Dynamic Avoidance of Illegal Nodes in Gnutella-like Pure P2P Networks

Kazuho Sato, Noriko Matsumoto, Norihiko Yoshida

Department of Information and Computer Sciences
Saitama University, Saitama 338-8570, Japan
{kazuho,noriko,yoshida}@ss.ics.saitama-u.ac.jp

Abstract. Pure P2P networks like Gnutella do not have a single point of failure. However, it is possible for an illegal (faulty or malicious) node to alter messages it transmits, so as to make the whole network malfunction. This altered message is forwarded by several nodes, thus it is difficult to identify and to eliminate the illegal node. This paper presents a method to avoid the influence of such illegal nodes by adjusting TTL of packets according to the evaluation of neighboring nodes. We show its effects by several simulation experiments.

1 Introduction

Pure P2P networks, such as Gnutella, are considered fault-tolerant because of their decentralized nature. It is not true actually; even a single illegal (faulty or malicious) node which alters messages while forwarding them can make the whole network malfunction. It is almost impossible to identify and eliminate such an illegal node, because there is no node which is always guaranteed to be correct.

In structured P2P systems, secure message forwarding, in which the message is forwarded correctly even in the presence of malicious nodes, is proposed [1]. However, this method cannot be applied to unstructured pure P2P networks.

On the other hand, there is a distributed fault diagnosis algorithm, which identifies illegal nodes by exchanging information among participant nodes [2]. However, it is difficult to apply this algorithm to P2P networks. Nodes must exchange messages for diagnosis, however there is no guarantee that these diagnostic messages are forwarded correctly.

This paper proposes a very simple method to avoid the anomaly influence of such illegal nodes without identifying them but make packets detour around them. Section 2 summarizes failures in Gnutella-like P2P networks. Section 3 proposes a method for dynamic avoidance of illegal nodes, and Section 4 presents some experiment results to show its effectiveness. Section 5 and 6 contain some discussions and concluding remarks.

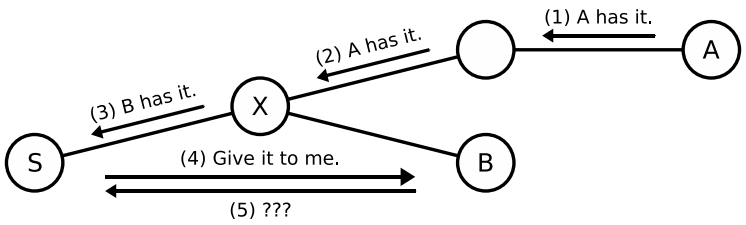

Fig. 1. Illegal Alteration of "QueryHit" Packet

2 Failures in Pure P2P Networks

In pure P2P networks, there can be some kind of failures. If a node stops working or a link is disconnected, a content may be lost, or the network may be separated into two disjoint subnets. This would reduce availabilities of contents, however the functionality of the network is not affected.

There may be a case that a node alters a query packet while forwarding. Suppose, as shown in Figure 1, that the node A replies a "QueryHit" packet of "A has the content" towards the query origin node S, and the node X in between alters the packet to "B has the content". Then, S cannot obtain the content. The functionality of the whole network may be affected by the single illegal (faulty or malicious) node. This is a fatal problem in pure P2P networks.

To make matters worse, this anomaly influence can only be detected by the query origin node S, and no node, even the illegal node itself neither, can identify who is illegal.

3 Dynamic Avoidance of Illegal Nodes

In the case of packet alteration by illegal nodes, the network cannot identify the illegal nodes, therefore it cannot eliminate the nodes, but can only detour around them [3]. Below is a very simple method to achieve this without introducing any additional special packets for diagnoses.

If a searching node cannot obtain the content according to a "QueryHit" packet, it means whether the content actually disappears during a short period, or someone behind the neighbor node which replies this "QueryHit" must be illegal. Consequently, the query origin node can avoid this anomaly thereafter, by making "Query" packets not to reach the illegal node any longer.

For this purpose, we take advantage of the TTL (time-to-live) parameter which limits the range of "Query" forwarding. When a query origin node cannot obtain the content according to a "QueryHit", the node reduces the TTL for the neighbor node when it multicasts "Query" packets thereafter. This applies both when the node originates queries, and when it forwards queries. This method reduces the possibility of "Query" packets to reach the illegal nodes.

