Augmentable Search Accomplishment in P2P Networks Using Guided Protocol Model (ASAPN-GPM)

J. Thirupathi, Assoc. Prof. Dept of Computer Science and Engineering, Samara University

Alemayehu Abere Lecturer Dept of Computer Science and Engineering, Samara University

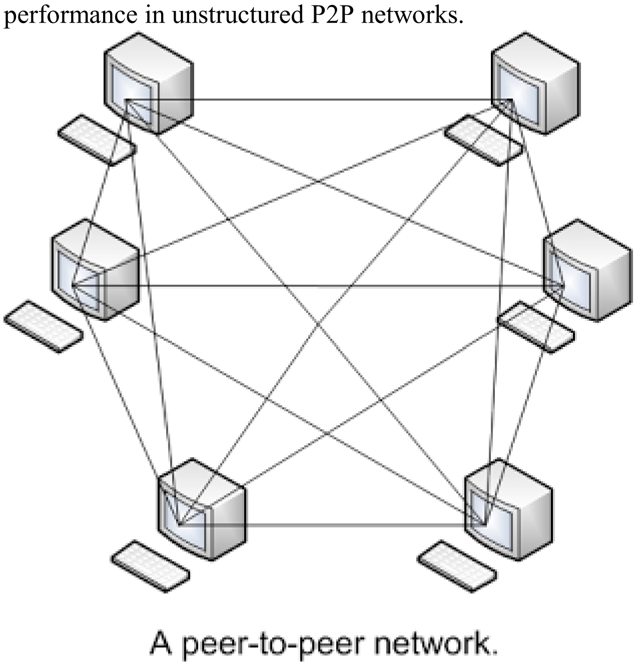
K. Ramakrishna, Asst. Prof. Dept of Computer Science and Engineering, Samara University

Abstract: Efficient search is a challenging task in unstructured peer-to-peer networks. In this paper Augmentable Search Accomplishment in P2P Networks Using Guided Protocol Model (ASAPN-GPM) is proposed that adaptively performs a query searching through guided protocol method. Unstructured P2P networks can be constructed very efficiently and are therefore considered suitable to the Internet environment. However, the random search strategies adopted by these networks usually perform poor with a large size networks. Most of the P2P networks construct loosely coupled overlay on the top of the internet based on the physical network constraints without taking user preferences or relationship into account. It leads to high inefficiency in their search algorithms, which mainly relies on the simple flooding or random walk strategies. In this paper, we present the architecture of Augmentable Search fully decentralized P2P structure as well as efficient protocol model for P2P search technique to improve the searching efficiency. A search protocol and a routing table updating protocol are further proposed in order to get the search process through self organizing the P2P network. Both theoretical and experimental analyses are conducted and defined effectiveness of our model.

Index Terms: **Peer Search, Guided Protocol Model Peer to Peer Networks, Search model,**

I. Introduction

A network of personal computers, each of which acts as both client and server, so that each can exchange files and email directly with every other computer on the network. Each computer can access any of the others, although access can be restricted to those files that a computer's user chooses to make available. Peer-to-peer networks are less expensive than client/server networks but less efficient when large amounts of data need to be exchanged. PEER-TO-PEER (P2P) networks establish loosely coupled application-level overlays on top of the Internet to facilitate efficient sharing of large amount of resources. One fundamental challenge of P2P networks is to achieve efficient resources discovery. In the literature, many P2P networks have been proposed in an attempt to overcome this challenge. Those networks can be largely classified into two categories, namely, structured P2P networks based on a distributed hash table and unstructured P2P networks based on diverse random search strategies. Without imposing any conditions over the network topology, unstructured P2P networks can be constructed very efficiently and have therefore attracted far more practical use in the Internet than the structured networks. Peers in unstructured networks are often termed blind, since they are usually incapable of determining the possibility that their neighbor peers can satisfy any resource queries. An undesirable consequence of this is that the efficiency of distributed resource discovery techniques will have to be compromised. In practice, resources shared by a peer often exhibit some statistical patterns. The fundamental idea of this paper is that the statistical patterns over locally shared resources of a peer can be explored to guide the distributed resource discovery process and therefore enhance the overall resource discovery performance in unstructured P2P networks.



A peer-to-peer network.

Manuscript received Nov 27, 2014.

