

Exact Performance Measures for Peer-to-Peer Epidemic Information Diffusion

Öznur Özkasap^{1,*}, Emine Şule Yazıcı², Selda Küçükçifçi^{2,**}, and Mine Çağlar²

¹ Department of Computer Engineering

² Department of Mathematics

Koç University, Istanbul, Turkey

{oozkasap, eyazici, skucukcifci, mcaglar}@ku.edu.tr

Abstract. We consider peer-to-peer anti-entropy paradigms for epidemic information diffusion, namely pull, push and hybrid cases, and provide exact performance measures for them. Major benefits of the proposed epidemic algorithms are that they are fully distributed, utilize local information only via pair-wise interactions, and provide eventual consistency, scalability and communication topology-independence. Our contribution is the derivation of exact expressions for infection probabilities through elaborated counting techniques on a digraph. Considering the first passage times of a Markov chain based on these probabilities, we find the expected message delay experienced by each peer and its overall mean as a function of initial number of infectious peers. In terms of these criteria, the hybrid approach outperforms pull and push paradigms, and push is better than the pull case. Such theoretical results would be beneficial when integrating the models in several peer-to-peer distributed application scenarios.

Keywords: peer-to-peer, epidemic, anti-entropy, counting, Markov chain.

1 Introduction

An efficient approach for information diffusion in distributed systems is to utilize epidemic algorithms that involve pair-wise propagation of updates. Epidemic algorithms are fully distributed and randomized approaches such that every peer in an information diffusion session picks a (subset of the other) peer(s) randomly for efficient propagation of updates, that happens through periodic rounds. The underlying epidemics theory for the biological systems study the spreading of infectious diseases through a population [1]. When applied to an information diffusion application, such protocols have beneficial features such as scalability, robustness against failures and provision of eventual consistency. We investigate variations of the epidemic algorithms used in the context of distributed information diffusion and derive exact performance measures for them.

* Research supported by TUBITAK (The Scientific and Technical Research Council of Turkey) under CAREER Award Grant 104E064.

** Research partly supported by TUBA (Turkish Academy of Sciences).

One of the first studies that applies epidemic methods to computer systems used the idea for spreading updates in a replicated database [2]. Several succeeding work utilized epidemic (or sometimes so called gossip-style, to reflect rumor propagation in a social network) communication in a variety of contexts such as large-scale direct mail systems [3], group membership tracking [4], support for replicated services [5], message garbage collection [6], failure detection [7], loss recovery in reliable multicast [8], and distributed information management [9]. [10] gives an overview of epidemic information dissemination where the focus is on four design constraints namely, membership, network awareness, buffer management, and message filtering.

There are different classes of epidemic processes one of which is referred to as anti-entropy. In anti-entropy, information diffusion progresses periodically via fixed length time periods, typically larger than the maximum round trip time between peers, and called rounds. In each round, every peer picks another site at random, and shares its state information. We study three approaches for update-exchange; namely, pull, push and hybrid cases as particular models of anti-entropy. When a peer holding information to be shared picks a peer lacking that information and if this triggers dissemination from the holding peer to the lacking peer, we say the dissemination occurs in pull fashion. In contrast, if information dissemination is triggered when a peer lacking a particular information picks a peer holding that information, it has occurred in push fashion. The third approach is a hybrid of the two where information diffusion takes place in both ways. In [11], pull and push anti-entropy approaches are compared through a binomial probability distribution for information flow where push approach is shown to be superior in terms of message latency.

In this study, we develop exact performance measures for the pull, push and hybrid information diffusion models of anti-entropy. The number of peers lacking information to be shared forms a Markov chain advancing in rounds. The transition probabilities on the chain are derived through elaborated counting techniques on a digraph exactly, with no resort to approximate probability distributions that rely on several independence assumptions. Indeed, we show that the binomial model for the pull approach in [11] is only approximate, whereas it is exact for the push approach. We analyze the exact dynamics through the Markov chain with no assumptions on any parameters. In particular, we compute the expected delay of each peer as well as per arbitrary peer with respect to the initial number of members holding information in the population. The hybrid approach outperforms pull and push paradigms, and push is better than the pull case. Such results would be beneficial when integrating the models in several distributed scenarios such as replicated servers, loss recovery, failure detection and group membership management.

