Available online at http://www.mecs-press.net/ijwmt

Efficient Techniques to Reduce Effects of Topology Mismatch and Heterogeneity in Unstructured P2P Networks

Dr. B Lalitha [a]

[a] *Assistant Professor in CSE, JNTUA College of Engineering, Anantapur, Andhra Pradesh, India*

Abstract

The formation of P2P logical networks oblivious to the structure of physical topology results in large amount of redundant network traffic. In addition to this mismatch problem, there exists a skew in properties of the participating peers which degrade the performance of P2P networks. So the current P2P systems call for effective overlay formation taking into consideration the underlying physical network topological properties and also inbuilt heterogeneity in participating peers. The heterogeneity of peers in the network can effectively used to bias neighbor selection and improve network performance by assigning more responsibility to nodes with higher capabilities. This paper presents two techniques to solve the problems of topology mismatch and heterogeneity. The proposed methods make use of bandwidth of peers and distance measures for overlay formation in the Gnutella network. The designed systems are tested with proper analysis and simulations to verify the correctness of the methods.

Index Terms: Peer-to-Peer, Gnutella, Topology Awareness, Bandwidth, Simulation, Network Traffic.

1. Introduction

Peer-to-Peer (P2P) networks are broadly acknowledged and are on a rise at rapid pace due the ability of the P2P networks to uphold the major distributed system networking problems such as scalability, heterogeneity and dynamic nature of the peers[13]. These P2P systems are operated not only for distributed high performance scientific computing but also for general purpose applications as they are becoming easiest and cheapest way to share resources and increase the computing capacity. It is evident from the studies [1] that P2P is highly popular, as 70% of the total Internet traffic is result of P2P systems.

In a P2P system each node acts as both as a client and a server as it gets services from other peers and also provide services to other participating peers. Overlay formation in the existing P2P overlays is independent of the underlying physical network i.e. internet. Peers communicate with other peers without considering their position or the relative distance from other peers in the network. This will result in a large communication

delay between the peers and poor utilization of the underlay network properties and resources.

Formation of overlay network in unstructured networks like Gnutella 0.4 [2] involves generation of lot of messages by the flooding based routing and even in the systems which consider the location for overlay formation generates significant amount messages to find the RTTs and delay between the peers. Although P2P networks provide many desirable properties for the users like anonymity, low maintenance costs and fault tolerance, they have some inherent problems in the current P2P networks:

a) Topology Mismatch Problem:

The inherent mismatch which exists between the P2P overlay network and the underlying Internet due to the formation of overlays without considering the underlay topological awareness results in redundant network traffic and low performance of the system.

Consider the P2P network depicted in fig 1 having 5 peers in the overlay network and 3 non-participating nodes along with overlay nodes in the underlay network.
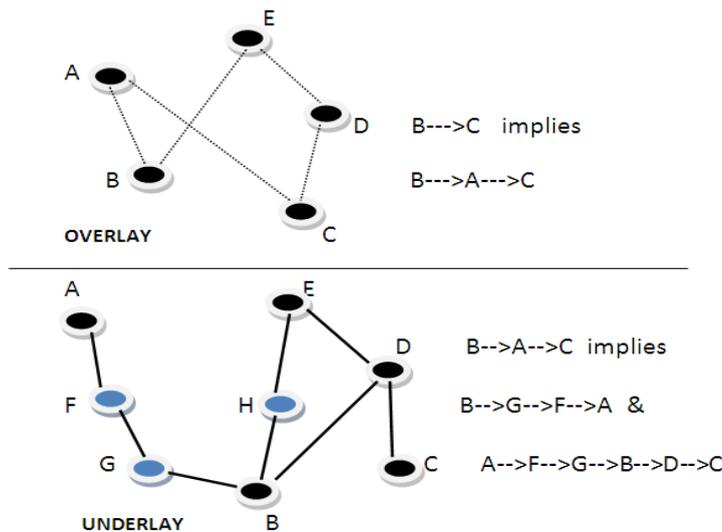

Fig.1. Topology Mismatch

As depicted in the figure each overlay network node has two other nodes as neighbors. The random selection of peers as neighbors in Gnutella without considering the underlay network properties lead to an overlay with low performance guarantee and redundant messages. In the network if a data item is to be sent from B to C then the overlay routing involves the path B→A→C. Since the actual routing occurs at the underlay level the actual data transmission path is B→G→F→A→F→G→B→D→C. As the resultant path shows that the routed path is not optimal for the data transmission. This is the topology mismatch problem where the overlay topology formation does not consider the underlay network properties like bandwidth, latency, location etc.

