k/21] \leq \exp(-\Omega(\varepsilon_m n f_k)) \leq n^{-\Omega(\log n)}, \tag{5.7}$$

where the last inequality follows from $f_k \geq \log^2(n)/n$. In the following, we assume

$$|\Gamma'_k| \geq \varepsilon_m n f_k/21. \tag{5.8}$$

We now show that either $|N_{k+1} \cap I_t| \geq 2^{t-2}$ for some $t \in [a,b)$, or with high probability, for sufficiently many $t \in [a,b)$, we have $|\Gamma_{k+1} \cap I_t| = \Omega\big(|\Gamma'_k|\sqrt{2^t/n}\big)$. Let $t \in [a,b)$. By E_2, I_t contains at least 2^{t-1} good nodes. Let S be initially the set of good nodes in $I_t \setminus N_k$. So $|S| \geq 2^{t-1} - |N_k \cap I_t|$. We consider the elements of Γ'_k one by one in any order. Let $\widetilde{\Gamma}$ be an initially empty set. Whenever we encounter some node $i \in \Gamma'_k$ that is connected to some node $j \in S$ via its second link (i.e., $\ell_{i,2} = j$), we remove j from S and include it into $\widetilde{\Gamma}$. Note that $\widetilde{\Gamma} \subseteq \Gamma_{k+1}$ throughout this process. Moreover, $\ell_{i,2}$ has not been revealed before i is considered ($\ell_{i,1}$ is independent from $\ell_{i,2}$ given all W_j values). Hence, as long as $|S| \geq 2^{t-2}$, we have for every node $i \in \Gamma'_k$,

$$\mathbb{P}[\ell_{i,2} \in S] \geq \frac{2^{t-2}}{10W_i\sqrt{2^{t+1}n}} \geq \frac{\sqrt{2^t}}{60\sqrt{n}}.$$

If $|S| < 2^{t-2}$ at some point, then $|N_{k+1} \cap I_t| \geq 2^{t-2}$ and we are finished. Otherwise

$$\mathbb{E}[|\Gamma_{k+1} \cap I_t|] \geq \mathbb{E}[|\widetilde{\Gamma}|] \geq |\Gamma'_k|\frac{\sqrt{2^t}}{60\sqrt{n}} =: \mu_t. \tag{5.9}$$

Let $Y = \sum_{1\leq i\leq|\Gamma'_k|} Y_i$ where $Y_1, \ldots, Y_{|\Gamma'_k|}$ are mutually independent 0/1-random variables with $\mathbb{P}[Y_i = 1] = \frac{\sqrt{2^t}}{60\sqrt{n}}$, where $1 \leq i \leq |\Gamma'_k|$. Note that $|\Gamma_{k+1} \cap I_t|$ stochastically dominates Y, i.e., $\mathbb{P}[|\Gamma_{k+1} \cap I_t| > t] \geq \mathbb{P}[Y > t]$. Now, if $\mu_t \geq 10\log n$, we have by Chernoff's bound $\mathbb{P}[Y \leq \mu_t/2] \leq n^{-5/4}$, and thus $\mathbb{P}[|\Gamma_{k+1} \cap I_t| \leq \mu_t/2] \leq \mathbb{P}[Y \leq \mu_t/2] \leq n^{-5/4}$. Taking a union bound over all indices $t \in [a,b)$ with $\mu_t \geq 10\log n$, we conclude that with probability at least $1 - \log_2(n)\cdot n^{-5/4} = 1 - \mathcal{O}(n^{-6/5})$, we have for all these $t \in [a,b)$,

$$\sum_{i\in\Gamma_{k+1}\cap I_t} \frac{1}{\sqrt{in}} \geq \frac{\mu_t}{2}\cdot\frac{1}{\sqrt{2^{t+1}n}} = \frac{|\Gamma'_k|}{120n\sqrt{2}} \geq \frac{\varepsilon_m f_k}{3564},$$

where the last inequality follows from (5.8). Let $\mathcal{T}$ be the set of indices $t \in [a,b)$ with $\mu_t \geq 10\log n$. Since $2^b \geq n/3$, we have $|\mathcal{T}| \geq \min\{2\log_2(\varepsilon_m f_k n/\log n) - 29, b-a\}$. Therefore $f_{k+1} \geq \min\{2\log_2(\varepsilon_m f_k n/\log n) - 29, b-a\}\varepsilon_m f_k/3564$. The failure probability is $n^{-\Omega(\log n)} + O(n^{-6/5}) = O(n^{-6/5})$. □