Previously, exact as well as asymptotical distributions have been studied for different epidemic models. In [12, 13], the epidemic process is defined on a random graph. In [14], the infection is spread through random contact in a manner less structured than a random graph and simpler than anti-entropy.

The paper is organized as follows. In Sect. 2, our models for the pull, push and hybrid anti-entropy are explained. The exact probability distributions are derived in Sect. 3. In Sect. 4, the Markov chain formulation and delay computations are given. Finally, Sect. 5 includes the conclusions and future work.

2 Model Descriptions

A popular distribution model based on the theory of epidemics is the anti-entropy [1]. According to the terminology of epidemiology, a peer holding information or an update it is willing to share is called *infectious*. A peer is called *susceptible* if it has not yet received an update. In the anti-entropy process, non-faulty peers are always either susceptible or infectious. Data diffusion progresses periodically via rounds of epidemics. In each round, every peer picks another site at random, and exchanges its state information with the selected one. We study the following approaches for update-exchange that execute in a fully distributed manner.

Pull Approach: When an infectious peer (holding data to be shared) picks a susceptible peer (lacking the specific data) randomly, this triggers data dissemination from infectious peer to the susceptible. Steps involved in the dissemination between two such peers is depicted in Fig.1(a) where infectious peer (on the left) has data labelled A. The infectious peer sends a digest (also referred to as gossip) message including its state information. On receiving digest and comparing it with its local data, the susceptible peer finds out it lacks A and sends a request for A back to the infectious. Upon getting request, infectious peer sends a retransmission of data A which causes the other peer to be infectious for A. In fact, each peer in the system performs state exchange periodically and concurrently with the others. Moreover, each peer may have a set of data in its local buffer. Therefore, a digest message generated by a peer would consist of the state information on the current contents of its message buffer. Spreading updates is triggered by susceptible peers when they are picked as gossip destinations by infectious peers.

Push Approach: If a susceptible peer picks an infectious peer randomly, and sends its state information, this triggers information dissemination from infectious peer to the susceptible. Steps involved in the dissemination between two such peers is depicted in Fig.1(b) where infectious peer (on the left) has data labelled A. The infectious peer on receiving digest and comparing it with its local data finds out that the digest owner lacks A and directly retransmits, or pushes data A which causes the other peer to become infectious for A. As illustrated in the figure, in the push approach, no request messages are used. Spreading updates is triggered by infectious peers when they are selected as gossip targets by susceptible peers.

Hybrid Approach: This is a hybrid of two approaches described above. As illustrated in Fig.1(c), when a peer sends its digest to a randomly selected peer in the population, this may trigger data dissemination at both peers. Consider the case where a peer has data A and the other has data B. When the former

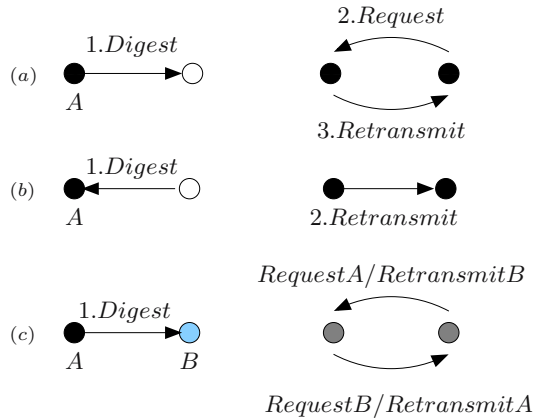


Fig. 1. Model Descriptions

selects the latter as the digest target in a given round, information A and B would be disseminated to the peer that lack it using pull-based or push-based approaches together. This approach is useful for delay sensitive applications since it decreases overall delay during data dissemination at the cost of possible duplicate data transmissions.