b) Inefficient Bandwidth Utilization:

The P2P applications like Gnutella uses the mechanism of flooding or pure broadcast for query search which consumes a lot of bandwidth due to the number of query messages generated in the network. Generally bandwidth is a property between two machines but practically the bandwidth between a peer and the rest of

internet is determined by the nodes direct link to the internet. As a result the bandwidth is a considered as a property of the peer. Studies [3][10] show that the average upstream bandwidth of nodes in a Gnutella network is around 1Mpbs, while 22% of peers have bandwidth below 0.1Mbps, 8% of peers have bandwidth above 10Mbps, and the highest observed bandwidth capacities are around 100Mbps. This inbuilt heterogeneity in the Gnutella network affects the performance of the system if all the nodes are treated equally.

The techniques presented in this paper try to solve the above two problems with focus on the bandwidth utilization as the bandwidth improvements in the system will have outstanding affects on the speed and performance of the system and also recent trends reveal that companies are more interested to improve the performance of the existing systems by monitoring and efficient utilization of bandwidth rather than increasing the amount of bandwidth due to high costs and resource consumption by the distributed devices.

The rest of the paper is structured as follows. Section 2 specifies the related work which led to problem definition. Section 3 introduces overlay formation technique using capacity of nodes and record route option of IP. The Section 4 illustrates an overlay formation technique using geographical latency measurements. Finally, we summarize in Section 5.

2. Related Work

2.1 Topology Mismatch Solutions

Numerous solutions [4] [5] [6] have been proposed to address the topology mismatch problem considering the properties of the underlying physical network like the location of the peer or delay and RTT between the peers.

One of the earlier solution is LTM [4] tries to overcome the mismatch by sending a detector message in the proximity of the node. The peers which receiving the detector will record the relative delay from the sending peer. The solution provided is simple but as the size of the network increases the more detector messages will be flooded in the network which affects the scalability and performance of the system.

To solve the mismatch problem SBO [5] uses the delay between the peers. But it is mainly used to reduce the optimization costs by distributing the optimization tasks to different peers. Peers in this system are characterized as red and white peers to distinguish the tasks performed by the peers. The demerits of this method are its complexity and message overhead involved for the overlay formation.

The solution in AOTO [6] use the formation of multicast spanning tree approaches which forms an overlay multicast tree from the source peer to all its neighbor nodes. Each peer maintains a cost table containing cost to reach all its neighbors. Peers exchange the tables periodically to update the costs. The demerit of this approach is the overhead involved in maintaining and exchanging the cost tables and also the formation of the overlay multicast tree while maintaining the overlay topology.

Another notable solution is Vivaldi [7] implements a distributed synthetic coordinate system and calculates the distance between two peers based on the coordinates assigned to the peer. But the overhead involved in implementing the coordinate system and calculating the relative distances which are not accurate compensates the merits provided by the solution.

2.2 Heterogeneity Solutions

The key method used to solve the problem of heterogeneity is the use of Super-peer topologies where a set of peers which are highly stable are provided additional responsibilities to improve the performance of the network. The super node approach was made popular by the introduction of KaZaA [9][12] which helped to build P2P systems that were scalable. Due to its success in the implementation this approach is introduced in the Gnutella systems.

The current stable version of Gnutella i.e. Gnutella 0.6 [8][12] uses the super peer topology where all the nodes are characterized as either ultra-peers or normal peers. The ultra-peer acts as server to the normal peers

and the set of ultra peers form an independent sub network within the overlay network and handle most of the system functionalities.

The use of super peer topologies introduces many problems which are needed to be taken care of. The problems are listed below:

- System must decide the total number of super peers needs to be present in the system.
- Super peer election algorithm deciding which peer to act as a super peer.
- Dynamically adjusting the super peer set as the peers join and leave network.
- Clients have to be distributed and migrated dynamically.
- All the operations must be performed in a distributed manner in the dynamic environment.

Most of these problems will be addressed if the overlay is developed keeping in mind the dynamic nature of the P2P system. But the real demerit of the super topologies is that if a super peer fails all the clients connecting to that super peer gets isolated from the P2P network. Even though heterogeneity in P2P systems is a challenge it is also an opportunity which can be exploited to our advantage by giving more preference to more stable nodes and improving scalability of the system and allowing dynamic movement in the system. The proposed techniques use this opportunity for formation of overlay network.