Proof of Lemma 5.11. We apply Lemma 5.14 consecutively for $k = 0, \ldots, K$, where $K = (1/2+\varepsilon)\log(n)/\log\log n$. The probability that the event considered in Lemma 5.14 holds for all $k = 0, \ldots, K$ is at least $1 - O(K\,n^{-6/5}) = 1 - O(n^{-7/6})$. In the following we assume this is the case.

Note that for $k = 0$, $f_k = w_u \geq \log^2(n)/n$ holds since u is useful. Also, by Lemma 5.12, we have $|C_0| \geq \varepsilon_m n/2$ with probability $1 - e^{-\Omega(n)}$. Assume that this is the case. Hence, we can apply Lemma 5.14 for $k = 0$. Since $\min\{2\log_2(\varepsilon_m f_k n/\log n) - 29, b-a\}\varepsilon_m f_k/3564 \geq f_k$ (for large enough n), the only way we fail to apply Lemma 5.14 for some k', where $0 < k' < K$, is when $|C_{k'}| < \varepsilon_m n/2$ or $|N_{k'} \cap I_t| \geq 2^{t-2}$ for some $t \in [a,b)$.

If $|C_{k'}| < \varepsilon_m n/2$, then there must be a k'', $0 \leq k'' \leq k'$, with $|\Gamma'_{k''}| \geq \frac{\varepsilon_m n/2}{K} \geq \frac{n}{\log n}$ for n sufficiently large. We stop the sequence at $\Gamma'_{k''}$ as soon as we encounter such a k''. Given the sequence $\Gamma_0, \Gamma'_0, \Gamma_1, \ldots, \Gamma'_{k''}$, the second links of the nodes in $\Gamma'_{k''}$ are mutually independent random variables. So the probability that no node in $\Gamma'_{k''}$ connects to 1 via its second link is at most

$$(1-w_1)^{|\Gamma'_{k''}|} \leq \big(1 - \tfrac{4}{\log(n)\sqrt{n}}\big)^{n/\log n} \leq \exp\big(-4\sqrt{n}/\log^2 n\big) = n^{-\Omega(\sqrt{n}/\log^3 n)},$$

where the first inequality follows from E_3. We can hence assume that $|C_k| \geq \varepsilon_m n/2$ for all $k = 0, \ldots, K-1$.

Similarly, if $|N_{k'} \cap I_t| \geq 2^{t-2}$ for some $t \in [a,b)$, there must be a k'' where $0 < k'' \leq k'$, with $|\Gamma_{k''} \cap I_t| \geq 2^{t-2}/K$. We stop the construction of the sequence $\Gamma_0, \Gamma'_0, \Gamma_1, \ldots$ at $\Gamma'_{k''}$. By (5.7) we have with probability $1 - n^{-\Omega(\log n)}$,

$$\begin{aligned}|\Gamma'_{k''}| &\geq \varepsilon_m n f_{k''}/21 \geq \varepsilon_m n|\Gamma_{k''} \cap I_t|/(21\sqrt{2^{t+1}n}) \\ &\geq \varepsilon_m n 2^{t-2}/(21K\sqrt{2^{t+1}n}) \geq \varepsilon_m 2^{t/2}\sqrt{n}/\log n \\ &\geq \varepsilon_m \log^{5/2}(n)\sqrt{n},\end{aligned}$$

