Journal of Theoretical and Applied Information Technology
28th February 2014. Vol. 60 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

TOWARDS POPULARITY AWARE HYBRID CACHING TO IMPROVE SEARCH IN SOCIAL RELATIONSHIP BASED P2P NETWORKS

[1] G. RAMACHANDRAN, [2] DR. K. SELVAKUMAR

[1,2] A Department of Computer Science Engineering, Annamalai University, Chidambaram, Tamilnadu,India, Pin :608002

E-mail: [1] gm_rama@yahoo.com , [2] kskaucse@yahoo.co.in

ABSTRACT

The widespread use of recent Peer-to-peer (P2P) file sharing has been mainly influenced by the scalability of their architecture and high versatile search mechanisms. However, 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 adaptable fully decentralized social based P2P overlay as well as efficient cognitive community P2P search technique to improve the searching efficiency. This paper proposes a Dynamic Overlay Adaptation (DOA) algorithm, which creates social communities by connecting group of peers having similar interest within the shortest path based on the social relationship between them. The basic premise of the creation of social communities is that generated queries are most probably satisfied within its own community. Therefore, it significantly improves the searching efficiency with a higher success rate and less response time. In addition to, this paper proposes a Popularity aware Hybrid Caching (PaHC) which caches the file based on its size to improve cache performance. To accomplish high data availability, popularity aware strong data consistency is proposed that eliminates the duplicate content replication among the peers. Experimental evaluation reveals the effectiveness and efficiency of the proposed approach in terms of success rate, network traffic, response time and cache hit ratio.

Keywords: *Unstructured P2P Networks, Peer Search, Social based P2P Overlay, Popularity Aware Hybrid Caching*

1. INTRODUCTION

In the last decade, the P2P system become popular widely due to its deployment in resource sharing over a large population of networked computers in a most efficient and cost effective manner. It facilitates efficient resource sharing in P2P network and establishes loosely coupled application-level overlay on top of the Internet [1]. These overlays are categorized into structured and unstructured networks. In a highly dynamic and unpredictable network, the structured P2P incurs high maintenance overhead to maintain the consistent and distributed index for the large scale network [2]. On the contrary, unstructured P2P systems dynamically adapt the continuous node membership changes without any stringent constraints over the network topology. However, peers and resources are randomly placed without any predetermined structure leads to incur either high search overhead or large network traffic for resource discovery [3]. Here, peers have no awareness about their neighbor peers that degrade the performance of P2P searching. Therefore, effective searching mechanism is essential to locate the resource which incurs low communication overhead. Consequently, clustering approaches are evolved to construct semantic overlay network in the form of “groups” or “communities” based on the user interest [4]. However, it requires a centralized server to create and maintain the community that incurs additional maintenance overhead.

Recently, the focus has shifted to social-based P2P sharing system, where peers are treated as social partners rather than consider them as solitary rational agents to improve searching efficiency by eliminating the problem of free-riding. In fast growing world, Online social networks offers a platform to establish and organize the social connection with their friends anywhere in the world based on shared affinities such as

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

hobbies, interests, location groups and overlaps in friendship circles [5]. This work exploits this notion to construct a social based P2P overlay on the top of the unstructured P2P network (e.g. Gnutella) for improving the performance of resource discovery. , peers are people and links between people are considered as social relationships. The main aim of this work is to create the community based on the social relationship. Therefore, it does not need extra overhead to maintain the relationship between peers in the network. In addition to, Cognitive community query searching is performed based on the social relationship to achieve effective searching with less response time and reduced network traffic overhead. To improve the cache performance, popularity aware hybrid caching is proposed to cache the data item based on its size. Furthermore, this paper proposes popularity aware strong data consistency to accomplish high data availability among the peers in the network.

