Analysis of peer cluster layers selection criteria for P2P contents distribution systems

UEDA Kazunori
School of Information
Kochi University of Technology
Kami, Kochi, JAPAN 782–8502
Email: ueda.kazunori@kochi-tech.ac.jp

AKASE Jun-ichi† and OKUBO Takuya‡
Electronic and Photonic Systems
Engineering Course
Kochi University of Technology
Kami, Kochi, JAPAN 782–8502

***Abstract*—In many peer-to-peer content distribution systems, queries for content are transmitted regardless of physical network topology. This mechanism results in content transfer along long paths on the Internet. Redundant long paths unnecessarily consume network resources. To solve this problem, the authors had previously proposed a system in which clusters were constructed on multiple logical layers. In this study, the authors analyze criteria with respect to content popularity rates in order to switch the clusters on the basis of several simulation results.**

I. Introduction

In peer-to-peer (P2P) content distribution systems, each peer, i.e., each node, can search the content it wants to obtain among the peers in the entire network without using conventional servers. However, there is an increase in the number of messages because query messages for content search are transferred repeatedly. To solve this problem, overlay networks are constructed in order to search content effectively.

Certain P2P content distribution systems employ peer clustering methods and use the logical networks of clusters as overlay networks. Clustering is a method for peer coordination and peers are classified according to their characteristics. Peers have content trends, which can be considered as their characteristics. A P2P file sharing software, Winny, adopts a clustering method based on content trends and the line speed of peers.

However, the physical location of peers is not considered in many P2P content distribution systems. If other peers select a peer that is located in a distant network, the path length from the sender to the receiver is large. This leads to the consumption of network resources because of the connection set up for content transfer.

In this paper, we outline our proposed P2P content distribution system. In this system, peers can switch overlay networks to search content according to its popularity. One of the overlay networks is a network based on clusters constructed according to the content trends of each peer. The other is an overlay network based on clusters constructed according to the physical location of peers.

Moreover, we verify a criterion for switching layers of overlay networks according to content popularity. In fact, we perform simulations under several conditions and discuss the results.

†Presently with World Business Kochi Center
‡Presently with NISSHO ELECTRONICS CORPORATION

II. Related Work

In a P2P network model, a node does not function as a server, and all nodes exchange information for services. Therefore, content distribution systems based on the P2P network model have certain advantages such as fault tolerance and scalability. However, searching for peers that have necessary information incurs significant computational costs. To solve this problem, overlay network construction or peer clustering methods are used for grouping peers; this prevents unnecessary message transmissions.

In this section, we explain P2P content distribution systems (also called P2P file sharing systems), node coordination systems, physical network considerations in P2P network applications, and an overlay network switching method for P2P content distribution systems.

A. P2P content distribution system/P2P file sharing system

Thus far, various P2P file sharing systems have been used by numerous users.

Napster was one of the most well-known P2P file sharing systems and it adopted a hybrid P2P network model. The system required a server to manage the index information of media files. In contrast, gnutella adopted a pure P2P network model. In the network built by gnutella, each peer had to search media files using query message transmissions by employing the flooding technique. This method led to the consumption of network resources such as network bandwidth.

Winny is also a P2P file sharing system that adopts the pure P2P network model[1]. In addition, Winny performs clustering according to the content trends of peers and constructs overlay networks based on clusters. Moreover, peers are classified according to the line speed of peers. These schemes are adopted for the effectiveness of content search. However, clustering methods are used regardless of physical network topology or the physical location of peers. Figure 1 shows overlay networks and clusters of the Winny system.

In the Winny network, clustering is performed based on the similarity of peers to classify them on the basis of content trends. The similarity of peers is calculated using the comparison between content search keywords of the peers.

