International Journal of Management, IT & Engineering
(ISSN: 2249-0558)

CONTENTS

Editorial Board

Mr. SANJAY ASATI
Dept of ME, M. Patel Institute of Engg. & Tech., GONDIA(M.S.)

Mr. G. Y. KUDALE
N.M.D. College of Management and Research, GONDIA(M.S.)

Editorial Advisory Board

November 2011

IJMIE

Volume 1, Issue 6

ISSN: 2249-0558

Title

COMPARING SEARCH ALGORITHMS OF UNSTRUCTURED P2P NETWORKS

Author(s)

Prashant K. Shukla
Asst. Prof.,
Deptt. of IT, SIRT,
Bhopal, M.P., India

Piyush K. Shukla
Asst. Prof.,
Deptt. of CSE, UIT,
RGPV, Bhopal, M.P., India

Prof. Sanjay Silakari
Deptt. of CSE, UIT,
RGPV, Bhopal, M.P., India

Abstract:

Computing has passed through many stages since the birth of the first computing machines. A centralized solution has one component that is shared by users all the time. All resources are accessible, but there is a single point of control as well as a single point of failure. A distributed system is a group of autonomous computers connected to a computer network, which appears to the clients of the system as a single computer.

Distributed system software allows computers to manage their activities and to share the resources of the system, so that clients recognize the system as a single, integrated computing facility. Opportunity to attach components improves the availability, reliability, fault tolerance and performance of the system. In such systems, the methods for minimizing communication and computation cost are significant. The widely used client–server model is an example of a distributed system. In this model, the servers are optimized to offer services to several clients. A peer-to-peer network is also a good example of a distributed system. In peer-to-peer network, we cannot force a peer to remain online until a file is completely downloaded; therefore it is important to choose a node, which is trustworthy by nature.

Keywords: Multi-tire differential search, unstructured P2P networks, OPNET

INTRODUCTION OF PEER TO PEER NETWORK:

A peer-to-peer, commonly abbreviated to P2P, is any distributed network architecture composed of participants that make a portion of their resources (such as processing power, disk storage or network bandwidth) directly available to other network participants, without the need for central coordination instances (such as servers or stable hosts). Peers are both suppliers and consumers of resources, in contrast to the traditional client-server model where only servers supply, and clients consume.

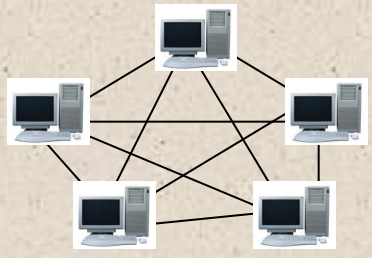

Fig.1 A peer-to-peer system of nodes without central infrastructure

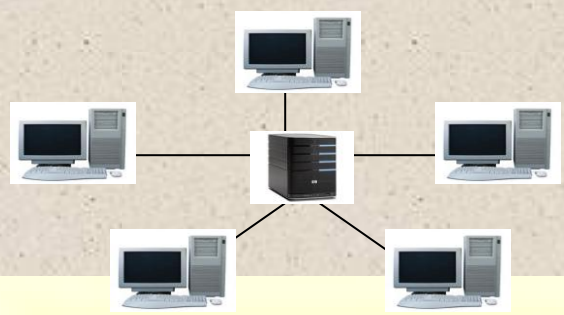

Fig. 2 Centralized server-based service model

PREVIOUSLY WORK HAS DONE ON UNSTRUCTURED P2P NETWORKS:

Qian Su Xuejie Zhang [1] propose an unstructured P2P resources search algorithm based on small-world model, in which peers perform k-means clustering on local resources separately, and then build a few similar-links between peers who own similar clusters and some random-links between non-similar peers. Fuyong Yuan Jian Liu Chunxia Yin [2] improves search efficiency and reduce unnecessary traffic in Gnutella, he proposes an algorithm to unstructured P2P network, and it consists of ranked neighbor caching scheme and query hit caching scheme. Yu Jin Yan Liu Hongwu Zhao [3] presents a new trust-based super node selection method that is complementary to the only-capability selection. He uses the concept of proxy trust to identify the behaviors of super nodes. Yongqiong Zhu Ruimin Hu [4] focused on building keyword search service over unstructured Peer-to-Peer (P2P) networks. Current state-of-the-art keyword search approaches for unstructured P2P systems are either blind or informed. Blind search methods such as flooding in Gnutella generate a large of redundant cloned messages and waste network bandwidth. Informed approaches such as routing indices can allow nodes to forward queries to neighbors that are more likely to have answers but it could not suitable to dynamic network. Rita [5] addresses the problem of the lack of reliability in peer-to-peer networks and proposes a number of algorithms that can provide reliability guarantees to peer-to-peer applications. Wei Hoo Chong Miet [6] proposes an alternative searching approach to optimize the existing

searching methods used in load balancing concept is the main consideration applied in this proposed system.

