2009 International Symposium on Nonlinear Theory and its Applications
NOLTA'09, Sapporo, Japan, October 18-21, 2009

Epidemic Communication Using an Erasure Code

Masaki Terasawa,† Tatsuhiro Tsuchiya and Tohru Kikuno

Graduate School of Information Science and Technology, Osaka University
1-5 Yamadaoka, Suita, Osaka, 565-0871 Japan
Email: t-tutiya@ist.osaka-u.ac.jp, kikuno@ist.osaka-u.ac.jp

Abstract—This paper proposes to incorporate an erasure code into epidemic-style broadcast in a wired network, aimed at improved performance. The idea is to encode a broadcasting message with an erasure code and then to broadcast each of the encoded blocks using an epidemic protocol. Analytical and numerical results of the stochastic properties of the proposed approach are presented.

1. Introduction

Epidemic-style broadcast (a.k.a *gossip*) has recently gained popularity as a potentially effective solution for disseminating information in large-scale applications in distributed settings [2, 5, 9]. Aimed at improving performance, this paper proposes a new epidemic-style broadcast technique which uses an erasure code.

The essential characteristic of epidemic-style broadcast is that information exchanges occur between sender nodes and randomly chosen receiver nodes. Information is disseminated throughout the system by multiple rounds of such communication. A node is said to be *infected* if it has already received the message.

This proactive use of redundant messages provides a means to ensure reliability in the face of failures. Also, the epidemic protocols are scalable: it is shown that the load on each node increases only logarithmically with the size of the network [6, 8]. Due to these properties, the epidemic protocols have been used in various contexts, such as replicated databases [1, 4], distributed failure detection [11], or distributed information management [10].

Since epidemic-style broadcast does not ensure that a message reaches all nodes, its performance is usually evaluated with respect the probability that a node receives a broadcast message. The receiving probability and communication efficiency are in a trade-off relationship: if each node selects more nodes to infect, then the receiving probability increases but the amount of incurred traffic also increases.

The proposed protocol aims to alter or, if possible, improve this trade-off relationship. The outline of the new protocol is as follows: When initiating a broadcast message, the original message is encoded, using an erasure code, into n blocks such that the original message can be decoded from any k out of the n blocks. Then each of the n blocks is disseminated using ordinary epidemic-style broadcast. Because of these distinguishing traits, the trade-off exhibits different characteristics from the original epidemic protocol. Later we will show when the new protocol exhibits better performance.

† Currently affiliated with Panasonic Corporation.

```
initiate broadcast of m:
  send m to f randomly chosen nodes;

when a node p receives a message m:
  if (p has received m for the first time)
    p sends m to f uniformly randomly chosen nodes;
```

Figure 1: Ordinary epidemic protocol

2. Epidemic-Style Broadcast

2.1. Basic Protocol

Figure 1 shows the protocol of the basic epidemic-style broadcast. In this protocol when a node initiates broadcast, the node sends the message to f randomly selected nodes. Message dissemination is carried out as follows: upon receiving a broadcast message for the first time, the node p randomly selects f nodes and forwards copies of the message to all these selected nodes. The value f is usually referred to as *fanout*.

We let $\pi(f)$ denote the probability that a node receives a broadcast message when the ordinary epidemic protocol is used with the fanout set to f. If receiver nodes are uniformly randomly chosen from all the nodes and if the network size is large, then we have the following fixed-point equation [6]:

$$\pi(f) = 1 - \exp^{-\pi(f)*f}$$

Now let us discuss the total communication cost c incurred by broadcasting one message. Assuming that the cost incurred by unicasting a single message is 1, we define c as the total number of messages used for broadcasting a message. Since every infected node sends f messages, we have:

$$c(f) = \pi(f) * f$$

```
initiate broadcast of m:
   encode m into n blocks M_1, ..., M_n using an erasure code;
   for each block M_i;
      send M_i to f randomly chosen nodes;

when a node p receives a message M_i:
   if (p has received M_i for the first time)
      Add M_i to Received;
      p sends M_i to f uniformly randomly chosen nodes;
      if (|Received| = k)
         decode m from Received;
```

Figure 2: Epidemic protocol using an erasure code

3. Proposed Approach

Erasure codes are usually used in some forms of forward error correction. Well-known erasure codes include Reed-Solomon and Tornado codes. An erasure code transforms a message into a longer message with n blocks, such that the original message can be recovered from any k blocks of these n blocks. We assume that we have an optimal erasure code such that the size of each encoded block is $1/k$ of the original message.