Mr. ***J Thirupathi***, *Assoc. Professor in Department of Computer Science and Engineering ,College of Engineering and Technology, Samara University ,Samara, ETHIOPIA,*

Mr. Alemayehu Abere *Lecturer and Head of the Dept. of Computer Science and Engineering, College of Engineering and Technology, Samara University, Samara, ETHIOPIA.*

Mr. ***K. Ramakrishna***, *Asst. Professor in Department of computer science and engineering, College of Engineering and Technology, Samara University ,Samara, ETHIOPIA,*

ISSN: 2278 – 7798

Three essential research issues have been identified and studied in this paper in order to save peers from their blindness. For ease of discussion, only one important type of resources, namely, data files will be considered.

This paper proposes a routing table updating protocol to support our search protocol through self organizing the whole P2P network into a small world. Different from closely related research works that are also inspired by the small-world model, in order to reduce the overall communication and processing cost, in this paper, the updating of routing tables are driven by the queries received by each peer. In a P2P network, queries handled by a peer may be satisfied by any peer in the network with uneven probability.

II. System Analysis

A . Existing System

Peers in unstructured P2P networks to choose their neighbors and locally shared files, using Flooding techniques. In purely unstructured P2P networks such as blind search through flooding mechanisms is usually explored for resource discovery. To find a file, a peer sends out a query to its neighbors on the overlay, until the query has traveled a certain radius. Despite its simplicity and robustness, flooding techniques, in general, do not scale. In large networks, the probability of a successful search may decrease dramatically without significantly enlarging the flooding radius.

The first research issue questions the practicality of modeling users' diverse interests

The second research issue considers the actual exploration of users' interests as embodied by UIM

The third research issue has been highlighted with the insight that the search protocol alone is not sufficient to achieve high resource discovery.

Deficiency of an Existing System:

- Blind Search.
- Future reference is not present in routing table.
- Delay due to absence of Routing Updating table.

B. Proposed System

In order to improve search performance, guided search. The key problem is what information is actually eligible to guide the search. used interest-based locality as the general search guidance. The basic assumption is that if a peer p0 has a particular file required by another peer p, and then p0 is likely to have other files to be requested by p in the future. According to previous successful queries, shortcuts from peer p to several peers p0 are established in order to expedite subsequent search processes.

Proposed System Features :

- Guided Search.
- Routing updating table.
- Fast Search Technique based on UIM.

III. Problem Formulation

A Objectives

A variety of techniques for locating resources (files) in P2P networks have been devised over the last few years. Initial approaches on a centralized architecture with designated indexing peers. Due to the single point of failure problem and the lack of scalability, recent research in P2P networks focuses more on distributed search technologies. Unstructured P2P networks blind search through flooding mechanisms is usually explored for resource discovery. Resource search, a peer sends out a query to its neighbors on the overlay. These neighbors, in turn, will forward the query to all of their neighbors until the query has traveled a certain radius. Despite its simplicity and robustness, flooding techniques, in general, do not scale. This process make a time consume, traffic and unwanted search.

B The Guided Protocol Model

The Search Protocol in this section, a file search protocol is presented to regulate the activities of every peer p in a P2P network upon receiving a query q.

Algorithm for Guided search protocol:

1. Add peer p to the search history hq
2. If filw fq is locally stored in peer p, inform peer pq that search is successful
3. Else if the size of hq is greater than TTL, inform peer pq that search is failed
4. Else

a. Order all the routing entries Ep=<pv,fv> of Rp decreasingly based on pr(fq|fe)

b. Iterate over every routing entry Ep=-<pe,fe> starting from the one with highest pr(fq|fe)

c. If no entry is chosen at step 4.b forward query q on to pe with highest pr(fq|fe)

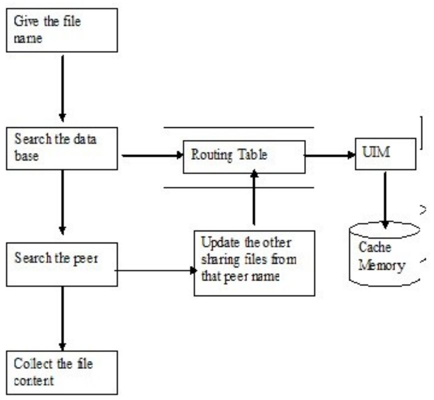
In case that the peer chosen to forward the query already exists in the search history hq, another peer with a relatively high probability of satisfying the query will be chosen instead. It follows essentially a greedy search strategy. Upon receiving a query q, a peer p will try to find, among its neighbor peers, the one that is most probable of having the requested file fq according to UIM. If all neighbor peers have been visited before, then the peer that enjoys the highest probability will be selected again for query processing.