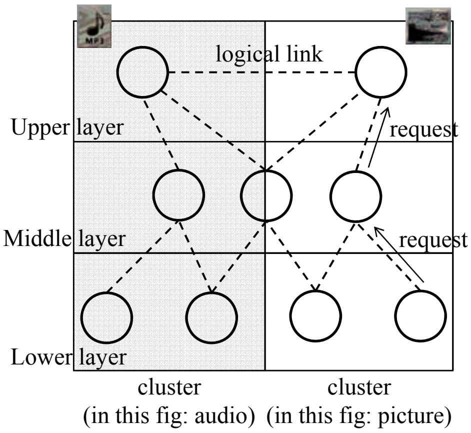

Fig. 1. Overlay networks and clusters of the Winny system

TABLE I. KEYWORDS COMPARISON

A \ B	abc	ab	bc
abcde	+3	+2	+2
ab	+2	+2	+0
bc	+2	+0	+2

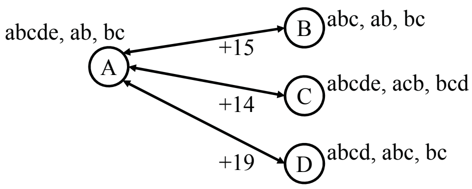

Fig. 2. Keywords similarity calculation in the Winny system

Three keywords are selected from the keywords sent by each peer for content search. Table I summarizes the similarity calculations among the keywords of two peers. In this example, the similarity between Peer A and Peer B is 15. This value can be calculated by adding all values listed in the table. Figure 2 shows similarities between Peer A and three other peers. In this example, the similarity between Peer A and Peer D is 19, and this value is the highest among the similarities between Peer A and other peers. A peer can join the most suitable cluster by repeatedly performing similarity calculations.

B. Node coordination system

Thus far, several node coordinating algorithms have been proposed.

PIC[2] is a distributed network coordinating system. PIC enables each node to construct a network coordinate space. Each node calculates the distance from a landmark and sets its coordinates to minimize the error calculated from the difference between the estimated distance and the measured distance from the landmark.

Vivaldi[3] is a node coordination method. Vivaldi enables each node to behave as an individual node to construct distributed network coordinates. The actual behavior of each node is similar to the nodes in PIC. Packets that are used for coordination are not transmitted independently; however, they are included in the normal data packets, i.e., piggy-back is adopted.

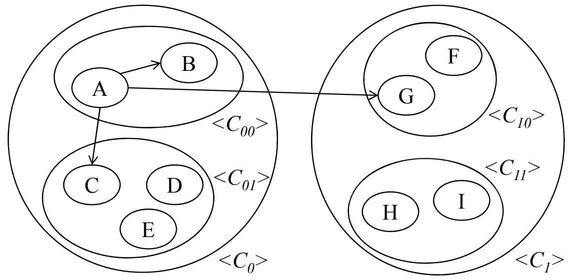

Fig. 3. Example of pointers to peers of other clusters

TABLE II. CLUSTER TABLE

Layer	joined cluster	another cluster 1	another cluster 2	...	another cluster n-1
1	$<C_0>$	$<C_1>$	$<C_2>$	...	$<C_n>$
2	$<C_{00}>$	$<C_{01}>$	$<C_{02}>$	...	$<C_{0n}>$
...	...	...	...	...	...
d	$<C_{d0}>$	$<C_{d1}>$	$<C_{d2}>$	...	$<C_{dn}>$

In a P2P based hierarchical clustering algorithm[4], clustering is executed hierarchically. In this method, peers in near networks are divided into the same cluster, i.e., peers in the same cluster are close to each other. Clusters are constructed hierarchically, so that a certain peer is closer to peers that belong to lower clusters than those, which belong to upper clusters.

Each peer has pointers to peers that belong to the same lowest cluster and pointers to a delegate peer of an upper cluster. Fig. 3 shows an example of pointers to each peer. In this figure, Peer A has pointers to peers that belong to other clusters. In the case of the upper layer (layer 1), Peer A belongs to cluster $<C_0>$ and consequently, Peer A has a pointer to Peer G that belongs to another cluster $<C_1>$. In the case of the lower layer (layer 2), Peer A belongs to cluster $<C_{00}>$ and the Peer A has pointers to Peer C that belongs to $<C_{01}>$, and Peer B that belongs to the same lowest cluster ($<C_{00}>$).

Each peer has a cluster table, which manages the pointers information. Table II is an example of a cluster table. In Table II, d is the number of layers and n is the number of sub clusters.