3 Exact Diffusion Probabilities

In this section, we will restrict our attention to the processes of distributing a single data message. Therefore, a peer with a copy of the data message is referred to as infectious; otherwise, it would be susceptible. Each step of this diffusion process can be represented by a digraph D where a peer in the population corresponds to a node of the digraph (We will assume there are n peers in the population). If the node u chooses to communicate with the node v then there will be an arc with the tail u and the head v in D . Since each node chooses exactly one node at each step the out degree of each node will be 1 in D . The number of all possible such digraphs with n nodes is $(n - 1)^n$. All of these digraphs are equally likely for each step of this process. Therefore, we will count the number of digraphs that infect i more nodes and take the ratio of this number with the number of all possible digraphs to find the probability of infecting i more nodes. Note that if there are k infectious nodes, after one step there will be $k + i$ infectious nodes, where $k + i = 1, 2, \dots, n$.

Let S be the set of all susceptible nodes and I be the set of all infectious nodes with $|I| = k$ and $|S| = n - k$. For simplicity, we will denote arcs with susceptible heads and infectious tails by IS -arcs, similarly arcs with infectious heads and susceptible tails, infectious heads and infectious tails, and susceptible heads and susceptible tails will be represented by SI -arcs, II -arcs, and SS -arcs, respectively. Note that D is the disjoint union of four subgraphs formed by IS -arcs, SI -arcs, II -arcs, and SS -arcs. Let $S(n, k)$ denote the Stirling number of

second kind defined as the number of all partitions of an n -element set into k nonempty subsets. For further information on these numbers see [15].

Pull Case: We form the digraph D as above. In the pull case, a susceptible node s will be infected if and only if there exists a IS -arc in D with the head s . Therefore, SI -arcs, II -arcs, and SS -arcs will not contribute to the number of new infectious nodes. Fig.2 (a) illustrates the pull case. We will determine the number digraphs representing a step that results in i more infectious nodes.

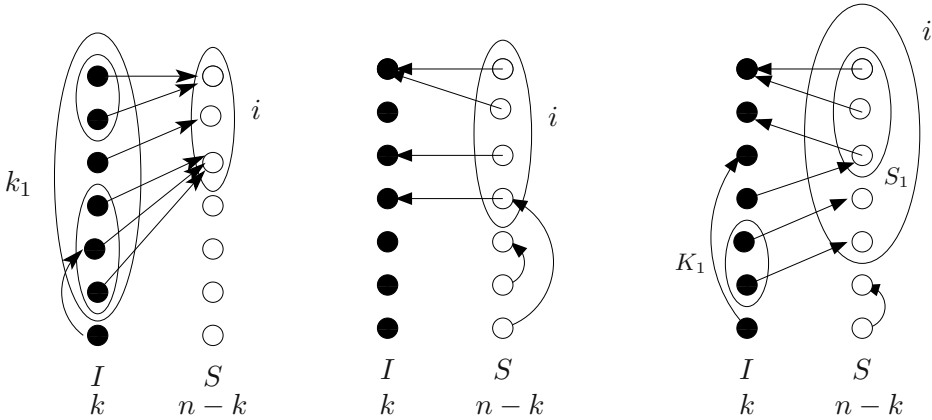


Fig. 2. (a) Pull Case (b) Push Case (c) Hybrid Case

The number of different possible subgraphs formed by SI -arcs and SS -arcs is $(n - 1)^{n-k}$, since each of the $n - k$ nodes in S can be adjacent to $n - 1$ different nodes.