C. The Protocol for updating routing tables:

In this section, a protocol for updating routing tables will be presented and analyzed. An uneven updating problem will also be highlighted, and a filtering mechanism will be further introduced to tackle this problem.

1. Repeat for every peer p in the search history hq
2. Add a new entry Ep=<pi,fq> to the routing table Rp of peer p
3. While the size of Rp is greater than Br
   a. With respect to each entry Ep=<pe,fe> , calculate the interest distance d(pe,p)
   b. Select an entry Ep=<pe,fe> with probability Proportional to d(pe,p)
   c. Remove Ep from the routing table Rp

ISSN: 2278 – 7798



Graphical representation of the "flow" of data through business functions or processes

The LRU strategy: The routing entry that is least recently used to forward queries will be dropped. .

The ECCR scheme: With a certain probability Pre, the least recently used routing entry will be dropped. Otherwise, the neighbour peer p0, which has the longest interest distance from peer p, will be removed from Rp. The distance-centric (DC) strategy: Either the peer p0, which has the longest interest distance from peer p, or another peer p00, which has the second longest distance, will be removed from Rp of peer p, depending on a probability Prd.

The implementation can be preceded through Socket in java but it will be considered as peer to peer communication .For proactive routing we need dynamic routing. So java will be more suitable for platform independence and networking concepts. For maintaining route information we go for SQL Server as database back end. The first research issue questions the practicality of modeling users' diverse interests. To solve this problem, we have introduced the user interest model (UIM) based on a general probabilistic modeling tool termed Condition Random Fields (CRFs) . With UIM, we are able to estimate the probability of any peer sharing a certain resource (file) fj upon given the fact that it shares another resource (file) fi. This estimation further gives rise to an interest distance between any two peers. The second research issue considers the actual exploration of users' interests as embodied by UIM. To address this concern, a greedy file search protocol is presented in this paper for fast resource discovery. Whenever a peer receives a query for a certain file that is not available locally, it will forward the query to one of its neighbors that has the highest probability of actually sharing that file .The third research issue has been highlighted with the insight that the search protocol alone is not sufficient to achieve high resource discovery performance.

IV Results and Analysis

Search Efficiency (SE) is an objective metric that affords general supervision to characterize the efficiency of the system. In the proposed scheme, Searching Efficiency (SE) is determined from three distinct factors such as Query Efficiency (QE), Responsiveness and Reliability.

SE = □ Query Hits [t] /t * 100% * (Success rate/ Query Message * R)

Query Efficiency (QE) Query Efficiency is evaluated from a number of Query Hits (QH) divided by the number of Query Messages (QM). In the best searching algorithm possess a large number of QH occurs even with few query messages.

Query Efficiency = Query Hits [t] * 100% / Query Message * R

Where QH is the query hits at any peers in hop h, T is TTL varies from (1 -7), QM is the total number of query messages generated during the query, R is the replication ratio of the queried object.

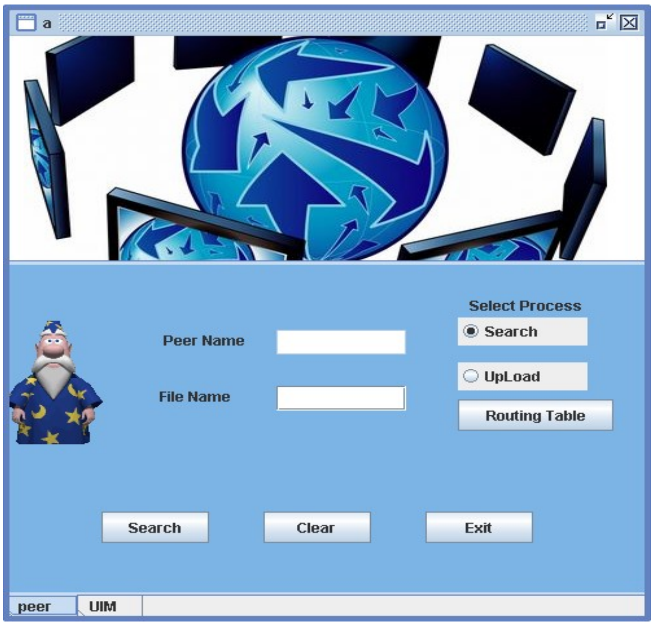
Network traffic

In an unstructured P2P networks, network traffic overhead is raised due to the total number of queries forwarded or received by peer during the search. Hence, average network traffic is the total amount of traffic generated during the whole number of queries processed in the network divided by the total number of queries totally generated by peers in the network.