Each peer has cluster information, which depends on the cluster that the peer belongs to. Cluster information includes the following elements: cluster ID, delegate peer, faraway peer list, all peer list (if the cluster is the lowest cluster), and a backup peer list. Cluster ID is the identifier for each cluster. Delegate peer is a peer that responds to messages received from peers of other clusters. Faraway peers list is the list of peers that belong to clusters that are away from the cluster that the peer belongs to. Backup peers list is a list of candidate peers, which can become an alternate delegate peer if a current delegate peer breaks down.

When a new peer joins a network, it performs clustering. If the number of sub clusters is smaller than the threshold, the new peer creates a new cluster and if the number of sub clusters is more than the threshold, the new peer is allocated to an existing cluster in the near network. This routine is repeated until the new peer is allocated to a cluster on the lowest layer or a new cluster is created.

C. Provision of network information by ISP

Several working groups have proposed architectures for the provisioning of network information. These architectures enable obtaining network information such as cost, using hint servers, which are provided by internet service providers (ISPs).

Provide portal for application (P4P)[5] has been proposed by the P4P Working Group (P4PWG) of Distributed Computing Industry Association (DCIA). In a P4P architecture, ISPs provide a hint server called "iTracker," which offers information about the distance, policy, or the capability of networks.

The Application-Layer Traffic Optimization (ALTO)[6], [7], [8] protocol has been published as the RFCs of Internet Engineering Task Force (IETF). Each peer can obtain the network cost information from ALTO servers and use the information to select a peer.

These architectures enable to achieve effective content retrieval in terms of path length, by obtaining network information. However, a hint server such as an iTracker needs to be set by ISPs. All P2P content distribution systems do not have the ability to use the network information.

III. Content distribution system with multiple clusters

We have proposed a system for P2P content distribution[9], [10]. This system was proposed in order to achieve an effective content search by using overlay networks that have clusters based on individual clustering algorithms.

In the following parts, the system overview is described.

A. Basic concept of the proposed system

The prototype system has adopted two clustering algorithms as a method to construct an overlay network based on clusters. One of the clustering algorithms adopted is a conventional peer clustering, which is based on contents or keywords each peer has. The other is the peer coordination algorithm, which uses round trip time (RTT) between peers as the network distance. Figure 4 shows layers used by this system. Overlay networks constructed by different clustering algorithms and a physical network are shown in the figure. In the figure, peers that are close to each other in a physical network belong to the same cluster on the overlay network based on network distance. For example, Peer A and Peer B are members of the same cluster. On the other hand, peers with similar content trends or search requests belong to the same cluster on the overlay network. For example, Peer A and Peer C are the members of the same cluster because they have the same content.

The overlay network, which has clusters based on network distance, is placed on the physical network layer. Moreover, the overlay network has clusters based on the content trends of each peer and the peer is placed on the overlay network

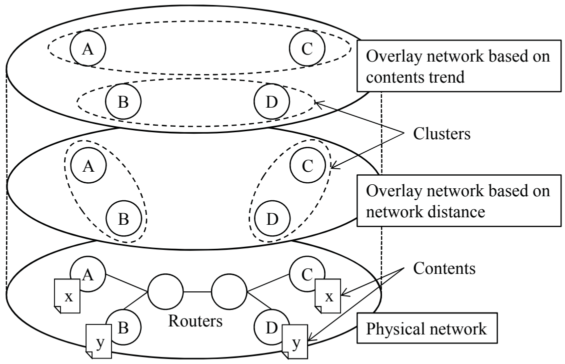

Fig. 4. Overlay networks of clusters layers

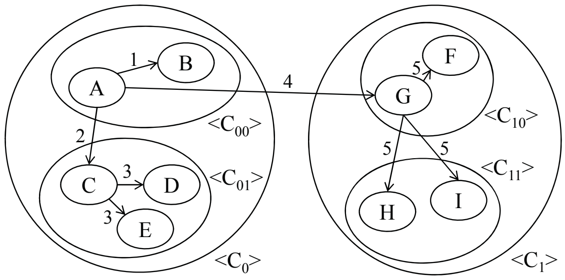

Fig. 5. Content search on the overlay network based on network distance

based on network distance. This system switches these overlay networks according to the popularity rate of searched content.