1.1 Contributions of the paper

This paper presents a design and implementation of PaHC approach in the social community based P2P File search paradigm by accumulating the social functionality. The contribution of this paper is five folds

i) Distributed Overlay Adaptation Algorithm constructs the well-connected social communities based on the notion of similar interest using social relationship.
ii) Interest based Multi-keyword indexing builds a novel multidimensional hash table like a structure for efficient query lookup rather than broadcasting of queries throughout the network.
iii) Cognitive community P2P searching performs effective query processing in three levels of hierarchical searching to improve query efficiency.
iv) Popularity aware hybrid caching mechanism is proposed which caches the data item based on its size and popularity to improve the search performance.
v) Popularity aware strong data consistency proposed that effectively replicates the content based on the popularity of the data item to improve high data availability.

1.2 Paper organization

The rest of the paper is organized as follows: In Section 2, we review the related work. Section 3 elucidates the proposed system architecture and also proposes three efficient algorithms such as Distributed topology adaptation algorithm, cognitive community searching and popularity aware hybrid caching to improve the effectiveness of P2P searching. The performance of the proposed system is evaluated extensively using performance metrics and also compares the performance of the proposed approach with existing approaches in Section 4. Finally, we concluded the work in Section 5.

2. RELATED WORK

This section provides a comprehensive review of some related works which applies the social network information to improve the performance of P2P searching. Accordingly, it reveals insights and rationale behind in each of the already existing approaches which provides motivation for this work. The author extends a new P2P searching approach from the traditional random walk search model to facilitate effective searching based on social information [6]. It models the online social relationship between peers according to strength of the relationship of the other peers in the network. However, the forwarding strategy relies on the knowledge about the mutual contribution among the peers. Rim Haw et al., proposed an efficient social P2P management scheme to perform interesting keyword based P2P searching [7]. Each peer encompasses an interesting keyword management table to maintain attractive keywords of all of its corresponding neighbor's peers in the network. In order to locate the resource, peer performs keyword lookup in its maintenance table. In case, lookup resource is not available among the neighbor peers then query is flooded over the network. However, it is difficult to update the maintenance table over a large scale dynamic environment

In [8], OverSoc approach constructs a structured P2P overlay to organize a social network efficiently which enables interactive networking activities in a decentralized manner. This approach provides a clean abstraction for privacy purposes at the same time eliminates the intrinsic limit of centralized client/server architecture deployed in social networks. Panagiotiset al published an article [9] proposed a small world theory to construct an acquaintance network with the set of peers having similar interest. This approach designed the Small World Indexing Model to store the descriptions of friends and acquaintances in order to find other small world peers having similar characteristics. Here, a central server gathers the description vectors of all peers and also establishes the overlay

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

links between each pair of peers based on the distance between them. In the case of a large scale networks, a centralized method incurs heavy network traffic and also difficult to handle churn in P2P systems efficiently.

Xu Cheng et al., developed a NetTube model to improve cooperation among peers based on a social network structure [10]. For effective video distribution, this approach constructs bi-overlay by exploiting Friend-to-Friend (F2F) social network relationship. It deploys a space/time efficient indexing scheme with data structure called as Bloom filter, which maintains a record of cached videos to locate the videos searched by the clients in the network. In addition to, a social network assisted pre-fetching strategy is introduced to realize fast and smooth transition with delay minimized video playbacks. Pouwelse et al., designed a Tribler to construct a robust and social overlay automatically on top of the most popular P2P file sharing paradigm BitTorrent. Here, friendship and trust phenomena establishes overlay links with peers who have similar interest [11]. This approach argued that no overhead required to maintain the membership relation between the peers in bloom filters in order to reduce the bandwidth requirement significantly for epidemic content distribution. This scheme relies on the Buddycast mechanism for user's similarity measurement by comparing their preference list which maintains a record of recently downloaded videos. However, this method acquires a problem for the new users who possess only a few downloads because it is difficult to get an accurate similarity measure. This distributed approach [12] proposed an overlay adaptation algorithm, which constructs the social-based overlay with a group of peers having similar interest. It is connected through the shortest path in a dynamic and distributed manner. Here, the distance measure properly quantifies the similarity of peers using a random walk technique in order to reduce overlay construction overhead significantly and also handle the dynamic churn efficiently. It also defines a user profiling method that exploits the social relationship between the peers to improve the effectiveness of the content query service.