The concrete operation of each node is as follows:

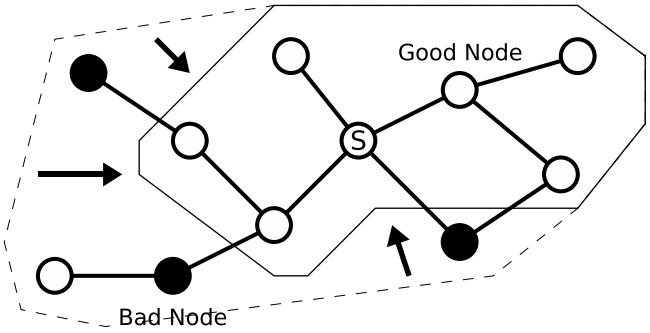

Fig. 2. Reachable Area Restriction

1. A node has a penalty counter for each neighbor node respectively. Its initial value is zero, and cannot be negative.
2. The node evaluates a "QueryHit" packet it receives. The node increases the penalty counter for the neighbor which replies an illegal "QueryHit" by one, while decreases the counter for the neighbor which replies a legal "QueryHit" by one, unless the counter gets negative.
3. When a node multicasts a "Query" to all the neighbors, the node decreases TTL of the "Query" packet according to each neighbor's penalty counter as:

$$\mathrm{TTL}_{neighbor} = \mathrm{TTL}_{default} - \mathrm{Penalty}_{neighbor}$$

When a node forwards a received "Query", the node decreases TTL as:

$$\mathrm{TTL}_{neighbor} = \mathrm{TTL}_{received} - 1 - \mathrm{Penalty}_{neighbor}$$

Following the above procedure, any packets become gradually unreachable to the illegal nodes, so that the network detours them to avoid their anomaly

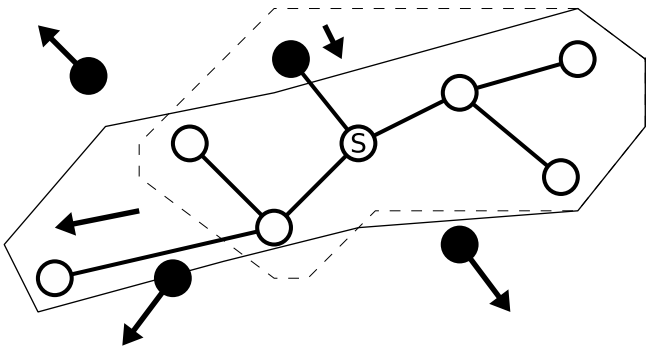

Fig. 3. Reachable Area Reorganization

influences. Figure 2 shows an example how the reachable area in the network becomes restricted to the legal nodes only. When some illegal nodes join or leave, the reachable area is reorganized eventually, as shown in Figure 3

4 Simulation Experiments

To verify the behavior of our proposed method, we conducted some preliminary experiments on top of our own simulator. Some principal parameters are: the number of nodes = 1,000, the number of illegal nodes = 50, the number of neighbors = 4, and the default TTL value = 7. Each node keeps one content, and searches randomly-selected contents. The simulator generates 1,000 searches as one step, and the illegal nodes alter "QueryHit" packets in a random manner.

Figure 4 through 8 show transitions of the numbers of (4) queries which illegal nodes received, (5) queries forwarded through illegal nodes, (6) errors of content retrieval due to altering, (7) successes of content retrieval, and (8) queries which correct nodes received. In each figure, we compare a run with our method and a run without it.

The result in Figure 4 shows that the queries which illegal nodes received are decreased by our method. This induces the decrease of queries forwarded through illegal nodes as shown in Figure 5. These results implies that our method suppresses alteration of "Query" and "QueryHit" packets. The result shown in Figure 6 shows that correct nodes avoid the influence of illegal nodes.

Figure 7 indicates that the number of successes in retrieval decreases once in the initial stage, but increases gradually thereafter. The decrease is caused by too much increase of the penalty counters because illegal nodes are not avoided sufficiently in the initial stage. After that, the illegal nodes are avoided sufficiently, so that the penalty counters in correct nodes get low, which bring increase of the success in retrieval.

The number of successes in retrieval with our method is lower than the number without our method. It is due to the shrinkage of area which queries can reach. The result in Figure 8 shows this.