Average network traffic = (Ti / (Qi)

Where Ti is the total amount of traffic generated by queries and Qi is the total number of queries generated by peer i.

User interests are varying dynamically according to the time. Therefore, it oscillates the popularity of the data item. Hence, the data popularity is updated periodically via incremental updates to achieve high data availability.



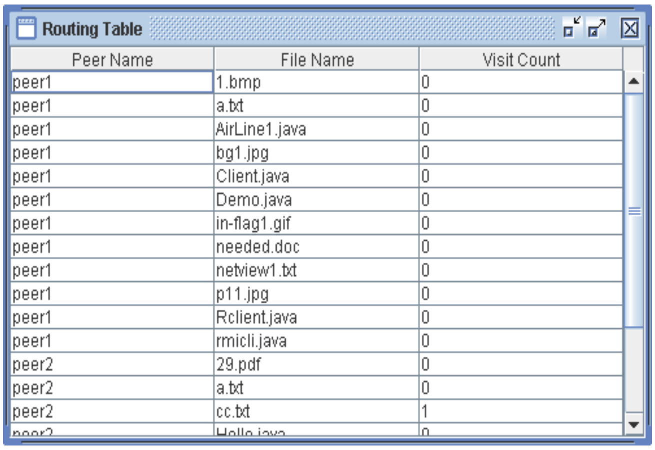
Output screen 1

In the above fig we have to give peer number and file name after entering file name select search that will give the result as how many peers it is travelling that is there in the next screen

So by using search efficiency we can find the efficiency of the system.

In order to determine the number of times requested of the data item within a certain time, a time stamp is assigned to each entry of the data to indicate the time of its last update. Here, an incremental update is successfully incorporated to update the latest changes in the data popularity within a certain time. Peers search the object in the network through message flooding. To flood a message an inquiry peer simultaneously sends the same message to the multiple recipients.

ISSN: 2278 – 7798



Peer Name	File Name	Visit Count
peer1	1.bmp	0
peer1	a.txt	0
peer1	AirLine1.java	0
peer1	bg1.jpg	0
peer1	Client.java	0
peer1	Demo.java	0
peer1	in-flag1.gif	0
peer1	needed.doc	0
peer1	netview1.txt	0
peer1	p11.jpg	0
peer1	Rclient.java	0
peer1	rmicli.java	0
peer2	29.pdf	0
peer2	a.txt	0
peer2	cc.txt	1
peer2	Hello.java	0

Output screen 2

Instead of forwarding the message to the neighbours first it searches in its local storage. If it can provide the requested object then it can either directly send the object or returns the objects to the overlay path where the message traverses. So that it can minimize the overlay path length between any two peers to reduce the query response time.

V CONCLUSION

In this paper, we have shown that the search performance in unstructured P2P networks can be effectively improved through exploiting the statistical patterns over users' common interests. Our solution toward enhancing search performance was presented in three steps: 1. A UIM has been introduced in order to capture users' diverse interests within a probability-theoretic framework. It leads us to further introduce a concept of interest distance between any two peers. 2. Guided by UIM, a greedy protocol has been proposed to drive the distributed search o queried files through peers' local interactions. 3. Finally, a routing table updating protocol has been proposed to manage peers' neighbor lists. With the help of a newly introduced filtering mechanism, the whole P2P network will gradually self organize into a small world.

Theoretical analysis has also been performed to facilitate our discussion in this paper. Specifically, the search protocol was shown to be quite efficient in small-world networks. Succeeding analysis further justifies that by using our routing table updating protocol, the P2P network will self organize into a small world that guarantees search efficiency.

To conclude, it is noticed that our solution presented in this paper can be extended from several aspects. For example, UIM can include extra attributes that characterize a network user (e.g., gender, age, and occupation). Meanwhile, live data from real P2P networks are expected to be collected and applied to further test the effectiveness of our protocols. These and other efforts will be left here as our future work.