where the last inequality follows from $2^t \geq \log^7 n$. So given $\Gamma'_0, \Gamma_1, \ldots, \Gamma'_{k''}$, the probability that no node in $\Gamma'_{k''}$ connects to 1 by its second link is at most

$$(1-w_1)^{|\Gamma'_{k''}|} \leq \big(1 - \tfrac{4}{\log(n)\sqrt{n}}\big)^{\varepsilon_m \log^{5/2}(n)\sqrt{n}} \leq \exp\Big(-4\varepsilon_m \log^{3/2} n\Big) \leq n^{-\Omega(\sqrt{\log n})},$$

where the last inequality holds since ε_m is a constant.

So assume now that $|C_k| \geq \varepsilon_m n/2$ and $f_{k+1} \geq \min\{2\log_2(\varepsilon_m f_k n/\log n) - 29, b-a\}\varepsilon_m f_k/3564 \geq f_k$, for all k, $0 \leq k < K$, where $K = \big(\frac{1}{2}+\varepsilon\big)\frac{\log n}{\log\log n}$. Then, by Lemma 5.13, we have $f_\ell \geq \log^3(n)/\sqrt{n}$, for some $\ell \leq K$. Again, by (5.7), we have $|\Gamma'_\ell| \geq \varepsilon_m n f_\ell/21 \geq \varepsilon_m\sqrt{n}\log^3(n)/21$ with probability $1 - n^{-\Omega(\log n)}$. Furthermore, given $\Gamma_0, \Gamma'_0, \Gamma_1, \ldots, \Gamma'_\ell$, the probability that no node in Γ'_ℓ connects to 1 by its second link is at most

$$(1-w_1)^{|\Gamma'_\ell|} \leq \big(1 - \tfrac{4}{\log(n)\sqrt{n}}\big)^{\varepsilon_m\sqrt{n}\log^3(n)/21} \leq \exp\big(-\tfrac{4\varepsilon_m}{21}\log^2 n\big) \leq n^{-\Omega(\log n)}.$$

The total failure probability is $O(n^{-7/6}) + e^{-\Omega(n)} + n^{-\Omega(\sqrt{n}/\log^3 n)} + n^{-\Omega(\log n)} + n^{-\Omega(\sqrt{\log n})} + n^{-\Omega(\log n)} = O(n^{-7/6})$. □

We can now use Lemma 5.11 to show that the rumor quickly proceeds from a useful node to node 1.

COROLLARY 5.15. *Let $W_1, \ldots, W_n$ be s.t. $E_1, \ldots, E_5$ are satisfied. Let G be a random graph from $G_m(W_1, \ldots, W_n)$.*

Let $v \in [n]$ be a useful node. With probability $1 - o(n^{-1})$, using the push-pull protocol with memory $M \geq 0$, a rumor present at v reaches node 1 in $\mathcal{O}(\log n / \log\log n)$ steps.

Proof. Consider a run of the process started with a rumor in u. Let $v \in \left[\frac{2}{3}n, n\right] =: C$ be a node of degree m. Note that $u \neq v$, since u is useful and thus $u \leq n/2$ by E_5.

The probability that v is fast is exactly $p := \frac{1}{m}(m-1)!/(m-1)^{m-1}$. This remains true if we condition arbitrarily on random decisions of other nodes during the run of the process. In consequence, the set of fast nodes is a random subset of the nodes of degree m in C with each such node being included independently with probability p.

Applying Lemma 5.11 with fast nodes being marked, we see that with probability $1 - o(n^{-1})$, there is a path of length $\mathcal{O}(\log(n)/\log\log n)$ such that every second node is fast. Even ignoring all rumor transmissions by nodes that are not fast, the rumor is propagated along the path in $\mathcal{O}(\log(n)/\log\log n)$ time steps. □

5.5 Informing all nodes

The following lemma allows us to invert the spread of the rumor: from node 1 to all other nodes. Note that for $M = 0$ the lemma has been shown in [9].