B. Content search on overlay network based on network distance

The proposed system switches between overlay networks, which are an overlay networks constructed using clusters based on content trends and an overlay network constructed using clusters based on network distance. Basically, the proposed system does not depend on specific clustering methods and can add or exchange clustering methods or algorithms. For future work, we will confirm which clustering method or algorithm is appropriate.

We assumed the P2P hierarchical clustering method[4] as a method of clustering based on network distance. The content search procedure followed on the overlay network based on network distance is as follows: 1) A peer sends a request for a content list to other peers that belong to the same lowest cluster. If the peers receive the request, they send the list to the peer. 2) If an index of contents that the peer has searched is included in the received content list, then the peer sends request to obtain the contents and finishes searching. 3) If the received content lists do not include contents the peer has searched, the peer sends the same request to delegate peers that belong to the upper clusters and sub clusters of upper clusters. If the received content lists do not include the index, delegate peers that belong to higher upper clusters are set as targets. These procedure is repeated until the content that the peer has searched is found or all content lists are completely searched.

Figure 5 shows the detailed content search procedure using the overlay network based on network distance. The content search procedure is as follows:

i Peer A starts to search.
ii Peer A sends a request for a content list to Peer B (arrow numbered 1 in Fig. 5). Peer B sends the content list to Peer A.
iii Peer A checks whether the received content list includes indexes of contents that Peer A has searched. If the received content list includes indexes of the contents, Peer A sends a request for the contents to Peer B.
iv If the received content list does not include indexes of the contents, Peer A sends a request for a content list to Peer C (arrow numbered 2 in Fig. 5). Peer C sends a content list to Peer A.
v Peer C sends the request to Peer D and Peer E (arrows numbered 3 in Fig. 5). Peer D and Peer E send content lists to Peer C. Peer C sends the content lists from Peer D and Peer E to Peer A. Peer A checks whether the received content lists include indexes of contents. If the received content lists include indexes of the contents, Peer A sends a request for the contents to Peer E or Peer D.
vi If the received content lists do not include indexes of the contents, Peer A sends a request for a content list to Peer G (arrow numbered 4 in Fig. 5). Peer G sends a content list to Peer A.
vii Peer G sends requests to Peer H and Peer I (arrowed numbered 5 in Fig. 5). Peer H and Peer I send content lists to Peer G. Peer G sends the content lists from Peer H and Peer I to Peer A, and Peer G sends a request to Peer F (arrowed numbered 5 in Fig. 5) in the same bottom cluster. Peer F sends a content list to Peer G. Peer G sends the content list from Peer F to Peer A.
viii Peer A checks whether the content lists include indexes of the contents.

Peers in our system search contents for peers in a near network. Consequently, peers can obtain contents from peers in a near network.

C. Content search on overlay network based on content trends

We assumed the content trends and hierarchical clustering method of Winny as a method of clustering based on content trends. The overlay networks of Winny network are shown in Fig. 1. In the Winny network, content search is performed as follows:

i Peers on the upper layer collect queries for contents
ii A peer on the lower layer sends a request message to peers on the middle layer
iii Peers on the middle layer transfers the request message to peers on the upper layer
iv Peers on the upper layer search contents by checking queries the peers have collected
v If a peer discovers searched content, the peer sends the result to the peer that sent the request message

The possibility that peers respond to a peer that sends a request message is high, because peers on the upper layer always collect queries for contents.

D. Switching overlay networks

In the proposed system, overlay networks are switched according to the content popularity. We assume that popular contents are requested many times. Content popularity is calculated by collecting search requests for content on the basis of this assumption.

TABLE III. SIMULATION PARAMETERS AND CONDITIONS

Parameter	Value
Number of peers	10,000
Categories of contents	10
Number of contents	200
Number of local networks	8

TABLE IV. CONTENT DISTRIBUTION

Order of popularity	Rate of peers that have the contents [%]
1st	100
2nd	50
3rd	33.3
. . .	. . .
k-th	1/k*100
. . .	. . .
10000th	0.01

We have considered a criterion for switching overlay networks depending on many conditions. Moreover, we have already confirmed that top 20 percent popularity, which yields 80 percent benefit (this law is known as Pareto's law), is not always suitable for the proposed system.