3. PROPOSED P2P ASSISTED QUERY SEARCHING IN ONLINE SOCIAL NETWORKS

This section presents a two-tier overlay construction based on social community using distributed overlay adaptation algorithm in a fully distributed fashion. It proposes a cognitive community query searching to perform three-level hierarchical similarity searching in the small metric space of super peer overlay. Interest based Multi-keyword indexing approach is proposed to build a novel multidimensional hash table like a structure for efficient query lookup within the small search scope. Furthermore, popularity aware hybrid caching is proposed to manage the cache performance effectively in a large scale P2P networks.

3.1 Social based P2P overlay construction

The proposed approach is a P2P assisted query searching in Online Social Network (OSN), where social based P2P overlay is constructed in the form of two-tier super peer network using DOA algorithm [12].

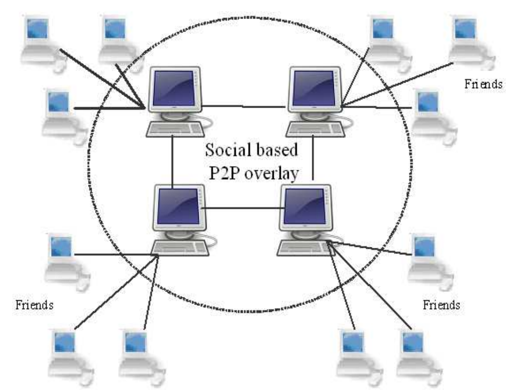

Figure 1: Social based P2P overlay

Figure 1. shows the constructed social based P2P overlay network. In the upper tier, a set of super peers having high connectivity, high bandwidth resources and computation power are unified together and form a P2P overlay. Moreover, the lower tier assembles the set of client peers in unstructured fashion where each peer establishes overlay links with the similar interest super peer distributively by exploiting social phenomena such as friendship and trust [8].

3.1.1 Distributed overlay Adaptation Algorithm

The distributed overlay adaptation algorithm is a social community based P2P overlay network. It is constructed as a two tier network where upper tier comprises a set of super peers and then lower tier composes of its associate client peers (friends). This algorithm dynamically computes the similarity score and adapts the topology of the P2P network. Therefore, the peers that share similar interests are formed into well-connected social communities based on social relationship. Indeed, peers having interest only on

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

some explicit types of content organized in the same social community increases search efficiency and success rate with reduced searching scope. Each peer computes the degree of similarity to form the social community. The cosine similarity measure determines the degree of similarity between the two peers P_1 and P_2 using the information collected by the social network's Friend-to-Friend (F2F) relationship among the peers in the network [12]. The degree of similarity is computed as follows

$$Sim(P_1, P_2) = \cos(\vec{w1}, \vec{w2}) = \frac{w_{1.} w_2}{w_1^2 * w_2^2} \quad (1)$$

Here, w1 and w_2 are the weight of user preferences for the peer P_1 and P_2 respectively. Based on the degree of similarity between the peers, the whole network is logically divided into a multiple number of social communities where each social community consists of only one high connectivity super peer and the set of similar interest client peers (friends) using social relationship. In each of the social community, one super peer is selected to maintain the bounded list of peer indices of its connected client peers (friends) in the overlay. It has high capability to serve a large number of files for the requesting peer with in the small searching scope that facilitates scalable P2P searching.

3.2 Interest based Multi-keyword Indexing

This approach proposes an interest based multi-keyword indexing that builds a novel multidimensional hash table like a structure for efficient query lookup. It completely eliminates the need for broadcasting of queries throughout the network. It implements a key-to-value map in their core to support exact match queries that provide relevant data for the user generated query. Indexing provides high expressiveness for structural representation of data and suitable for application characteristics in large scale P2P networks. It is extremely easy to implement and maintain a hash table that stores the indexed information as a key value pairs where the key denotes the interesting keywords (I_k) in the filename F and the corresponding value of indexed attributes representing the file name. Interesting keywords are derived from the content of the text documents or file name. Thus, the index entry for a keyword points out all the filename that contains the content related to the keyword [13]. The set of interesting keywords in a peer p is represented as S = {Ik_1, Ik_2, Ik_3}. Furthermore, hashing identifier h (Ik) is assigned to each keyword Ik using MD5 hashing algorithm. In order to identify the file over the set of keywords, the hashed identifiers of keywords are combined as follows