Figure 2 shows the protocol of the proposed epidemic-style broadcast. The outline is as follows: The originator of a broadcast message uses an erasure code to encode the message into n small messages. Then the node disseminates each of the n messages using the ordinary epidemic protocol. The receiver node can decode the original broadcast message if it receives any k different messages of the n messages.

The probability that a node receives a broadcast message is given by the following equation.

$$\pi_{k,n}(f) = \sum_{i=k}^{n} \binom{n}{i} \pi(f)^i (1 - \pi(f))^{n-i}$$

Let $c_{k,n}(f)$ denote the communication cost incurred by broadcasting one message by the proposed protocol with fanout f. Note that the size of each message is now $1/k$ of the original message. Thus we assume that the message cost caused by unicasting an encoded block is $1/k$. (Recall that we assume that the cost caused by forwarding an original message is 1.) Consequently we have:

$$\begin{aligned} c_{k,n}(f) &= n * c(f) * \frac{1}{k} \\ &= n * \pi(f) * f * \frac{1}{k} \end{aligned}$$

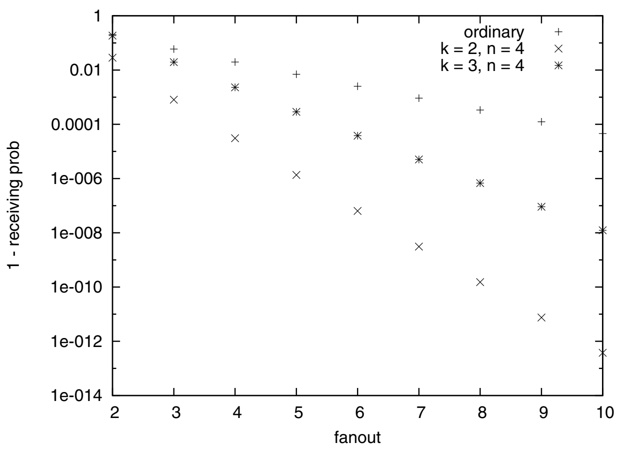

Figure 3: The probability of not being infected ($1-\pi(f)$, $1-\pi_{k,n}(f)$) as a function of fanout f

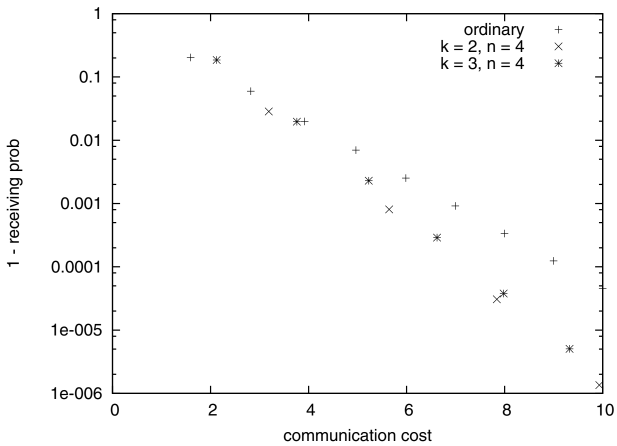

Figure 4: The probability of not being infected ($1-\pi(f)$, $1-\pi_{k,n}(f)$) vs communication cost ($c(f)$, $c_{k,n}(f)$)

4. Numerical Results

Figure 3 shows how the probability of not receiving a broadcast message decreases as the fanout increases for the two different protocols. The probability rapidly reaches 0 as the fanout increases.

We omit the graphs that depict communication cost as a function of the fanout, because the cost increases almost exactly proportionally to the fanout. This can be explained by the fact that all parameters except f and $\pi(f)$ are constants in the formulas of c and $c_{k,n}$ and that $\pi(f)$ is very close to 1 as shown in Figure 3.

Figure 4 depicts how the two protocols trade off the performance and the communication cost. The horizontal axis represents the communication cost, while the vertical axis represents the probability of failing to receive a broadcast message. A point in the graph represents the values for both measures when the fanout is fixed.