In next section, we consider the criterion for switching overlay networks with several simulation results.

IV. ANALYSIS OF PEER CLUSTER LAYER SELECTION CRITERION

A criterion about switching overlay networks is required in order to use the proposed system. However, it seems that the criterion depends on many conditions. Therefore, we performed simulations to verify how to decide the criterion.

A simulation system is implemented on the P2P agent platform PIAX[11], [12]. PIAX is an emulator which is created using JAVA and provides functions, which allows to emulate agent behaviors on peers easily.

We performed simulations with this emulation environment for verification. The search hit rate, search hit rate in a local network, and number of messages were measured from the simulation results.

A. Simulation conditions

Table III lists the parameters and conditions of the simulations. The location of peers was set based on real locations of Japanese cities. Moreover, the distribution of peers in networks was set, based on the population of each city.

We assumed that the distribution of contents depends on Zipf's law. Zipf's law states that 1/k of all peers have the k-th popular contents. Table 1 lists the actual content distribution obtained in the simulations.

The simulation procedures are as follows:

i All peers join the network.
ii Contents are distributed randomly.
iii Each peer receives information from a reference peer and constructs overlay networks, which are overlay networks based on content trends and network distance. In this

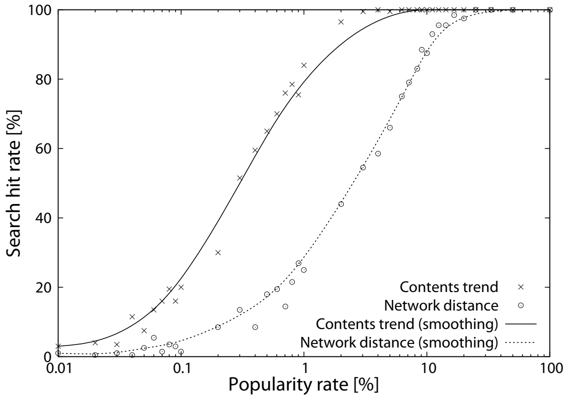

Fig. 6. Search hit rate

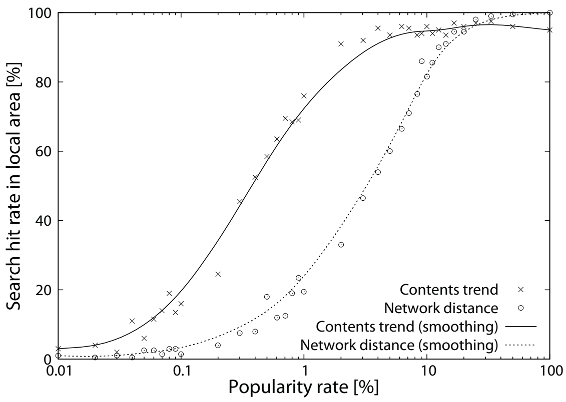

Fig. 7. Search hit rate in local network

overlay network construction process, each peer rejoins the network.

iv Peers on the lower layer in content trends clusters send the contents information that peers have, to the peers on the upper layer in content trends clusters.

v A peer selected randomly starts, to search contents.

vi After all search processes of peers are complete, statistical data is calculated.

B. Simulation Results and Discussion

Figure 6 shows the search hit rate. From this result, we confirmed that content search with overlay network based on content trends is better than the conventional content search method. The trend found in this result is obvious.

Figure 7 shows the search hit rate in a local network. From this result, we confirmed that content search can be finished in local networks if the popularity rate is higher than about 15 percent.

Figure 8 shows the number of messages consumed for content search. From this result, we confirmed that content search with overlay networks based on network distance requires many request messages in order to find the content. The trend found in this result is also obvious.

