TI-CHORD: AN IMPROVED CHORD MODEL BASED ON TOPOLOGY AND INTEREST LAYER

Qin Dong
School of Computer Science and Technology
Beijing University of Posts and Telecommunications
Beijing, P.R. China
e-mail:dongqin.bupt@gmail.com

Jingling Zhao
School of Computer Science and Technology
Beijing University of Posts and Telecommunications
Beijing, P.R. China
e-mail: jing_lingzh@sina.com

***Abstract*—In original Chord model, the semantic property of the content in the model is not taken into account. Besides, a node's logical ID is independent of its physical location, bringing tremendous delay to network routing. Aiming at the system instability of structured P2P model which is caused by the heterogeneity of nodes in model, this paper proposes a Structure P2P network based on the Chord (TI-CHORD). TI-CHORD divides the entire network into different groups according to the interest. Groups are in the form of organization ring, and nodes in one group have the same interest. Each group generated a super node independent of physical topology and all the super nodes constitute the super group. Compared with the original Chord, the TI-CHORD enhances the query efficiency and decrease the impact of dynamic environment.**

Keywords- P2P; Chord; group; interest; physical topology

I. INTRODUCTION

A Peer-to-Peer (P2P) networked system is one in which several individuals participate in the construction of an independent network. Such a network might be centralized or decentralized, meaning such a network might be constructed with or without a central authority. Such a system performs application level routing on top of IP routing. Based on the organization, a P2P network can be classified as structured and unstructured. Structured P2P systems are those in which nodes organize themselves in an orderly fashion while unstructured P2P systems are those in which nodes organize themselves randomly. Structured P2P systems boast an efficient lookup mechanism by means of Distributed Hash Tables (DHTs) while most of the unstructured P2P systems use broadcast search technique.

Chord is a typical DHTs-based lookup protocol of P2P network designed by MIT in 2001. Chord has some outstanding characters, e.g., simplicity, provable correctness, provable performance. However, chord is still deficient. First of all, chord leaves out the consideration of the queries' semantic character that helps to improve the query efficiency. Secondly, the system is sensitive to the dynamics of the network, and the maintenance costs greatly increase when the nodes join and leave frequently. Thirdly, the construction of a P2P network in Chord does not take into account underlying network topology.

This paper proposes a topology-interest structured P2P system—TI-Chord, of which the nodes are configured by prefix. Therefore; routing process of The TI-Chord takes both content's interest and physical network into account. TI-CHORD divides the entire network into different groups according to the interest. Groups are in the form of organization ring, and nodes in one group have the same interest. Each group generated a super node independent of physical location and all the super nodes constitute the super group.

Through these schemes, P2P search is restricted to a group of nodes, in which nodes have the same interests, and this allows faster search and relieves other nodes from processing search requests.

The rest of the paper is organized as follows. Related work is discussed in SectionⅡ. We describe our TI-Chord model in detail in SectionⅢ. In SectionⅣ, we evaluate routing performance of TI-Chord. We give the conclusion in SectionⅤ.

II. RELATED WORK

Recently, hierarchical super-peer systems[2] have been proposed to improve search efficiency. They utilize the heterogeneity existing in P2P networks and adopt hierarchy in the form of Super-nodes (Fast Track [4]). These powerful nodes maintain the indices for other nodes, therefore searching can be carried out only among these more powerful ones. The introduction of a new level of hierarchy in the system increases the scale and speed of a query lookup.

Semantic Web [5] attempts to define the metadata information model for the World Wide Web to aid in information retrieval and aggregation. It provides general languages for describing any metadata. Currently, many P2P applications [6, 7] have leveraged semantic web technologies to add semantics to P2P systems, and thus improving the effectiveness of content and query representation, and the efficiency of content searching.

To resolve mismatching between overlay topology and physical network topology in P2P systems, there are many types of approaches proposed.

978-1-4244-7941-2/10/$26.00 ©2010 IEEE

PChord[8] achieves better network proximity than the original Chord by using proximity lists, i.e. the list of the close nodes that a node discovers in its lifetime. The next hop is decided by the entries in both the proximity list and finger table. Although this approach achieves better routing efficiency than Chord while keeping lightweight maintenance costs, it has problems of slow convergence and inefficiency in the case of churn, where the lifetime of a node in the overlay is relatively short. And also, when the number of nodes is large enough, the performance of PChord get worse dramatically.

III. TI-CHORD MODEL

This section gives a detailed explanation of TI-CHORD model. At first, we introduce group theory and how to construct the overlay network, and then, explain the lookup algorithm. At last, we demonstrate our approaches on how to maintain the system.

A. Group theory

Group theory is introduced, to reduce efficiency of the node joining and leaving the system. According to the interest, the network could be divided to many groups. Each group generate a node to make up the super group. Each node in the super group is the code whose physical position is nearest to the other node in the group (called super node). Super node in the super group is containing multiple original nodes.