Key (S) = h (Ik_1) h (Ik_2), h (Ik_3) - (2)

3.2.1 Index Advertising

Each peer maintains a local index that is the key value pairs of its own local files in a hash table. In order to achieve collaboration in the P2P network for supporting membership queries, an index advertising mechanism is proposed that allow friends advertise its state information such as file name and its corresponding keywords to its connected super peer using the operation publish (Keywords, Hostname). Each super peer encompasses an aggregated hash table like data structure, which is a simple space-efficient randomized data structure. It maintains the compact representation of a set of keywords and caches file in its social community. It consists of n-array bits to store records provided for locating the resources directly using m independent hash functions {h_1 (Ik_1), h_2 (Ik_2)… h_m (Ik_n)} without any complex query searching schemes [14]. Thus, each super-peer has a global image about the set of keywords and caches file in its social community. Therefore, it has enough information to forward the query to the corresponding client peer within the small search scope. It consumes less bandwidth for information distribution and leads to achieve high scalable query searching.

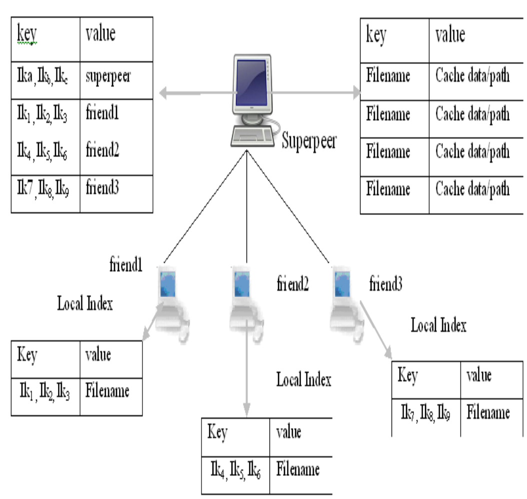

Figure 2: Interest based Multi-keyword Indexing

Journal of Theoretical and Applied Information Technology
28th February 2014. Vol. 60 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3.3 Towards Popularity aware Hybrid aching

Most of the existing caching approach caches all the query response in the entire peers along the query returning path to serve the future queries [15]. However, it incurs high maintenance overhead due to the high replication of the data item and its large size. In order to achieve optimized caching with high flexibility, this paper proposes a popularity aware hybrid caching approach that selectively caches the query responses based on the size of the data item in the highly connectivity super peer of each community. Intuitively, it yields a high cache hit rate with reduced network traffic rather than caching of all passing query responses. In addition to, it proposes a popularity aware replication mechanism that achieves strong data consistency for high availability and also minimizes the replication of duplicate content among the peers in the network.

3.3.1 Hybrid Caching

Hybrid caching enhances the cache performance that combines the advantage of both caching technique such as cache data and cache path. At the same time, it avoids the weakness based on the size of the data item. Most of the existing caching scheme works in the assumption that all the data items are equal in size. This work proposes the hybrid caching approach that caches the high popularity data item over the high connectivity super peer, in each peer of the social community based on its size.

```
For every data item Di
    for (i=1; i<=n; i++) {
        if (size (Di) < 1MB) {
            cacheData( ) {
            cache Di → super peer in social community;
            update for every short TTL;
            }
        }
        else {
            cachePath( ) {
            cache path (Di) → super peer in social community;
            update for every long TTL;
            }
        }
    };
```

Figure 3: Hybrid Caching

Based on the size of the data item, the super peer undergoes two cooperative techniques such as cache data and cache path techniques in order to cache the data item. The cache data technique caches the data item in its cache, whereas the cache path technique caches only the path of the data items. In this approach, the size of the data item is an influential factor to decide whether to cache data or its path to achieve better cache performance. If the size of the data item is greater than the threshold (T), then super peers perform cache path technique to cache the path of the peer in which the data item resides. Furthermore, the data of the path is updated for the long interval of time due to the less dynamics of the P2P system. On the other hand, if the size of the data item (text files) less than a certain threshold, the super-peer undergoes cache data technique to cache the whole data item (text files) and it is expected to have a shorter update interval (TTL), therefore, updated frequently [16].