Now we need to count the number of different possible subgraphs that can be formed by IS -arcs and II -arcs. Let k_1 be the number of IS -arcs. Note that k_1 has to be at least i since each IS -arc infects at most one new node in S also there are $\binom{k}{k_1}$ such k_1 -subsets of I . We have $k - k_1$ II -arcs. The number of different possible subgraphs formed by these arcs is $(k - 1)^{k-k_1}$. Finally we will count the number of different subgraphs that can be formed by IS -arcs. Among $n - k$ susceptible nodes there are $\binom{n-k}{i}$ different i -subsets of S that may be infected. There are $S(k_1, i)i!$ different ways for k_1 nodes to infect exactly i new nodes. So Therefore, the number of different subgraphs that can be formed by IS -arcs and II -arcs is $\sum_{k_1=i}^k \binom{k}{k_1} (k - 1)^{k-k_1} \binom{n-k}{i} S(k_1, i)i!$. Hence, the probability of infecting i more nodes in the next step given $|I| = k$ is

$$p(i|k) = \frac{(n - 1)^{n-k} \binom{n-k}{i} i! \sum_{k_1=i}^k \binom{k}{k_1} (k - 1)^{k-k_1} S(k_1, i)}{(n - 1)^n}$$

$$= \frac{\binom{n-k}{i} i! \sum_{k_1=i}^k \binom{k}{k_1} (k-1)^{k-k_1} S(k_1, i)}{(n-1)^k}$$

where $k = 2, 3, \dots, n-1$ and $i = 0, 1, \dots, n-k$.

For the case $k = 1$, we can easily see that $p(0|1) = 0$ and $p(1|1) = 1$.

In [11], a binomial probability distribution model is assumed for the infection probabilities at each step. It turns out that this is only approximate since the distribution found above is not binomial.

Push Case: In the push case, a susceptible node s will be infected if and only if there exists a SI -arc with the tail s . Therefore, IS -arcs, II -arcs, and SS -arcs will not contribute to the number of new infectious nodes.

Fig.2 (b) illustrates the push case. The number different possible subgraphs formed by IS -arcs and II -arcs is $(n-1)^k$. Since i new nodes will be infected there are i SI -arcs and $\binom{n-k}{i}$ different i -subsets of S . For each SI -arc there are k different choices for the head of the arc, therefore there are $\binom{n-k}{i} k^i$ different possible subgraphs formed by these arcs. Finally, as there are $n-k-i$ SS -arcs, the number different possible subgraphs formed by SS -arcs is $(n-k-1)^{n-k-i}$. Hence, the probability of infecting i more nodes after the step given $|I| = k$ is

$$p(i|k) = \frac{(n-1)^k \binom{n-k}{i} k^i (n-k-1)^{n-k-i}}{(n-1)^n} = \frac{\binom{n-k}{i} k^i (n-k-1)^{n-k-i}}{(n-1)^{n-k}}$$

where $k = 1, 2, \dots, n-2$ and $i = 0, 1, \dots, n-k$.

Clearly, when $k = n-1$, we get $p(0|k) = 0$ and $p(1|k) = 1$. The probability distribution above can be rewritten as

$$\binom{n-k}{i} \left(\frac{k}{n-1}\right)^i \left(\frac{n-k-1}{n-1}\right)^{n-k-i}$$

which can now be recognized as binomial distribution with parameters $n-k$ and success probability $k/(n-1)$. This coincides with the distribution modeled in [11] through probabilistic arguments. The only difference is that the number of possible nodes among which an infectious node chooses to communicate is rounded as n in [11]. In fact, it is $n-1$ as given in the present analysis.

Hybrid Case: In the hybrid case, a susceptible node s will be infected if and only if there exists either a SI -arc with the tail s or an IS -arc with the head s . Therefore, II -arcs and SS -arcs will not contribute to the number of new infectious nodes.

Fig.2 (c) illustrates the hybrid case. There are i new infectious nodes and $\binom{n-k}{i}$ different i -subsets of S . Let S_1 be the set of the tails of SI -arcs where $|S_1| = i_1$. There are $\binom{i}{i_1} i_1$ -subsets of each i -set. The number different possible subgraphs formed by SI -arcs and SS -arcs is $\binom{n-k}{i} \binom{i}{i_1} k^{i_1} (n-k-1)^{n-k-i_1}$.