LEMMA 5.16. *Assume for the push-pull protocol with memory $M \geq 0$ that if the rumor starts in node u, it reaches node v in k rounds with probability p. This implies the reverse statement: if the rumor is initiated by v, then it reaches u in k rounds with probability p.*

Proof of Lemma 5.16. We define a *snapshot* of the process to be a vector $(s_1, \ldots, s_n)$ where s_i denotes a neighbor of i. We interpret a snapshot as the ordered set of nodes that were contacted in one time step. Thus, we can represent a (possibly infeasible) run of the process by a (finite) vector of snapshots $S = (S_1, \ldots, S_\ell)$. Moreover, a run $S = (S_1, \ldots, S_\ell)$ is feasible if and only if for all $i > 1$, $k \in [n]$, and $j \in [1, \min\{\deg(k) - 1, M\}]$, we have $(S_i)_k \neq (S_{i-j})_k$.

Let $\mathcal{S}$ be the set of all series S of snapshots such that if the rumor starts in node u and the process follows S, then node v is informed for the first time after $|S|$ steps. Since v is informed with probability one after finitely many steps, $\mathcal{S}$ naturally defines a probability space: the probability of a series $S \in \mathcal{S}$ is the probability that the process follows S (infeasible series have probability 0). Note that the probability that node v is informed in k steps is just $\sum_{\substack{S \in \mathcal{S} \\ |S| = k}} \mathbb{P}[S]$.

We give an automorphism $\phi\colon \mathcal{S} \to \mathcal{S}$ s.t. for all $S \in \mathcal{S}$,

(i) $|S| = |\phi(S)|$,

(ii) $\mathbb{P}[S] = \mathbb{P}[\phi(S)]$,

(iii) if the rumor starts in node v and the process follows $\phi(S)$, then node u is informed after $|\phi(S)| = |S|$ steps.

Thus if the rumor is initiated from node v, then the probability that node u is informed in k steps is $\sum_{\substack{S \in \mathcal{S} \\ |S| = k}} \mathbb{P}[S]$. Equality then follows by a symmetric argument.

The automorphism is defined as follows. For $S = (S_1, \ldots, S_\ell)$, let $\phi(S) = (S_\ell, S_{\ell-1}, \ldots, S_1)$, i.e., ϕ simply inverts the series. It remains to check that all three properties are indeed satisfied. Property (i) and (iii) are immediate. For (ii), there exist real numbers $\rho_0, \rho_1, \ldots, \rho_M \in [0, 1]$, such that for all series, we have $\mathbb{P}[S] = \rho_0 \rho_1 \cdots \rho_{M-1} \rho_M^{|S|-M}$ if S is feasible, and $\mathbb{P}[S] = 0$ otherwise. Since S is feasible iff $\phi(S)$ is feasible, (ii) follows. □

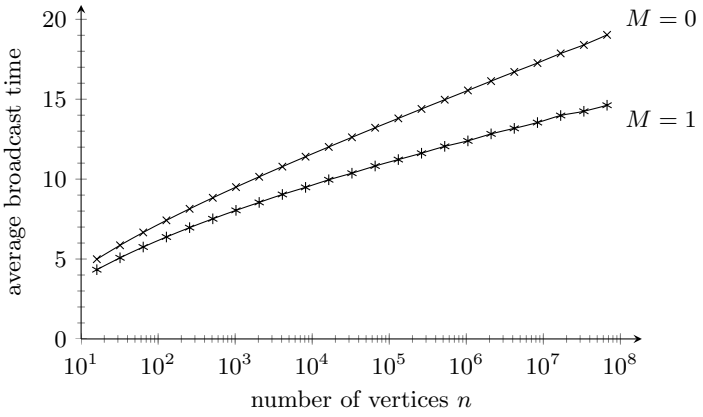

Figure 1: Empirical comparison of the classic push-pull strategy (corresponding to $M = 0$), and the push-pull strategy with $M = 1$ for preferential attachment graphs G_m^n with $m = 2$. The observed average runtimes are consistent with the proven asymptotic bounds of $\Theta(\log n)$ for $M = 0$ and $\Theta(\log n / \log\log n)$ for $M = 1$.