3. Bandwidth-Aided Overlay Formation

In general the capacity of a node is characterized by the number of queries that the node can handle per second. Normally for low-bandwidth nodes the query processing capacity is determined by the nodes available bandwidth. On the other hand for the nodes with higher speed connection, along with bandwidth other properties such as speed of the CPU, disk latency etc. will affect the node's capacity. The bandwidth-aided system implementation ignores the affects of CPU speed, disk latencies and the capacity is assumed as a function of available bandwidth. The overlay topology formation of the system is described as follows:

When a node enters the network it uses the bootstrapping mechanism similar to the one found in Gnutella to find the other nodes of the network. Each peer maintains a cache containing the list of all other peers present in the network. The cache entries will contain the IP address, port number and the capacity values of the other peers. These caches are exchanged periodically with the neighbors and updated accordingly[14]. The old and stale entries are marked as dead.

3.1. Heterogeneity

The main aim of the overlay topology formation is to ensure that high capacity nodes have higher degree and the lower capacity nodes are at short reach of higher capacity nodes. To achieve this purpose each node computes a *stability level*. This quantity is a function node's current set of neighbors, their capacities, degree and preconfigured maximum and minimum number of neighbors. This measure takes values between 0 and 1 that signifies how stable the node is with the current set of neighbors. As long as the node is not stable, it will continuously configure its neighbors and improves its level of stability. The stability value can be given as follows:

$$S(X) = \frac{\sum_{\forall i \in Nbr_x} \left(\frac{c_i}{Nbr_i}\right)}{c_x}$$

The measure S(X) returns a value between 0 and 1. If the value is 0, the node X is not stable and continues to search for neighbors with better capacity values. If the value is 1, the node X is stable and only tries to calculate the stability value periodically to keep the node stable.

3.2. Topology Mismatch

In order to solve the topology mismatch problem, this technique uses the *record route* option of IP protocol. The record route is the option 7 of the internet protocol which makes the nodes traversed during communication to add their IP addresses to record route data. The *pointer* field in record route will point to the position to insert the next IP address. The *length* field defines the number of IP addresses to be inserted in the record route data field. If the route data area is full i.e. pointer exceeds the length, an ICMP parameter problem message will be sent to the sender host. This *Length* field is used to restrict the scope of choice of neighbors for a node. The format of record route option is as follows:

OPTION TYPE (8 bits)	LENGTH (8 bits)	POINTER (8bits)	ROUTE DATA (variable)

Fig.2. Record Route Format

3.3. Topology Formation

When a node starts up it has minimum preconfigured set of neighbors and it is assigned the maximum number of neighbors it can have based on its bandwidth. Now the node starts gathering more neighbors in this scope and improves its level of stability. The node will continuously reconfigure its neighbors until the node will have optimal set of neighbors. This process is described in algorithm 1. To add a new neighbor, a node randomly selects a small set of entries from its own cache that are not already neighbors or dead entries. From these entries the node selects those nodes with greater capacity than itself and sends the connection request message. If no such eligible entry exists the node randomly selects one entry from the cache and sends the connection message.

> Let $Ci \leftarrow$ Capacity of Node i
> $NbrA \leftarrow$ Number of Neighbors of A
> $Max\_NbrA \leftarrow$ Maximum neighbors Node A can have
> $Min\_NbrA \quad \leftarrow$ Minimum Neighbors Node A should have
> $S \leftarrow$ Stability Value
>   If $Nbr_A < Min\_Nbr_A$ then
>   Return 0
>   Value $\leftarrow$ 0
> $\forall X \varepsilon \ Nbr_X \ do$
>   Value $\leftarrow$ Value $+ \frac{x}{Nbr_x}$
> $S \leftarrow \frac{value}{c_x}$
>   If $S>1$ or $Nbr_A > Max\_Nbr_A$ then
> $S \leftarrow 1$
>   Return S

Algorithm 1: Stability level

During the connection establishment phase each node makes decision of whether to accept the other node as neighbor or not based on its capacity with the existing neighbors and with the new one. To accept the new node,

an existing neighbor will be dropped based on the capacity value. The neighbor acceptance is based on the maximum number neighbors that the node can accept. While dropping the neighbors, the degree of the dropping node is also taken into consideration in order to avoid dropping connections of already poorly connected nodes in favor of highly connected ones.