Let K_1 be the set of nodes that are the tails of the IS -arcs whose heads are in $S \setminus S_1$, where $|K_1| = k_1$. There are $\binom{k}{k_1}$ different ways to choose K_1 . These k_1 arcs will infect $i - i_1$ new nodes and there are $S(k_1, i - i_1)(i - i_1)!$ different ways to do this. Finally the remaining $k - k_1$ arcs can be chosen in $(k - 1 + i_1)^{k - k_1}$ different ways. Therefore, the number of different possible subgraphs formed by IS and II -arcs can be calculated as

$$\Theta_{k,i}(k_1, i_1) = \sum_{k_1=i-i_1}^k \binom{k}{k_1} (k - 1 + i_1)^{k - k_1} S(k_1, i - i_1)(i - i_1)!$$

Hence, the probability of infecting i more nodes in the next step given $|I| = k$ is

$$p(i|k) = \frac{\binom{n-k}{i}}{(n-1)^n} \sum_{i_1=0}^i \binom{i}{i_1} k^{i_1} (n - k - 1)^{n - k - i_1} \Theta_{k,i}(k_1, i_1)$$

where $k = 2, 3, \dots, n - 2$ and $i = 0, 1, \dots, n - k$.

Now, we will consider the end points. If $k = n - 1$, $p(0|k) = 0$ and $p(1|k) = 1$. If $k = 1$, $p(0|1) = 0$ and $p(i|1) = \frac{\binom{n-1}{i} i (n - 2)^{n - i - 1} (n - 1)}{(n - 1)^n}$ for all $i \geq 1$. There can be $\binom{n-1}{i}$ different i -subsets of S and i different possibilities for the head of the SI -arc, call this node u . There are $n - 1$ possibilities for the arc with the tail u . The arcs coming out of the rest of the $i - 1$ nodes will have heads in I and there is a unique way to do this. Finally the remaining $n - i - 1$ arcs can be chosen in $(n - 2)^{n - i - 1}$ different ways.

4 Expected Delay Per Peer

Many stochastic models of epidemic processes are based on the fact that the number of infectious peers, equivalently the number of susceptibles as the population size n is fixed, forms a Markov chain [1]. In existing models, the transition probabilities are modeled according to a probability distribution or left as rates to be estimated from the network due to the complexity of the problem. What is accomplished in Sect. 3 is that we have analyzed the true dynamics taking place at each transition of the Markov chain with no assumptions on any parameters. Using the analytical expressions derived for the transition probabilities, we find the message delay in this section.

The Markov chain under consideration is $\{I_t : t = 0, 1, 2, \dots\}$ where I_t denotes the number of infectious processes at time t . The transition probabilities $P_{kj} = P\{I_{t+1} = j | I_t = k\}$ can be obtained from $p(i|k)$ given in the previous section by

$$P_{kj} = p(j - k|k) \quad j = k, k + 1, \dots, n$$

where $j - k$ is the number of newly infected peers. Clearly, $P_{kj} = 0$ if $j < k$. In [11], the authors show that the delay experienced by each peer can be found by considering the first passage time of I_t to a specific set of states. Let $s_{i\bar{j}}$ denote

the first passage time from state i to the set of states $\bar{j} = \{j, j + 1, \dots, n\}$, for $i = 1, 2, \dots, j - 1$. The expected time for the j^{th} infection to occur, or the mean delay that the j^{th} member to receive the message experiences, is given by $s_{k\bar{j}}$ when the initial number of infectious peers I_0 is k . In order to find $s_{k\bar{j}}$ even for a single k value, one needs to solve the set of equations

$$(I - P_{\bar{j}})S_{\bar{j}} = \underline{1} \tag{1}$$

where $P_{\bar{j}}$ is the upper left $j - 1 \times j - 1$ portion of matrix P , I is the $j - 1 \times j - 1$ identity matrix, $S_{\bar{j}} = [s_{1\bar{j}}, s_{2\bar{j}}, \dots, s_{j-1,\bar{j}}]^T$ and $\underline{1}$ is a vector of 1's [11]. Since P is upper triangular, solving system (1) does not pose any numerical difficulties.