3.3.2 Popularity based Strong Data consistency

In order to achieve strong data consistency for high data availability, replication of cached files is performed based on the popularity of the data items for which request received from the user within a certain time. Therefore, super-peer in the social community overlay maintains the cached files that comprise of two vital tasks: replicating data item basis on their popularity and updating the data popularity according to time.

i) Popularity computation using Bit Vector

The popularity of the data item is computed from the number of times that was requested by the other peers in the network. For that, query history is needed to trace out the unique ID of the data item in the previously transmitted query hit messages. Popularity computation is difficult to estimate the number of times accessed by other peers in the network due to the overall data distribution in the overlay network. Therefore, a bit vector B_D of length L is associated with each data item (D) to quantify the popularity. When each time data item is accessed, the bit vector B_D is shifted one bit to the right and append 1. Thus, the popularity of the data item is computed from the number of bits set into 1 during query processing. The popularity of the data item is decomposed according to time so that the bit vector B_D is shifted one bit to the right and append 0 periodically. If the number of a set of bits in B_D exceeds the system threshold, an index of the data item is maintained in a hash table. Otherwise, the index of the data item removed from the hash table. Depending upon the threshold value, perform the periodic update of data item according to query processing.

Journal of Theoretical and Applied Information Technology
28th February 2014. Vol. 60 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

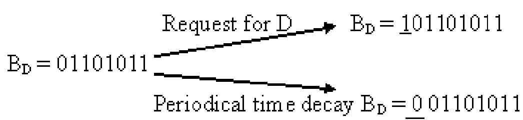

Figure 4: Popularity Computation using Bit Vector

ii) Periodical update of data popularity

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. 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. Thus, this approach periodically updates the data popularity from recent query history.

3.4 Cognitive Community Query Searching

In this approach, two-tier hierarchical structure is shaped well to perform an effective three levels of searching. In this approach, super-peers are efficiently configured to maintain the set of interesting keywords and it caches file of connected client peers (friends). It facilitates effectively forward the query to friends having similar interest in the social community without much communication overhead and traffic. Since, there is a high possibility for matching files found within the same community of two hops from the requestor. It performs keyword matching between the multiple keywords $\{k_1, k_2 \ldots k_n\}$ in the query Q and set of keywords $\{Ik_1, Ik_2 \ldots Ik_n\}$ in the hash table to discover matched query results. It achieves optimal tradeoff between the network traffic and success rate by effectively maintaining the index in the case of high dynamic environment. In addition to, the F2F social network provides social phenomena such as friendship and trust to achieve quite effective searching.

i) Local level search

The source peer initiates a query that invokes the first round search in order to locate a file. Here, requested peer checks its own hash table to find whether it has the keywords in the user generated query. If the first round of search fails, the second round of the search is initiated [17].

ii) Community level search

In the second level of search, requested peer forwards the query into the super peer of its located community. Indeed, each super-peer has a bloom filter to maintain two types of information includes i) Community index aggregates the set of interesting keywords of all friends in its social community and ii) the response index comprises the indices of cached query results using hybrid caching. When super-peer receives a query from its friend, it will look up both the community index as well as response index to selectively forward the query to its connected friend who has the matched hash value of the required file name in the generated query. Thus, the query will be satisfied with in the same community with less response time. However, it is still possible that some friends in the community have matched file name. In the worst case, the third round search is invoked. The workflow of the search protocol is in Figure 5.

```
friend1 Generate Query (Keyword Iki: Filename F: TTL)
Process Query (Keyword Iki: Filename F: TTL) {
    Local_Search Query (Iki, F) {
        Search keyword in local index hash table of friend1;
        check if (LI contains IkI)
        return → F;
    }
    // community level
    else Community_Search Query (friend1,Iki, F) {
        forward query (Iki, F) → superpeer sp1 in its community;
        search F in bloom filter hash table of sp1 ();
        check if (sp1 (RI) contains F) { // In response index table
            sp1 retrun F → pi; }
            else if (sp1 (CI) contains F) { // In Community index table
                sp1 forward Q → corresponding friendi;
                friendi return F → sp1;
                sp1 forward F → friend1 ;}
    }
    // super-peer level
    else superpeer_search query (pi:peer, F: file_name) {
        flood query (friend1, f) → all superpeer (spj) in overlay;
        search F in bloom filter hash table of spj ();
        check if (spj (RI) contains F) {
            spj return F → spi;
            sp1 forward F → friendi; }
            else if (spj (CI) contains F) {
                spj forward → corresponding friendj
                friendj return F → spj;
                spj forward F → sp1
                sp1 forward F → friend1 ;}
    }
}
```