B. Overlay construction

The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin in this template measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

1) Naming Mechanism

The consistent hash function assigns each node and keys an m-bit identifier using SHA-1[10] as a base hash function. Different from the original chord, the identifier consists of two parts. As shown in Table 1, each node in TI-Chord ring has a unique identifier named NodeID. It consists of G-ID and L-ID. G-ID is produced by hashing the node's topic, while L-ID represents node's physical Topology which determines its location in the super group of TI-Chord ring. A key's identifier called KeyID is gotten by the same way.

TABLE I. NODEID (M= M1 + M2)

G-ID=Hash(Topic)(m1 bit)	L-ID=Location (m2 bit)

TABLE II. KEYID (M= M1 + M2)

G-ID=Hash(Topic)(m1 bit)	K-ID=Hash(Key) (m2 bit)

2) Overlay Architecture

TI-CHORD divides the entire network into different groups according to the interest. Groups are in the form of organization ring, and nodes in one group have the same interest. Each group generated a super node independent of physical location and all the super nodes constitute the super group. The super group is composed of super nodes, and each super node is containing many nodes in one group(also called virtual node).In the super group one node's physical location is closer to the other super nodes. This topology shows that the intra-class request and query can finish in the group, the request in super group are forwarded by group of super nodes in the super group. In order to support the routing, common node store the information of the common node in the same group and the super node of the group, and the super node is not only need to maintain the information of the super node list items, still need to store information of other common nodes in the same group.

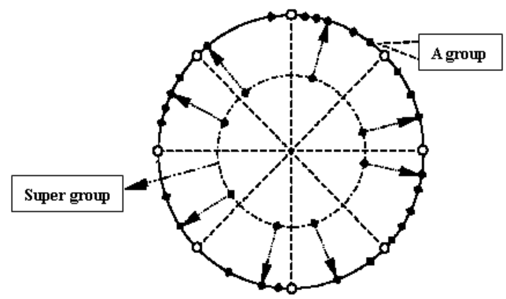

Figure 1. Overlay Architecture

C. *Lookup Mechanism*

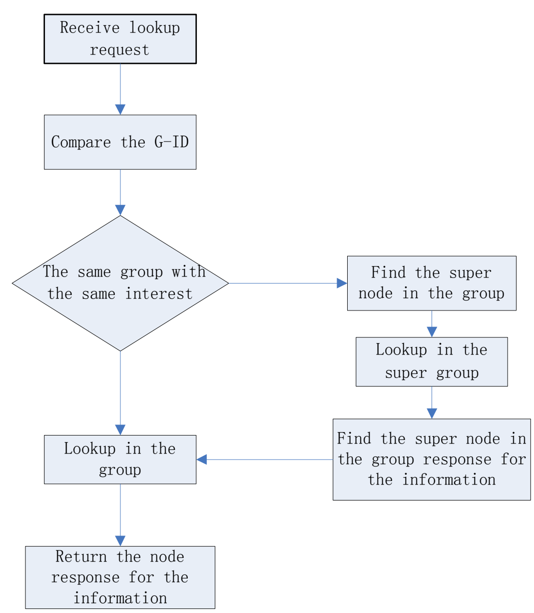

Figure 2. Lookup Procedure

When node n initiates a lookup request, it compares the G-ID of the key with its G-ID. If G-ID (n) = G-ID (key), that is, the key is stored in the node's group. And then, using K-ID (key) as the target, node n starts a lookup by the means of original Chord in the node's group. As a result, the node responsible for the key will be found.

If G-ID(n) ≠ G-ID(key), that is, the key is not stored in the node's group. In this condition, node n forwards the query request to the super node of the node's group. Then, using G-ID (key) as target, the super node starts a lookup in the super group to look for the super node responsible for the key according to the strategy in Chord. When the super node is found, the request will be forwarded to the super node of the super group. After that, using K-ID (key) as the target, the super node will look for the node responsible for the key in the group in which it is. Finally, the responsible node for the key is found.

D. Nodes join and departure

Because we adopt the mechanism in the improved Chord, the impact of nodes join and departure on the operation of the whole network will be reduced.

When node n is about to join the network, it firstly get its G-ID through NCF. If there is a group who's G-ID is equal to the G-ID (n), that is, there has been a group in which the nodes have the same interests as the joining node. Then, node n will join the group, and becomes a member of the group, using L-ID (n) as its ID. If there is no group having the same G-ID as node n, a new group whose G-ID is G-ID (n) will be established and node n is naturally become the super node of the new group, however, as nodes having better capability join the group, the super node can be chosen over again.

When node n that is not the super node is about to leave the network, only the node's group will be affected, and the maintain mechanism of original Chord will be implement in the group. If the leaving node is a super node, the backup node of the node's group will substitute it. At the same time, a new backup node will be selected from other nodes in the group. Consequently, as long as there are nodes in the super group, the impact of the node's departure will be confined within the group, which greatly decreases the overhead of maintaining the whole system.

In the improved Chord, the number of group members is not less than one with comparatively high probability, thus the major problem of nodes join and departure can be solved within the super group.

IV. SIMULATION

In P2P overlay network, average routing hop and average routing delay are indicator of routing performance. The average routing hop is the average forward times invoked by a routing process and the average routing delay is the average time delay over the path from source node to target node for routing process. In order to evaluate the performance of the TI-Chord, we perform large-scale experiments by simulation to the TI-Chord and traditional Chord, respectively. We constructed our simulator on Omnet++ which is a public-source, component-based, modular and open-architecture simulation environment with strong GUI support and an embeddable simulation kernel.