The delay experienced by each peer is an important performance measure from user perspective. In Fig.3, the peers in the order they receive the message are plotted against expected number of rounds for different starting number k of infectious peers, for $k=1,25,75$ and $n=100$. That is, $s_{k\bar{j}}$ appears in the x -coordinate for $j = k, k + 1, \dots, n$ and $I_0 = k$. Hybrid approach performs significantly better than the others. Although push approach is only slightly better than pull case in terms of mean delay when $k = 1$, its total time to disseminate to the whole population is much lower than the other. The delay is clearly lower for push case when $k > 1$. On the other hand, some peers have lower expected delay in the course of the information diffusion process when $k = 1$ such as the 10th to 15th peers in the order of receiving the message.

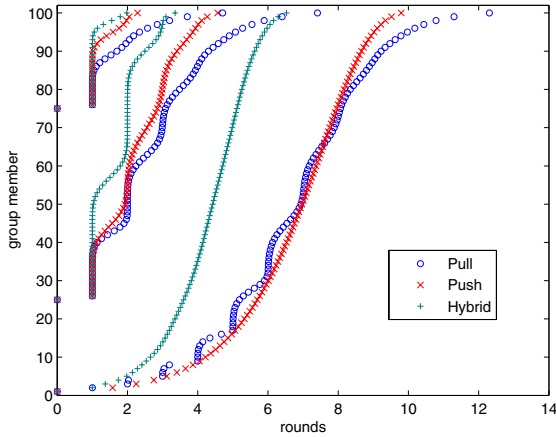


Fig. 3. Peers ordered according to their expected delays given in rounds, for 1, 25 and 75 starting number of infectious peers and $n=100$

Mean delay experienced per susceptible peer is depicted in Fig.4. In terms of this performance measure, the pull and push approaches behave similarly for small k , and the hybrid approach behaves like push case as k increases to n .

In [11], binomial distribution is used for modeling one step of epidemics for pull and push approaches. This happens to be the exact distribution for only

push case as indicated in section 3.2. On the other hand, the model of pull case estimates the true probabilities pessimistically in [11]. The pull approach does not behave as poorly as predicted by the binomial model there.

The advantage of epidemic dissemination, in particular the anti entropy paradigm, is its scalability with respect to population size. That is why the foregoing analysis is repeated for $n=200$. The results being similar to $n = 100$ are not given here. In order to demonstrate scalability, we have tabulated the expected time of dissemination to the whole population and the mean delay per peer in Tab. 1. These values are consistent with the prediction of asymptotical results for epidemic processes that the delay values increase only logarithmically as n increases.

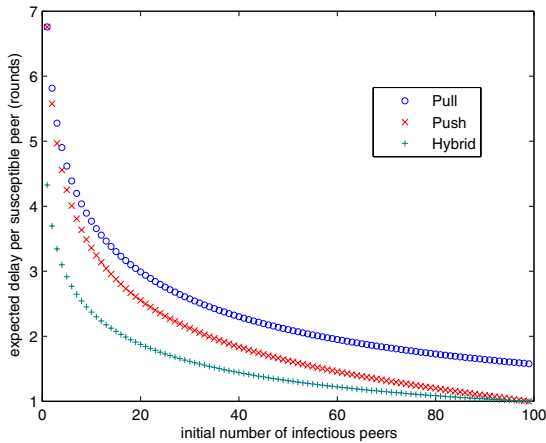


Fig. 4. Mean delay per susceptible versus initial number of infectious peers for $n=100$

Table 1. Expected time to dissemination and mean delay when $I_0 = 1$

	Time to dissemination		Mean delay per peer	
	$n=100$	$n=200$	$n=100$	$n=200$
Pull	12.30	14.05	6.76	7.75
Push	9.79	11.03	6.75	7.75
Hybrid	6.53	7.40	4.33	4.96

5 Conclusions and Future Work

We have derived the exact probability distributions for the pull, push and hybrid information diffusion models of anti-entropy. To the best of our knowledge, this study is the first one deriving exact distributions which would be helpful in performance analysis of these epidemic diffusion models. In contrast, previous results rely on simplified models of epidemics usually requiring estimation of

several parameters. Our findings show that the binomial model used previously for pull case is not accurate whereas the model for push case is exact. There exists no previous probability model for the hybrid case, the exact distribution of which is derived in this paper.