3.4. Search Protocol

Unlike Gnutella, this technique uses the biased random walk for searching the data items in the P2P network. Biased random walk is the method of forwarding the incoming query to highest capacity neighbor instead of forwarding to a random neighbor of all the neighbors. In order to limit the random walk at some point of time, the TTL values are used their by limiting the number of hops a query can progress. To avoid redundant paths the nodes store the QID of the query and the neighbors to which it has already forwarded the query, which reduces the chance that a query is forwarded twice in the same path. The search protocol is depicted in the fig 3. It shows how the query is forwarded using the biased random walk where the bias is towards the high capacity nodes, which are represented using larger circles.
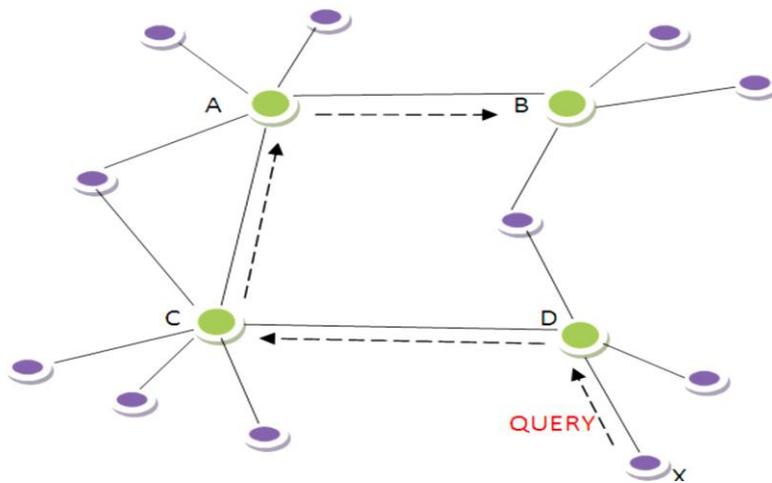

Fig.3. Search Query

3.5. Simulation

The simulation setup involves the following parameters. First each node in the system is assigned a capacity value which represents the number of messages it can process per unit time [11]. The capacities are assigned based on Gnutella distributions derived from [3]. In order to evaluate the effectiveness of the proposed system, we compare the simulate results with that of Gnutella 0.4 and Gnutella 0.6. Each node in the Gnutella 0.4 system is modeled with 4 neighbor nodes. In the Gnutella 0.6 each ordinary peer has 3 ultra-peer connections and each ultra-peer has 10 ultra-peers as neighbors. The node in the proposed system is configured with minimum of 3 neighbors and the maximum is based on the capacity of nodes. On an average the nodes in the bandwidth aided system have 20 neighbors.

To evaluate the effectiveness of proposed overlay formation technique we consider the response time of the queries which were successful. This is depicted in Fig 4.
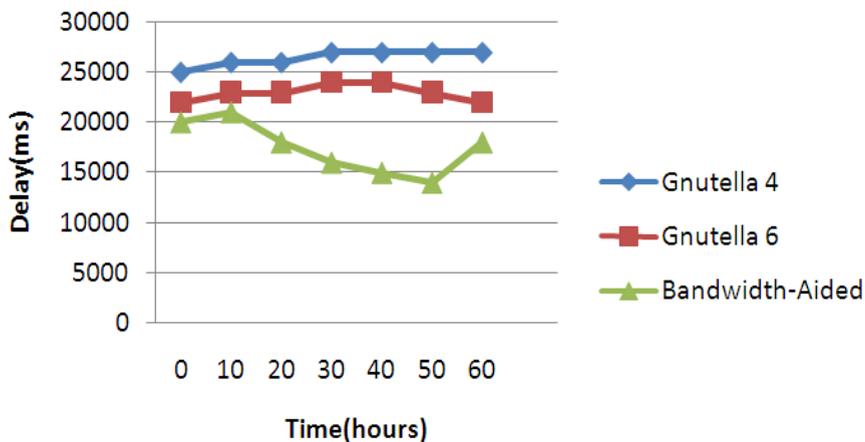
Fig.4. Query Delay

To better understand the distribution of load among the nodes, we consider the load at the nodes for each of capacity values. This is depicted in Fig 5.
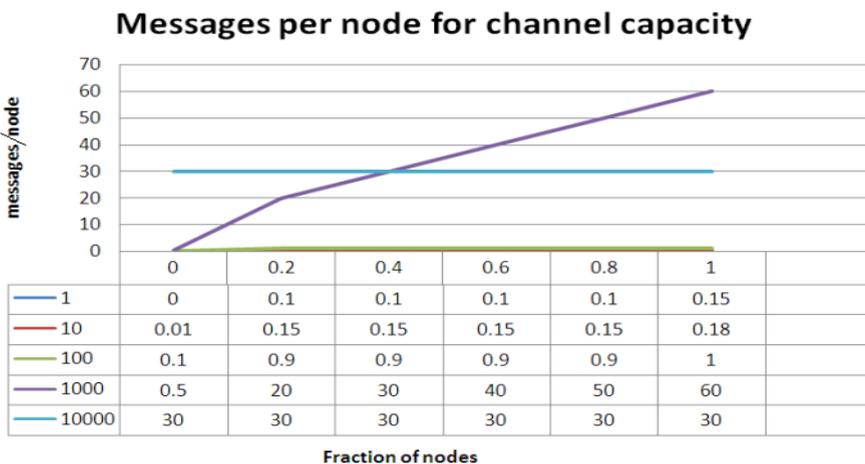