NOTEWORTHY PREVIOUS WORK:

The mainly related work to our paper is diff search algorithm which works as follows:

Research shows that only a small portion of volunteering peers provides the majority of the service in P2P networks. Based on the ultra peer overlay which contains the main portion of shared files, we propose the Diff Search algorithm. In the Diff Search algorithm, a query consists of two round searches. In the first round search, the query is only sent to the ultra peer overlay.

If the first round search fails in the ultra peer overlay, the second round search will be evoked to query the entire network. Here, we assume the entire P2P network is connected, which is consistent with the observation in the prerequisite of the Diff Search algorithm is that the ultra peer overlay consisting of content-rich peers is well formed in a P2P network. We show that the Diff Search algorithm can be self-utilized to shape an ad hoc P2P network into a two-tier hierarchical structure which clearly separates the content-rich ultra peer overlay from other peers.

To understand the working of the proposed work, we illustrate the ultra peer overlay construction by an example as follows: If a Diff Search query fails in the first round search, but succeeds in the second round, this implies two possible cases:

• The queried object is not shared by any ultra peers such that it can only be found by querying the entire network.

• The queried object is shared by ultra peers and some of the replies in the second round search are responses from the ultra peers, which implies that the failure of the first round search is caused by the incompleteness of the two-tier hierarchical structure:

Either because the leaf node is not connected to the ultra peer overlay or the ultra peer overlay is partitioned into multiple disjointed clusters. Otherwise, the query should have succeeded in the first round search in the ultra peer overlay. Then there is no need of the second round search.

In the second case, the incompleteness of the two-tier structure is detected and can be repaired by adding new connections between the querying peer and responding ultra peers.

Since any first round search failure due to the incompleteness of the hierarchical structure can be self repaired by the Diff Search algorithm, the two-tier structure can be incrementally evolved from a random status through frequent search behaviors initiated from a large number of querying peers. Our experiment shows that the algorithm converges very fast and the two-tier hierarchical structure can be evolved after each peer sends out about 20 queries on average.

MULTI-TIER DIFFERENTIAL SEARCH ALGORITHM:

This paper proposed the new technique for unstructured P2P networks to reduce the overheads traffic and to get the required data efficiently from the network. The proposed concept shows the better efficiency than previously described methods and diff. search methods which are based on single or two tire systems in addition applied semi-supervised decision approach make the system to take very refine decisions in categorization of ultra peers. At the end the simulation results shows the 90 percent reduction in overheads traffic in our system as compare to previous diff. search which shows 70 percent.

In this paper, we propose the Multitier Differentiated Search algorithm which is a modification over Diff Search algorithm to improve the search efficiency of unstructured P2P networks by giving higher querying priority to peers with high querying reply capabilities. Our experiment shows that the Multitier Differential Search algorithm can improve the search efficiency & can reduce the overhead traffic.

Algorithm	Differential Search Algorithm	Multi-tier Differential Search Algorithm
Memory	It required more Secondary memory (hard disk)	It required less Secondary memory(hard disk)
Speed	Speed of processing is slow	Speed of processing is more due to distributed request handling system

Bandwidth	Large Bandwidth required	Normal Bandwidth required
Flooding	More flooding	Less flooding
Reliability	Less reliable	More reliable

TABLE 1: COMPARISON BETWEEN THE DIFFERENTIAL SEARCH ALGORITHM AND MULTI-TIER DIFFERENTIAL SEARCH ALGORITHM

Our work is based on the observations that query reply capabilities are extremely unbalanced among peers (as some peers reply many times than other): according to empirical observation given in papers seven percent of peers in the Gnutella network share more files than all of those other peers can offer and 47 percent of queries are responded to by the top 1 percent of peers. This shows that only a small portion of peers actually delivers the large percent of traffic. This analysis is base for development of diff search algorithm which confine the search in only that peers they called them ultra peer overlay in our work we are extending the work proposed by diff search algorithm by breaking the system into multi layer ultra peer overlay each overlay consists of peers of similar query replaying capability so we increase the search level step by step instead of single step also the decision about creating the ultra peer is improves by applying semi supervised learning which provides the practically efficient overlay structure.

PROPOSED WORK:

In our proposed multi-tire differential search (MTDS) algorithm, the peers with high query answering capabilities will have higher priority to be queried. Because the query answering capabilities are extremely unbalanced among peers, a high query success rate can be achieved by querying only a small portion of a network. The search traffic is significantly reduced due to the shrunken search space. Our trace analysis and simulation show that the Multi-Tire Diff Search algorithm can save up to 60 percent of search traffic.