We have computed the expected delay of each peer as well as per arbitrary peer exactly, depending on the initial number of infectious members in the population. The hybrid approach outperforms pull and push paradigms, and push is better than pull case. Such results would be beneficial when integrating the models in several distributed scenarios such as replicated servers, loss recovery, failure detection and group membership management. Since the probability distributions found in this paper are exact, any possible discrepancies with real measurements of networks or their simulations can lead us to scrutinize other aspects such as overhead associated with each approach. For instance, the duplicate messages associated with the hybrid case due to both pull and push deliveries are worth counting in order to determine any trade off with the superiority of this case.

Dissemination of only one message has been considered. Initialization with a bigger volume of content such as in file sharing applications can be analyzed, the possibility of gossiping to more than one peer and partial knowledge of membership could also be incorporated as future work. Finally, evaluating the probabilities $p(i|k)$ are not computationally intensive, but accuracy is a concern for large group sizes although can be handled with the state of the art computing abilities. On the other hand, asymptotic expressions for large n would be useful.

References

1. N. T. J. Bailey. *The Mathematical Theory of Infectious Diseases and its Applications*. Charles Griffin and Compan, London, 1975.
2. 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 *Proc. of the Sixth ACM Symp. on Principles of Distributed Computing*, 1–12, 1987.
3. A. Birrell, R. Levin, R. Needham, M. Schroeder. Grapevine: Providing availability using lazy replication. *ACM Transactions on Computer Systems*, 10(4):360–391, 1992.
4. R. Golding and K. Taylor. *Group Membership in the Epidemic Style*. Technical Report, UCSC-CRL-92-13, University of California at Santa Cruz, 1992.
5. R. Ladin, B. Lishov, L. Shriram, and S. Ghemawat. An exercise in distributed computing. *Communications of the ACM*, 25(4):260–274, 1982.
6. K. Guo. *Scalable Message Stability Detection Protocols*. Ph.D. dissertation, Cornell University, Dept. of Computer Science, 1998.
7. R. van Renesse, Y. Minsky, and M. Hayden. A gossip-style failure detection service. In *Proceedings of Middleware98*, 55–70, 1998.
8. K. Birman, M. Hayden, O. Ozkasap, Z. Xiao, M. Budiu, and Y. Minsky. Bimodal multicast. *ACM Transactions on Computer Systems*, 17(2):41–88, 1999.
9. R. van Renesse, K. Birman, W. Vogels. Astrolabe: A robust and scalable technology for distributed system monitoring, management, and data mining. *ACM Transactions on Computer Systems*, 21(2):164–206, 2003.

10. P. Eugster, R. Guerraoui, A.-M. Kermarrec, and L. Massoulié. Epidemic information dissemination in distributed systems. *IEEE Computer*, 60–67, May 2004.
11. M. Çağlar, Ö. Özkasap. A chain-binomial model for pull and push-based information diffusion. *Proc. of IEEE ICC*, 2006.
12. I. B. Gertsbakh. Epidemic Process on a Random Graph: Some Preliminary Results. *J. Appl. Prob.*, 14:427–438, 1977.
13. Jerzy Jaworski. Epidemic Processes on Digraphs of Random Mappings. *J. Appl. Prob.*, 36:780–798, 1999.
14. Boris Pittel. On Spreading a Rumor. *SIAM Journal on Applied Mathematics*, 47(1):213–223, 1987.
15. J. H. van Lint and R. M. Wilson. *A Course in Combinatorics*. Cambridge University Press, 1992.