Proof of Theorem 3.1. By Lemmas 5.1, 5.2, 5.3, and 5.4, assumptions $E_1, \ldots, E_8$ hold with probability $1 - o(1)$. Hence we can assume that this is the case. By Lemma 5.8 and Lemma 5.10, with probability $1 - o(1)$, we have for all nodes v that a rumor initiated by v reaches a useful node u in $\mathcal{O}(\log^{3/4}(n) \log\log n)$ time steps. By Corollary 5.15, a rumor starting from a useful node u reaches node 1 in $\mathcal{O}(\log(n)/\log\log(n))$ time steps with probability $1 - o(n^{-1})$.

Corollary 5.15, Lemma 5.16, and a simple union bound show that with probability $1 - o(1)$ after another $\mathcal{O}(\log(n)/\log\log n)$ time steps, all useful nodes are informed. Similarly, Lemmas 5.10 and 5.16 together with a union bound prove that another $\mathcal{O}(\log^{3/4}(n) \log\log n)$ rounds suffice to inform all nodes with probability $1 - o(1)$.

The total failure probability is bounded by the sum of the probability that anyone of $E_1, \ldots, E_8$ does not hold and the probability that a node does not get informed within $\mathcal{O}(\log(n)/\log\log n)$ time steps conditioned on $E_1, \ldots, E_8$. Since both probabilities are $o(1)$, the result follows. □

The proof of Theorem 3.2 is equivalent except that it uses Lemma 5.9 instead of Lemma 5.10.

6. DISCUSSION

We have shown that for preferential attachment graphs the classic push-pull strategy needs $\Theta(\log n)$ rounds to inform all vertices. The slightly improved version which avoids that a vertex contacts the same neighbor twice in a row only needs $\Theta(\log n / \log\log n)$ rounds, which is best possible since the diameter is of the same order of magnitude.

In order to show that this asymptotic speed-up is visible for preferential attachment graphs of realistic size, we have implemented both protocols and empirically determined the average time needed for graphs G_m^n with $m = 2$ and $n = 2^4, 2^5, \ldots, 2^{26}$. Figure 1 shows the averages over ≥ 1000

runs ($\geq 100,000$ runs for $n \leq 2^{22}$). It is clearly visible that the push-pull strategy with memory $M = 1$ is significantly faster. More precisely, for $n = 2^{26}$ the classic push-pull strategy ($M = 0$) needs 19.02 ± 0.79 rounds, while the push-pull strategy with memory $M = 1$ needs only 14.61 ± 0.52 rounds. Interestingly, adding more memory gains very little. The push-pull strategy with $M = 2$ finishes after 14.18 ± 0.42 rounds, the one with $M = 3, 4, 5$ gives 14.12 ± 0.34.

Acknowledgments

We are thankful to the reviewers for helpful comments that improved the presentation of the results significantly. The first author was partially supported by the German Science Foundation (DFG) via grant DO 749/4.

References

[1] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286:509–512, 1999.

[2] N. Berger, C. Borgs, J. T. Chayes, and A. Saberi. On the spread of viruses on the internet. In *16th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 301–310, 2005.

[3] B. Bollobás and O. Riordan. Robustness and vulnerability of scale-free random graphs. *Internet Mathematics*, 1:1–35, 2003.

[4] B. Bollobás and O. Riordan. Coupling scale-free and classical random graphs. *Internet Mathematics*, 1:215–225, 2003.

[5] B. Bollobás and O. Riordan. The diameter of a scale-free random graph. *Combinatorica*, 24:5–34, 2004.

[6] B. Bollobás, O. Riordan, J. Spencer, and G. Tusnády. The degree sequence of a scale-free random graph process. *Rand. Struct. & Algo.*, 18:279–290, 2001.