Figure 5: Cognitive Community Query Searching

Journal of Theoretical and Applied Information Technology
28th February 2014. Vol. 60 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

iii) Super peer level search

In the third round of search, the super peer floods the query among the neighboring super peers until the criterion TTL>0 get satisfied. That is the TTL in the query is decremented for each time the query is transmitted. The query will be dropped when TTL >0.

4. PERFORMANCE EVALUATION

In this section, a well-designed PaHC search mechanism is simulated using peer-sim and the performance of the proposed scheme is evaluated in terms both efficiency and user satisfaction. Efficiency of searching scheme is mainly determined from the better utilization of resources such as bandwidth and processing power and user satisfaction is evaluated from four performance metrics such as network traffic, success rate, response time and cache hit ratio [17]. It constructs network topology efficiently with 10,000 peers where 2000 social communities are available. Among that, each community is organized with a single high connectivity super-peer and associated with a set of client peers (friends).

4.1 Success Rate

The success rate directly relies on how efficiently processing the queries without fail. SR = $\sum S_p$ / Qp here S_p is the number of successful searches by peer p and Qp is the number of generated queries by peer p.

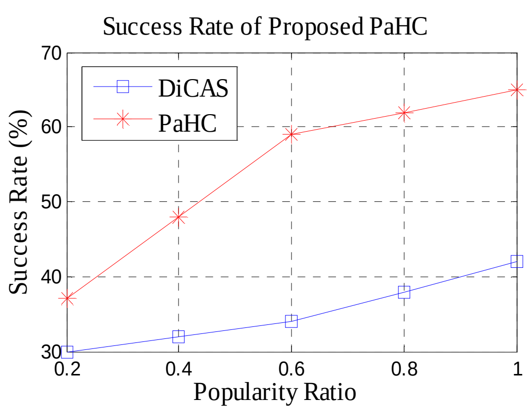

Figure 6: Success Rate of the proposed PaHC algorithm

Figure 6 illustrates that the PaHC approach achieves high success rate than existing approaches according to the increasing ratio of the popularity of the cached data item. If the popularity ratio of the data item increases, super-peer caches the most popular data items rather than unpopular data items. It leads to high data availability that improves the percentage of queries which can be responded by at least one query reply. The above analysis above shows that the PaHC improves the success rate at the price of high popularity computation cost. Thus, it achieves optimal tradeoff exists between the success rate and computation cost by effectively computing the popularity of the data item in the case of a large scale P2P network.

4.2 Average Network traffic

Average network traffic is the total amount of traffic generated divided by the total number of queries totally generated by peers in the network. ANF = $\sum T_p$ / Q_p, here T_p is the total amount of generated traffic and Q_p is the total number of generated queries.