FOLLOWING CONCLUSIONS ARE DRAWN:

1. Multi-tier Diff Search algorithm significantly improves the performance of the network by using ultra peer overlay which basically created on the basis of maximum number of request serve in small time interval. By giving higher priority to such peers, which are more consistent than other peers we find better result.
2. Improves network efficiency by serving more request created by the client as compared to the request arrived by the client.
3. Reduces the flooding problem which creates due to the number of request generated by clients to the server
4. Time taken by the server to serve the request reduces due to the ultra peer overlay as a result it reduces the delay. So response time is also reduced.
5. Ultra peer overlay provides the scalability, reliability and interoperability of the network.

FUTURE SCOPE OF THE WORK:

During the simulation by using Multi-tier Diff Search algorithm it has been found that the ultra peer overlay improves the efficiency, speed, scalability, reliability and interoperability of the network.

1. Server security can be implemented by using firewall. Firewall allows only that data which is secure otherwise it denies the packet. Server security is required to get better performance and consistent network. We can work to find better technique to provide security to the server.

2. In Multi-tier Diff Search algorithm we are using the semi –supervised learning process on the basis of that learning clients automatically converted into the ultra peer overlay. We can improve the learning process by using other technique to get the better performance and select more consistent peers in the network.

3. We can improve the directory duplicity so that system become more faster.

REFERENCES:

- Qian Su Xuejie Zhang "A Peer-to-Peer Resources Search Algorithm Based on Small-World Model", Communications, Circuits and Systems Proceedings, 2006 International Conference, Guilin, 25-28 June 2006, pp: 1557 – 1561
- Fuyong Yuan Jian Liu Chunxia Yin" A Scalable Search Algorithm on Unstructured P2P Networks" Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing 2007, SNPD 2007, Eighth ACIS International Conference, Qingdao, July 30 2007-Aug. 1 2007
- Yu Jin Yan Liu Hongwu Zhao "Trust-based super node selection in peer-to-peer systems" Future Computer and Communication ICFCC), 2010 2nd International Conference on, Wuhan, 21-24 May 2010
- Wang Ping Qiu Jing Qiu Yu Hui" A Search Algorithm Based on Referral Trust in Unstructured P2P Systems" Electronic Commerce and Security, 2009. ISECS' 09, Second International Symposium on, Nanchang, 22-24 May 2009
- Chen Wang Li Xiao"An Effective P2P Search Scheme to Exploit File Sharing Heterogeneity", Parallel and Distributed Systems, IEEE Transactions on, Michigan, Feb. 2007
- L. A. Adamic, R. M. Lukose, and B. A. Huberman, "Local search in unstructured networks," Review chapter to appear in Handbook of Graphs and Networks: From the Genome to the Internet, S. Bornholdt and H.G. Schuster (eds.), Wiley-VCH, Berlin, 2004.
- S. Androutsellis-Theotokis and D. Spinellis. "A survey of peer-to-peer content distribution technologies," ACM Computing Surveys, 36(4):335371, December 2003
- D. S. Bernsteing, Z. Feng, B. N. Levine, and S. Zilberstein. "Adaptive Peer Selection," Proceedings of the 2nd International Workshop on Peer-to-Peer Systems (IPTPS), Berkeley, California, February 2003.
- BitTorrent. http://www.bittorrent.com/.
- E. Cohen and S. Shenker. "Replication Strategies in Unstructured Peer-to-Peer Networks." Proceedings of the 2002 conference on Applications, technologies, architectures, and protocols for computer communications (ACM Sigcomm 2002), pp. 61-72, 2002.

- S. El-Ansary, L. O. Alima, P. Brand, and S. Haridi, "Efficient broadcast in structured P2P networks," 2nd International Workshop on Peer-to-Peer Systems (IPTPS '03), Berkeley, CA, February 2003.
- A. C. Fuqua, T. Ngan, and D. S. Wallach. "Economic Behavior of Peer-to-Peer Storage Networks," Workshop on Economics of Peer-to-Peer Systems (Berkeley, California), June 2003.
- M. Gupta, P. Judge, and M. Ammar. "A Reputation System for Peer-to-Peer Networks." In Proceedings of the NOSSDAV'03 Conference, Monterey, CA, June 1-3 2003. pp 67]
- D. Qiu and R. Srikant. "Modeling and Performance Analysis of BitTorrent- Like Peer-to-Peer Networks," Proceedings of ACM SIGCOMM, Portland, Oregon, September 2004.
- http://en.wikipedia.org
- http://www.gnutella.com
- http://rfc-gnutella.sourceforge.net/src/rfc-0_6-draft.html
- http://rfc3684.x42.com