VI FUTURE ENHANCEMENT

Our future work considers the implementation and evaluation of the proposed suboptimal solution on a wireless system. To this end, we will consider a wireless sensor network where the nodes are constrained in computing power as well as power consumption.

REFERENCES

[1] S. Tewari and L. Kleinrock, "Optimal Search Performance in Unstructured Peerto- Peer Networks with Clustered Demands," IEEE J. Selected Areas in Comm., vol. 25, no. 1, 2007.

[2] M. Khambatti, K. Ryu, and P. Dasgupta, "Structuring Peer-to-Peer Networks Using Interest-Based Communities," Proc. Int'l Workshop Databases,Information Systems and Peer-to-Peer Computing (P2PDBIS), 2007.

[3] E. Cohen, A. Fiat, and H. Kaplan, "Associative Search in Peer-to-Peer Networks: Harnessing Latent emantics," Proc. IEEE INFOCOM, 2008.

[4] J. A. Pouwelse, P. Garbacki, J. W. A. Bakker, J. Yang, A. Iosu D. Epema, M.Reinders, M. R. van Steen, and H. J. Sips, "Tribler: A socialbased based peer to peer system," Proceedings of the Fifth International Workshop on Peer to- Peer Systems (IPTPS), 2006.

[5] Ching-Ju Lin Yi-Ting Chang, Shuo-Chan Tsai, Cheng-Fu Chou, "Distributed Social-based Overlay Adaptation for Unstructured P2P Networks", Proceedings of the IEEE International Symposium on Global Internet, pp:1-6, 2007.

[6] Christian von der Weth, Anwitaman Datta " Multiterm Keyword Search in NoSQLSystems" IEEE Internet Computing, Vol 16, No1, pp 34-42, 2012.

[7] B. Bloom, "Space/Time Trade-offs in Hash Coding with Allowable Errors," Communications of the ACM, Vol. 13, No. 7,pp. 422–426, 1970.

[8] Chen Wang, Li Xiao, Yunhao Liu, and Pei Zheng, "DiCAS: An Efficient Distributed Caching Mechanism for P2P Systems" IEEE Transactions on Parallel and Distributed Systems, Vol 17, No 10, pp 1097-1109, 2006.

[9] E. Herskovits, "Computer-Based Probabilistic-Network Construction,"

[10]R.A.Horn and C.R. Johnson, Matrix Analysis. Cambridge Univ.Press, 1990.

[11] H. Jin, X. Ning, and H. Chen, "Efficient Search for Peer-to-Peer Information Retrieval Using Semantic Small World," Proc. Int'l Conf. World Wide Web (WWW '06), pp. 1003-1004, 2006.

[12] M. Khambatti, K. Ryu, and P. Dasgupta, "Structuring Peer-to-Peer Networks Using Interest-Based Communities,"Proc.Int'l Workshop Databases, Information Systems and Peer-to-Peer Computing (P2PDBIS), 2003.

[13]J. Kleinberg, "The Small-World Phenomenon: An Algorithmic Perspective," Proc. 32nd ACM Symp. Theory of Computing (STOC), 2000.

Mr. Jangapally Thirupathi, post graduated in C.S.E M.Tech from JNTUH and graduated in C.S.E B.Tech from JNTUH, AP, INDIA, having 8+ years of teaching experience, presently working as an Associate Professor in Department of computer science engineering, College of Engineering and Technology in Samara University ,Samara, ETHIOPIA, Research interests include, mobile ad-hoc networks, cloud computing data mining and information security .

Mr. Alemayehu Abere Working as Head of the Dept Computer Science and Engineering, Samara University, Samara Ethiopia.. He is having 8+ years of teaching experience. His research interests includes Cloud Computing, Data Mining, Information Security, Software Testing, Wireless Networks and Software Quality.

Mr. K. Ramakrishna, post graduated in (CSE-SE) M.Tech from JNTUH and graduated in (IT) B.Tech from Kakatiya university, Warangal ,AP, INDIA, he has 5+ years of teaching experience, He is working presently as an Asst .Professor in Department of computer science engineering ,College of Engineering and Technology in Samara University ,Samara, THIOPIA, his research interests include mobile ad-hoc networks ,information security and communication

ISSN: 2278 – 7798