There are six different network topologies in our experiment. During this simulation, 60 query requests are invoked per node and destination node are generated at random.

A. Routing hop in TI-Chord

Fig3 shows routing hops comparison between TI-Chord and traditional Chord. We set the number of nodes 1000, 4000, 8000, 12000, 16000, and 20000. From the fig, It is obvious that the needed hops become less in TI-Chord.

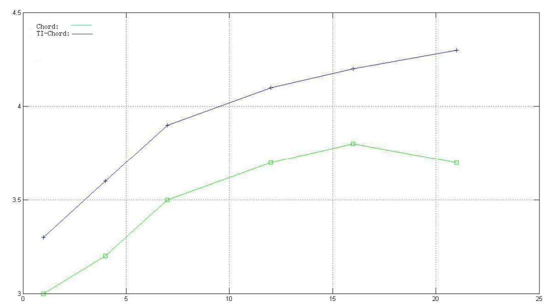

Figure 3. Routing hops in TI_Chord and Chord

B. Routing delay in TI-Chord

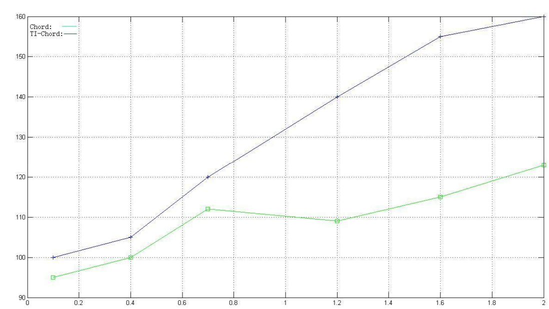

Figure 4. Routing delay in TI-Chord and Chord.

Fig4 shows routing delay comparison between TI-Chord and traditional Chord. The set is same as hop comparison. TI-Chord routing delay is shorter than traditional Chord which means our design's routing performance is better than the original. This is because TI-Chord take the semantic property of the content into account when building chord ring, so when node search resource it looks up locally first.

V. CONCLUSION AND FUTURE WORK

To deal with ignorance of the semantic property of the content and physical topology information in P2P, this paper presents an improved Chord model, named TI-Chord. Two core concepts are integrated into the system structure: super group and super node. The system facilitates the landmark +RTT method to generate topology information. The result of experiments show that the TI-Chord, compared with traditional Chord, has obvious improvements in the routing delay and the hops of overlay networks.

Our work provides an initial study of the TI-Chord system. Although simulation results are encouraging, there are many open problems to be addressed. We will address this problem in the future. We will also present the hot point problem in TI-Chord.

REFERENCES

[1] I. Stoica, R. Morris, D. Karger, M. Kaashoek, and H.Balakrishnan."Chord: A scalable peer-to-peer lookup service for internet "IEEE/ACM Transactions on Networking, Vol.11, No.1, February 2003

[2] B.Yang and H.Garcia-Molina, Designing a Super-Peer Network, Proc.19th Int'l Conf. Data Engineering, Los Alamitos, CA, March 2003

[3] W.P.K. Yiu, X. Jin,and S.H.G. Chan, "VMesh: Distributed SegmentStorage for PeertoPeerInteractive Video Streaming", pp1717, Journal onSelected Area of Communication, Dec 2007.

[4] B.Yang, and H.Garcia-Molina, Designing a Super-Peer Network, Proc. 19th Int'l Conf. Data Engineering, IEEE Computer Society Press, Los Alamitos, CA, March 2003

[5] T Berners-Lee, J Handler, O Lassila, The semantic web. Scientific American, 2001,284(5):34–43.

[6] M. Schlosser, M. Sintek, S. Decker and W.Nejdl. A Scalable and Ontology-Based P2P Infrastructure for Semantic Web Services. Second International Conference on Peer-to-Peer Computing (P2P'02) September 05-07,Linkoping, Sweden 2002.

[7] W. Nejdl, M. Wolpers, W. Siberski, A. Loser,I. Bruckhorst, M. Schlosser, and C. Schmitz.Super-Peer-Based Routing and Clustering Strategies for RDF-Based Peer-To-Peer Networks", In Proceedings of the Twelfth International World Wide Web Conference (WWW2003), Budapest, Hungary, May 2003.

[8] L. H. Dao, J. Kim. AChord: Topology-aware Chord in anycast-enabled networks. In Proceedings of the 2006 International Conference on Hybrid Information Technology (ICHIT 06), 2006.

[9] Zhao Jingling, Xiao Yonggang, Liao Qing. HTC-CHORD: AN IMPROVED CHORD MODEL BASED ON TOPIC-CLUSTER AND HIERARCHIC LAYER, Proceedings of IC-BNMT2009

[10] C. Wu, B. Li, and S. Zhao, 'Characterizing PeertoPeer StreamingFlows', pp1612, Journal on Selected Area of Communication, Dec 2007.