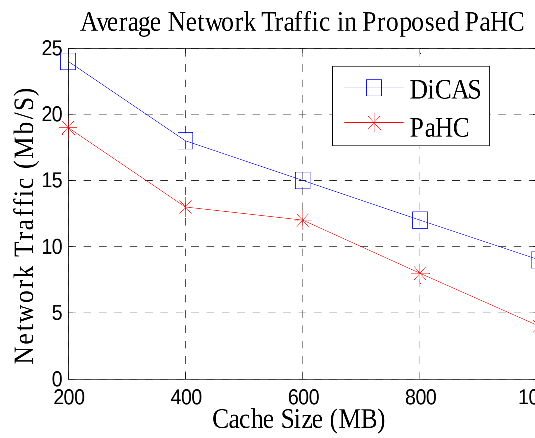

Figure 7: Average Network Traffic of proposed PaHC

In general, P2P searching incurs high network traffic overhead because of the propagation of a large number of data and control messages for query processing. Whereas in the proposed scheme, the network traffic converges due to the less searching scope and high cache hit ratio. Figure 7 depicts that proposed scheme incurs considerably 20% less network traffic than other existing searching schemes. At the beginning, PaHC incurs high traffic obviously, but further increasing the cache size can help to acquire less traffic. Thus, it achieves optimal tradeoff exists between the network traffic and maintenance cost by effectively maintaining in the case of a large scale P2P network.

4.3 Response Time

Response time clearly is the minimum number of hops of the underlying network required to forward the query reply back to the requested peer. The proposed scheme reduces query delay

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

significantly due to caching a large number of popular data items. It helps to reduce the average hop count. Therefore, requested peer gets query responses from the three hops of client peers within its social community. When the cache size is greater than 600MB, PaHC considerably incurs low response time with 15 percent lesser hops than existing searching schemes.

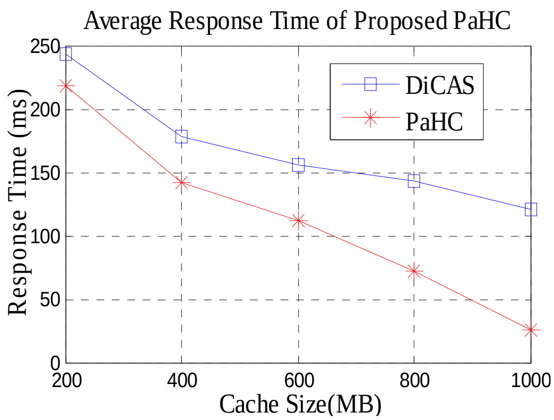

Figure 8: Average Response Time of proposed PaHC

Figure 8 clearly depicts that the proposed approach acquires much lesser hops for query processing than any other existing searching schemes. As the increasing size of the cached data item, gradually improves the response time by providing query responses with in lesser hops for the requested query. The enhanced PaHC approach can reduce the network traffic by an order of magnitude, decrease the response time by 25 percent than existing searching schemes.

4.4 Cache Hit Ratio

It is defined as the ratio of the number of generated queries get satisfied through cached information rather than query forwarding. CHR = $\sum C_p / Q_p$, here C_p is the number of successful searches routed through a cache and Q_p is the number of generated queries by peer p.

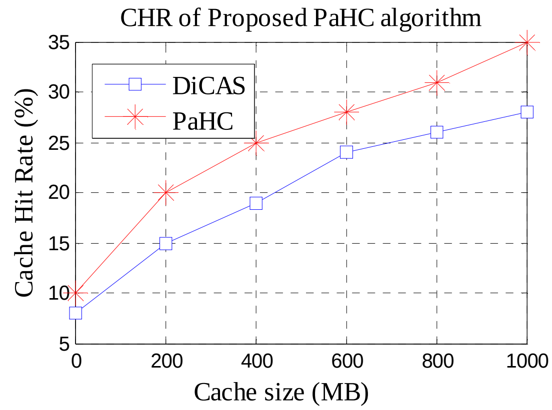

Figure 9: Cache Hit Rate Of The Proposed Pahc Algorithm

Figure 9 depicts that the PaHC approach achieves high cache hit ratio to significantly improve the searching efficiency. If the number of cached data item increases, most of the queries get satisfied within a cache that leads to acquire high success rate and low access latency. Indeed, in the proposed approach the more number of popular data items are cached in the super-peer that reduces the searching scope within lesser hops of minimum average delay. As the result of high searching efficiency, the proposed approach boosts the performance of P2P searching than the existing caching scheme.