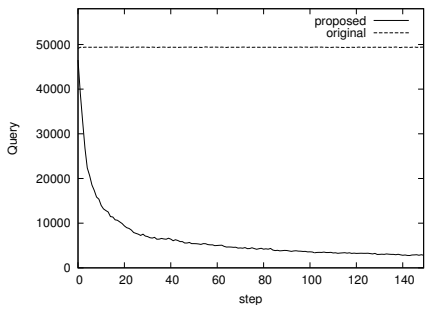

Fig. 4. Queries Illegal Nodes Received

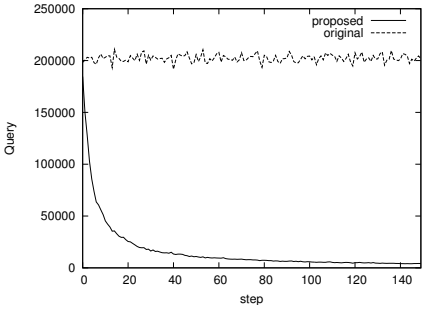

Fig. 5. Queries through illegal nodes

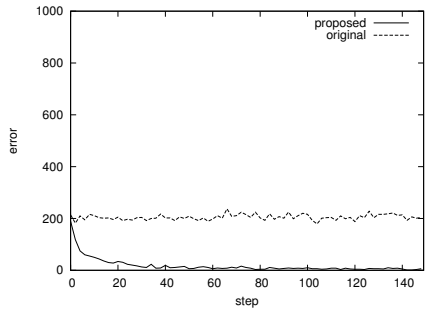

Fig. 6. Errors in Retrieval

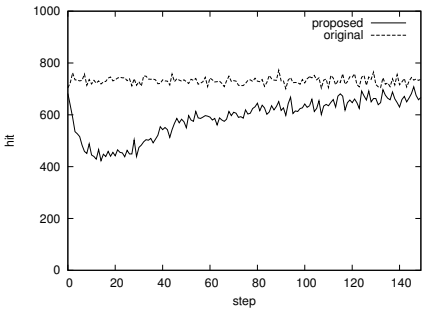

Fig. 7. Successes in Retrieval

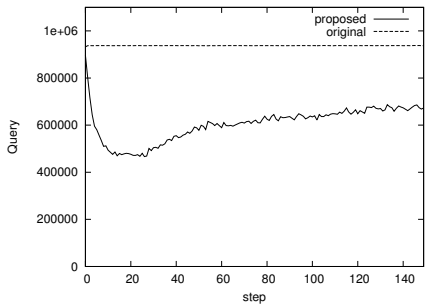

Fig. 8. Queries Correct Nodes Received

5 Discussions

Martinovic, et al. mentioned some similar idea on anomaly avoidance in outline in their paper on neighborhood evaluation [4], however they did not present any detailed consideration or evaluation for the idea.

An issue which still remains to be solved is the improvement of the number of search successes. The results of simulation show that our method shrinks the search area where the "Query" reaches. To improve our method, some other search method, e.g. Expanding Ring [5], is considered a possible alternative.

The method we propose is to adjust the TTL values of "Query" packets, and there would be a possibility that illegal nodes would alter TTLs in the packets. However, any correct node could detect packets with unusual TTLs.

Our proposed method cannot avoid nodes which alter messages intermittently. It is necessary counting the sum of the illegal "QueryHit" as well as counting the penalty in order to address this issue. Namely, a node accumulates respective illegal "QueryHit" which each node forwards, and decreases TTL according to this accumulated value.

6 Conclusions

P2P systems like a Gnutella have possibilities of alteration of message which makes system malfunction, namely, causes search failures. This paper proposed a very simple method to avoid illegal nodes that alter messages. This method uses penalty counters for neighbor nodes respectively, and decreases TTL of Query according to penalty so as to avoid reaching to illegal nodes. The simulation showed that this technique allows correct nodes to decrease the messages illegal nodes forwarded, that is, our technique is effective in avoiding illegal nodes. The simulation, however, showed that the number of Queries decreased too much, so that the numbers of search successes reduced. It is remaining work to address the decreasing the Queries.

Acknowledgments

This research was supported in part by JSPS in Japan under Grants-in-Aid for Scientific Research (B) 17300012, and by MEXT in Japan under Grants-in-Aid for Exploratory Research 17650011.

References

1. M. Castro, P. Druschel, A. Ganesh, A. Rowstron, and D. S. Wallach, "Secure Routing for Structured Peer-to-Peer Overlay Networks", Proc. Usenix Symp. on Operating Systems Design and Implementation, 2002
2. P. Preparata, G.Metze, and R.T.Chien, "On the Connection Assignment Problem of Diagnosable Systems", IEEE Trans. Electromag. and Comput., Vol.EC-16, No.6, pp.848–854, 1967
3. K. Sato and N. Yoshida, "Dynamic Avoidance of Illegal Nodes in P2P System" (in Japanese), IPSJ/IEICE Information Technology Letters, Vol.4, pp.307-309, 2005
4. I. Martinovic, C. Leng, F. A. Zdarsky, A. Mauthe, R. Steinmetz, and J. B. Schmitt, "Self-protection in P2P Networks: Choosing the Right Neighbourhood", Proc. 1st Int. Workshop on Self-Organizing Systems, 2006
5. Q. Lv, P. Cao, E. Cohen, K. Li, and S. Shenker, "Search and Replication in Unstructured Peer-to-Peer Networks", Proc. 16th ACM Int. Conf. on Supercomputing, 2002