[7] Business Insider. Facebook has more than 600 million users, Goldman tells clients, Jan. 2011. `http://read.bi/eIuu3L`.

[8] F. Chierichetti, S. Lattanzi, and A. Panconesi. Rumor spreading in social networks. In *36th International Colloquium on Automata, Languages and Programming (ICALP)*, pages 375–386. Springer, 2009.

[9] F. Chierichetti, S. Lattanzi, and A. Panconesi. Almost tight bounds for rumour spreading with conductance. In *42nd ACM Symposium on Theory of Computing (STOC)*, pages 399–408, 2010.

[10] C. Cooper and A. M. Frieze. Crawling on web graphs. In *34th ACM Symposium on Theory of Computing (STOC)*, pages 419–427, 2002.

[11] C. Cooper and A. M. Frieze. The cover time of the preferential attachment graph. *Journal of Combinatorial Theory, Series B*, 97:269–290, 2007.

[12] A. J. Demers, D. H. Greene, C. Hauser, W. Irish, J. Larson, S. Shenker, H. E. Sturgis, D. C. Swinehart, and D. B. Terry. Epidemic algorithms for replicated database maintenance. *Operating Systems Review*, 22: 8–32, 1988.

[13] B. Doerr, T. Friedrich, and T. Sauerwald. Quasirandom rumor spreading: Expanders, push vs. pull, and robustness. In *36th Inter. Colloq. on Automata, Languages and Programming (ICALP)*, pages 366–377, 2009.

[14] R. Elsässer. On the communication complexity of randomized broadcasting in random-like graphs. In *18th ACM Symposium on Parallelism in Algorithms and Architectures (SPAA)*, pages 148–157, 2006.

[15] R. Elsässer and T. Sauerwald. On the power of memory in randomized broadcasting. In *19th ACM-SIAM Symposium on Discrete Algorithms*, pages 218–227, 2008.

[16] U. Feige, D. Peleg, P. Raghavan, and E. Upfal. Randomized broadcast in networks. *Random Structures and Algorithms*, 1:447–460, 1990.

[17] A. D. Flaxman, A. M. Frieze, and J. Vera. Adversarial deletion in a scale-free random graph process. *Combinatorics, Probability & Computing*, 16:261–270, 2007.

[18] N. Fountoulakis and K. Panagiotou. Rumor spreading on random regular graphs and expanders. In *14th Inter. Workshop on Randomization and Comput. (RANDOM)*, volume 6302 of *LNCS*, pages 560–573, 2010.

[19] A. M. Frieze and G. R. Grimmett. The shortest-path problem for graphs with random arc-lengths. *Discrete Applied Mathematics*, 10:57–77, 1985.

[20] A. M. Frieze and M. Molloy. Broadcasting in random graphs. *Discrete Applied Mathematics*, 54:77–79, 1994.

[21] C. Gkantsidis, M. Mihail, and A. Saberi. Conductance and congestion in power law graphs. In *ACM Inter. Conference on Measurement and Modeling of Computer Systems (SIGMETRICS)*, pages 148–159, 2003.

[22] R. Karp, C. Schindelhauer, S. Shenker, and B. Vöcking. Randomized rumor spreading. In *41st IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 565–574, 2000.

[23] J. Krums. Twitter message, 2009. `http://twitpic.com/135xa`.

[24] J. Leskovec, K. J. Lang, A. Dasgupta, and M. W. Mahoney. Statistical properties of community structure in large social and information networks. In *17th International conference on World Wide Web (WWW)*, pages 695–704, 2008.

[25] M. Mihail, C. Papadimitriou, and A. Saberi. On certain connectivity properties of the internet topology. *J. Comput. Syst. Sci.*, 72:239–251, 2006.

[26] B. Pittel. On spreading a rumor. *SIAM Journal on Applied Mathematics*, 47:213–223, 1987.

[27] T. Sauerwald. On mixing and edge expansion properties in randomized broadcasting. *Algorithmica*, 56:51–88, 2010.