As shown in Figure 4, these two protocols exhibit different performance-cost relationships. When communication cost is small, the ordinary epidemic protocol exhibits slightly better performance than the new protocol. On the other hand, if the communication cost exceeds around 5.0, the erasure code-based protocol shows better performance. The two cases $(k, n) = (2, 4)$ and $(k, n) = (3, 4)$ exhibit almost the same performance-cost relationship.

5. Optimization Using Network Coding

This section presents an optimization technique for the proposed erasure code-based protocol. The optimization uses *network coding* [7]. Network coding refers to a scheme where coding is done at the interior nodes in the network, not only at the sender and receivers.

The idea of the optimization is as follows: When a node forwards a broadcast message, if that message is the first message that the node received, then the node simply forwards it, just as in the protocol of Figure 2. Otherwise, that is, when the node has had more than one message, it forwards, instead of one of the n blocks, a linear combination of the blocks over some finite field.

For example, suppose that we use $k = 2$ and that node p and q already have received blocks M_1 and M_2, respectively. Now suppose that another node r had received M_1 and just has received M_2. Then r forwards a linear combination of M_1 and M_2, instead of M_2. If p and q receive r's message, both can decode the original message using the new message. Note that this is not the case if r sent M_2, in which case only p could decode the original message.

The application of network coding to epidemic information dissemination has already been studied in, for example, [3], although in very different contexts.

6. Conclusion

This paper discussed the applicability of an erasure code to epidemic-style broadcast protocols. Numerical results show that the proposed technique can alter the performance-cost trade-off. These results also suggest that the erasure code-based epidemics can be more effective than the ordinary approach when permissible communication cost is relatively high. The difference is, however, somewhat subtle and thus it seems too rushed to conclude that the proposed approach is superior. For further improvement we introduced an optimization which uses network coding. The analysis and evaluation of this optimization are left for future work.

Acknowledgments

This work was supported in part by the MEXT Global COE program (Center of Excellence for Founding Ambient Information Society Infrastructure).

References

[1] D. Agrawal, A. El Abbadi, and R. Steinke. Epidemic algorithms in replicated databases. In *Proceedings of the Sixteenth ACM Symposium on Principles of Database Systems*, pages 161–172, 1997.

[2] K. Birman, M. Hayden, O. Ozkasap, Z. Xiao, M. Budiu, and Y. Minsky. Bimodal multicast. *ACM Transactions on Computer Systems*, 17(2):41–88, May 1999.

[3] S. Deb, M. Médard, and C. Choute. Algebraic gossip: A network coding approach to optimal multiple rumor mongering. *IEEE Transactions on Information Theory*, 52(6):2486–2507, June 2006.

[4] 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 *Proceedings of the Sixth Ann. ACM Symp. Principles of Distributed Computing (PODC)*, pages 1–12, Aug. 1987.

[5] P. T. Eugster, R. Guerraoui, S. Handurukande, A.-M. Kermarrec, and P. Kouznetsov. Lightweight probabilistic broadcast. In *Proceedings of the 2001 International Conference on Dependable Systems and Networks (DSN '01)*, pages 443–452, July 2001.

[6] P. T. Eugster, R. Guerraoui, A.-M. Kermarrec, and L. Massoulié. Epidemic information dissemination in distributed systems. *IEEE Computer*, 37(5):60–67, May 2004.

[7] S.-Y. R. Li, R. W. Yeung, and N. Cai. Linear network coding. *IEEE Transactions on Information Theory*, 49(2):371–381, Feb. 2003.

[8] B. Pittel. On spreading a rumor. *SIAM Journal on Applied Mathematics*, 47(1):213–223, Feb. 1987.

[9] Q. Sun and D. Sturman. A gossip-based reliable multicast for large-scale high-throughput applications. In *Proceedings of the International Conference on Dependable Systems and Networks (DSN 2000)*, pages 347–358, June 2000.

[10] R. van Renesse, K. P. Birman, and W. Vogels. Astrolabe: A robust and scalable technology for distributed system monitoring, management, and data mining. *ACM Transactions on Computer Systems (TOCS)*, 21(2):164–206, 2003.

[11] R. van Renesse, Y. Minsky, and M. Hayden. A gossip-style failure detection service. In *Proceedings of the IFIP International Conference on Distributed Systems Platforms and Open Distributed Processing (Middleware '98)*, pages 55–70, Sept. 1998.