5. CONCLUSION

This paper proposed PaHC, a novel peer-to-peer assisted searching framework by leveraging benefits from the online social network. The proposed approach addressed a series of key design to realize the robust query searching scheme in a large scale P2P network. It described how DOA algorithm can help to build a robust social overlay automatically on top of the unstructured network. This work is particularly interested in developing more efficient and scalable interest based Multi-keyword indexing to speed up the P2P query searching. In addition to, popularity aware hybrid caching strategy is proposed to boost the searching performance, and it also reduces the cache nodes overhead. It effectively relieves the high maintenance overhead due to popularity based hybrid caching at the same time delivers lower query delay, better load balance and higher cache hit ratios when many peers request popular data. Experimental results showed that the proposed algorithm provides high reliability, scalability and the simple implementation mechanism while comparing with the existing searching mechanism.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

REFERENCES:

[1] Choon Hoong Ding, Sarana Nutanong, and Rajkumar Buyya, "Peer-to-Peer Networks for Content Sharing", *Technical Report on Grid Computing and Distributed Systems Laboratory*, University of Melbourne, Australia, 2003.

[2] William Acosta Surendar Chandram, "Unstructured Peer-to-Peer Networks Next Generation of Performance and Reliability", *Proceedings of the 24th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '05)*, 2005.

[3] Arturo Crespo and Hector Garcia-Molina "Semantic Overlay Networks for P2P Systems", *Lecture Notes in Computer Science Agents and Peer-to-Peer Computing* , Volume 3601, pp:1-13, 2005.

[4] Gang Chen, Chor Ping Low, and Zhonghua Yang, "Enhancing Search Performance in Unstructured P2P Networks Based on Users' Common Interest", *IEEE Transactions on Parallel and Distributed Systems*, Vol 19, No 6 , pp 821 – 836, 2008.

[5] Sonja Buchegger and Anwitaman Datta, "A case for the P2P infrastructure for social networks – opportunities and challenges", *Proceedings of the Sixth International Conference on Wireless On-demand Network Systems and Services*, USA, 2009.

[6] Jing Xie, King-Shan Lui, "Modeling Random Walk Search Algorithms in Unstructured P2P Networks with Social Information", *IEEE International Conference on Communications, 2009.*

[7] Rim Haw Choong Seon Hong, "A Social P2P Networking Based on Interesting Keywords", *Proceedings of the International Conference of Information Networking*, pp 509 – 512, 2011.

[8] David Isaac Wolinsky, Pierre St. Juste, P. Oscar Boykin, Renato Figueiredo, "OverSoc: Social Profile Based Overlays", *Proceedings of the IEEE International Workshop on Enabling Technologies: Infrastructures for Collaborative Enterprises (WETICE)* pp: 205 – 210, 2012.

[9] P. Androutsos, D. Androutsos, and A. Venetsanopoulos, "Small World Distributed Access of Multimedia Data: an Indexing System that Mimics Social Acquaintance Networks", *IEEE Signal Processing Magazine*, Vol 23, No 2, pp 142– 153, 2006.

[10] Xu Cheng Jiangchuan Liu, "NetTube: Exploring Social Networks for Peer-to-Peer Short Video Sharing", *Proceedings of the IEEE International Conference on Information Communications(INFOCOM)*), pp 1152 – 1160, 2009.

[11] 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 social-based based peer to peer system," *Proceedings of the Fifth International Workshop on Peer to- Peer Systems (IPTPS),* 2006.

[12] 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.*

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

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

[15] 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.

[16] Liangzhong Yin, Guohong Cao, "Supporting Cooperative Caching in Ad Hoc Networks", *IEEE Transactions on Mobile Computing*, Vol 5, No 1, pp 77-89, 2006.

[17] Vicent Cholvi, Pascal Felber and Ernst Biersack, "Efficient search in unstructured peer-to-peer networks", *Proceedings of the sixteenth annual ACM symposium on Parallelism in algorithms and architectures*, 2004.

[18] Hsinping Wang and Tsungnan Lin, "On Efficiency in Searching Networks", *Proceedings of the 24th Annual Joint Conference of the IEEE Computer and Communications Societies*, 2005.