However, we have considered that the peers may use the overlay network based on network distance when peers search

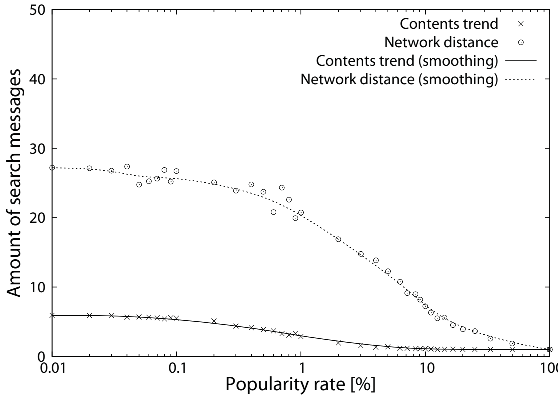

Fig. 8. The number of search message

content with a popularity rate higher than 15 percent. Then, it seems that the path length of data transmission is smaller. A smaller path length can reduce the packet transmission wastage on non-essential paths. We have considered this, as it has advantages; however, certain messages are consumed for content search.

V. Conclusion

This paper introduced an outline of peer-to-peer content distribution systems based on multiple logical layers constructed by peer clustering. Moreover, the criterion for switching layers to select a smaller path on the proposed peer-to-peer content distribution system were discussed. From simulation results, we have verified trends of each perspective that were estimated in advance. However, we have concluded that the criterion we had set was not always suitable for this purpose.

For future work, we plan to perform simulations under other scenarios and environments. Moreover, we will adopt other clustering methods and verify the characteristics of systems with logical layers based on those clustering methods.

References

[1] I. Kaneko, *The Technology of Winny (in Japanese)*. ASCII, 2005.

[2] M. Costa, M. Castro, R. Rowstron, and P. Key, "PIC: practical internet coordinates for distance estimation," in *Distributed Computing Systems, 2004. Proceedings. 24th International Conference on*, Mar 2004, pp. 178–187.

[3] F. Dabek, R. Cox, F. Kaashoek, and R. Morris, "Vivaldi: a decentralized network coordinate system," in *ACM SIGCOMM Computer Communication Review*, vol. 34, no. 4, Aug. 2004, pp. 15–26.

[4] T. Ueda, K. Abe, H. Ishibashi, and T. Matsuura, "Decentralized hierarchical internet hosts clustering (network system management technologies, management and operation of distributed systems and the internet)," *IPSJ Journal*, vol. 47, no. 4, pp. 1063–1076, Apr. 2006.

[5] H. Xie, Y. R. Yang, A. Krishnamurthy, Y. Liu, and A. Silberschatz, "P4P: Provider portal for applications," in *ACM SIGCOMM Computer Communication Review*, vol. 38, no. 4, Oct. 2008, pp. 351–362.

[6] Seedorf, J. and Burger, E., "Application-Layer Traffic Optimization (ALTO) Problem Statement," IETF RFC 5693, Oct. 2009.

[7] Kiesel, S., Previdi, S., Stiemerling, M., Woundy, R., and Yang, Y., "Application-Layer Traffic Optimization (ALTO) Requirements," IETF RFC 6708, Sep. 2012.

[8] Alimi, R., Penno, R., and Yang, Y., "ALTO Protocol," IETF Internet-draft draft-ietf-alto-protocol-16.txt, May 2013.

[9] K. Ueda and T. Okubo, "Peer–to–Peer contents distribution system using multiple clusters," in *Network Operations and Management Symposium (APNOMS), 2012 14th Asia-Pacific*, Sep. 2012, pp. 1–6.

[10] T. Okubo and K. Ueda, "Peer-to-Peer contents delivery system considering network distance," in *Network Operations and Management Symposium (APNOMS), 2011 13th Asia-Pacific*, Sep. 2011, pp. 1–4.

[11] PIAX, http://www.piax.org/.

[12] M. Yoshida, T. Okuda, Y. Teranishi, K. Harumoto, and S. Shimojo, "PIAX : A P2P Platform for Integration of Multi-overlay and Distributed Agent Mechanisms," *IPSJ Journal*, vol. 49, no. 1, pp. 402–413, Jan. 2008.