Social Networks Spread Rumors in Sublogarithmic Time

Benjamin Doerr

Max-Planck-Institut Informatik, Germany

Mahmoud Fouz

Universität des Saarlandes, Germany

Tobias Friedrich

Max-Planck-Institut Informatik, Germany

Abstract

It has been observed that information spreads extremely fast in social networks. We model social networks with the preferential attachment model of Barabási and Albert (Science 1999) and information spreading with the random phone call model of Karp et al. (FOCS 2000). In a recent paper (STOC 2011), we prove the following two results. (i) The random phone call model delivers a message to all nodes of graphs in the preferential attachment model within $\Theta(\log n)$ rounds with high probability. The best known bound so far was $O(\log^2 n)$. (ii) 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.

Keywords: rumor spreading, preferential attachment, randomized algorithms

1 Introduction

The social network service *Facebook* is currently the most-visited website on
the internet with more than 15 million page views per minute. Online social
networks have become an important part of the daily life of millions of people.
An important form of social interaction is spreading information. We are
interested in how fast a rumor spreads in social networks. This fundamental
property has obvious applications in viral advertising, social marketing, online
political campaigns, and peer-to-peer 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 [2–4] 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.

   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 to all n
vertices of a graph. In contrast to this, Chierichetti, Lattanzi, and Panconesi
[5] showed that this 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*. Here, each vertex in each
round contacts a random neighbor and learns the rumor if its contact knows
the rumor already. There is a certain symmetry between the two models.
This was observed for a quasirandom version of the two models in [9], but
similar arguments also hold for the two random models discussed so far. In
consequence, the above results also hold for the pull model.

   Karp, Schindelhauer, Shenker, and Vöcking [13] pointed out that for com-
plete 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-pull strat-
egy* 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 [8, 11, 13]. Elsässer [11]

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. [5] showed that with this strategy, $O(\log^2 n)$ rounds suffice.

So far it was 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 can be bounded by $O(\Phi^{-1} \log^2(\Phi^{-1}) \log n)$ [6]. Unfortunately, the conductance of the preferential attachment model seems not known. Several power law graphs have a conductance of $\Phi = \Omega(\log^{-1} n)$ [7, 12] and this has also been observed empirically for real social networks [14]. Mihail, Papadimitriou, and Saberi [15] showed that certain graphs that are very similar to PA-graphs have constant conductance. If this was true also for the true PA-model, a bound of $O(\log n)$ would follow.

Our results: We prove in [10] 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 $O(\log n/\log\log n)$ rounds suffice (cf. Theorem 3.1). This is asymptotically optimal as the diameter of a PA-graph is $\Theta(\log n/\log\log n)$ [4]. This result can be seen as an explanation why rumor spreading in actual social networks is extremely fast.

2 Precise Model and Preliminaries

Preferential attachment (PA) graphs were first introduced by Barabási and Albert [1]. We follow the formal definition of Bollobás et al. [3, 4]. 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 $V := \{1, \ldots, n\}$ 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 $e_n^1 = \{n, v_1\}, \ldots, e_n^m = \{n, v_m\}$ 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 [4].

By bounding the probability that the graph contains no loops, 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 [3] showed that the diameter is only $\Theta(\log n / \log \log n)$.

We examine the following broadcasting protocol.

Definition 2.2 Let $M \geq 0$ be a fixed parameter. Assume that every vertex can store the last M vertices it contacted. 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.

3   Statement of Results

Our main result is that PA 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 $O(\log n / \log \log n)$ rounds.*

Our proof uses several arguments of Bollobás and Riordan [3] 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. 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 1 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. We first analyze the time needed until the rumor reaches a *useful node*. Roughly speaking, a node

is useful if its degree is at least polylogarithmic. We show that for $M \geq 2$, a useful node is reached in only $O(\log \log n)$ rounds. By a more involved argument, we also show that for $M = 1$, $O(\log^{3/4}(n) \log \log n)$ rounds suffice.

The core of the proof consists of showing that once a useful node u has been informed, within $O(\log n / \log \log n)$ 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. [5] that social networks have a connected subgraph of linear size and diameter $O(\log n)$ in which every node has degree $O(\log n)$. Using this property, the authors showed a running time of $O(\log^2 n)$ for the classic push-pull protocol.

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

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

Theorem 3.2 *The classic push-pull protocol broadcasts a rumor from any node of G_m^n to all other nodes in $O(\log n)$ rounds with probability $1 - o(1)$.*

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

References

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

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

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

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

[5] 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.

[6] 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.

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

[8] 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.

[9] 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.

[10] B. Doerr, M. Fouz, and T. Friedrich. Social networks spread rumors in sublogarithmic time. In *43rd ACM Symposium on Theory of Computing (STOC)*, 2011. To appear.

[11] 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.

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

[13] 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.

[14] 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.

[15] M. Mihail, C. Papadimitriou, and A. Saberi. On certain connectivity properties of the internet topology. *Journal of Computer and System Sciences*, 72:239–251, 2006.