Fig.5. Message Overhead

The Bandwidth-Aided topology awareness system is a highly scalable unstructured P2P system which improves the performance of the Gnutella network by considering the peer bandwidth for the topology formation and distributing the load to the high capacity nodes. This technique makes the search more scalable and reduces the query response time which resulted in drastic improvement in the query success rate with small overhead.

4. Geographical Latency Based P2P overlay

This technique involves the implementation of a P2P overlay reducing the overhead caused due to topology

mismatch and heterogeneity. The overhead introduced by the topology mismatch is optimized by limiting the scope of neighbor selection using the geographical latency measures. The problem of heterogeneity and message overhead is overcome by building a scalable and robust overlay considering the capacity of the participating peers of the P2P system.

The structure of the overlay has to be adapted according to the capabilities of the participating peers. Otherwise the lowest performance peers with low processing power and throughput will become bottleneck to the system. The proposed system involves the use of a modified super peer topology and uses a *value* metric for the super peer election. The issues considered for the construction of this modified super peer network include

- Super peer election algorithm
- The number of super peers.
- The number of clients for each super peer.
- Issues related to the dynamic nature of the P2P network.

4.1. Super Peer Election

The super peer election algorithm constitutes the major function to find which peer in the network has eligibility to act as the super peer in the network. Gnutella 0.6 specifies that a peer can become a super peer if the peer has non- firewalled link with 20 KB/s download bandwidth and 10 KB/s upload bandwidth running a suitable operating system.

The super peer election in the proposed technique depends on the *value* metric. The *value* metric is a function calculated at the peer locally and gives the peer's ability to give resources to the P2P system and provide services to other participating peers. This metric is a domain specific and imbibe the requirements of the application built above the overlay network. For example a file sharing application can set the value metric based on the function of peers upload bandwidth. All the peers with a certain super peer threshold will act as the super peers and the remaining peers will become the clients.

4.2. Number of Super Peers and Clients

The number of super peers is very critical for the performance of the network. For example a file sharing system must have the number of super peers smaller enough to reduce the search traffic and large enough to handle the load of the clients. In the proposed technique the number of super peers is directly controlled by the super peer threshold. Larger the threshold smaller the number of super peers and vice versa. Since the number of super peers is considered by the threshold there will be no reconfiguration in the connections of the network.

4.3. Dynamic Nature of the Network

Due to the dynamic nature of the P2P networks if a super peer leaves the network all the clients which are attached to the super peer gets isolated from the network[15]. This is a demerit in the Gnutella 0.6 specification. In order support this dynamic behavior of the network the clients in the proposed system are allowed to connect to more than one super peer, thereby allowing the clients to be in the network even if one of the super peers leaves the network.

The modified super peer topology constitute the peers which have high *value* are logically present as the super peers of the topology while peers with lower *value* will be present as clients . In this topology *value* thresholds are defined to distinguish low *value* peers from the higher ones. Peers with similar thresholds are present close to each with small hop count in the overlay. A general overlay topology is depicted in the Fig 6.
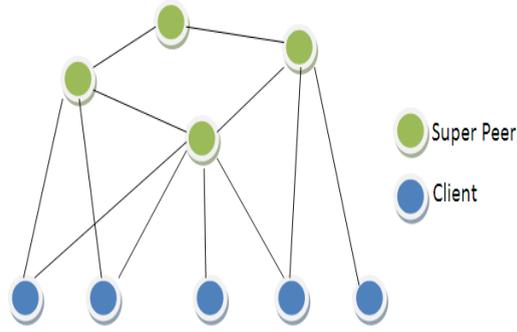
Fig.6. Overlay Topology

4.4. Topology Mismatch

In order to solve the topology mismatch we first estimate the geographical latency between the peers using the following formula

$$RTT\ (a,\ b) = 62ms + 0.02ms/km * dist\ (a,\ b)$$

Where *dist* (*a*, *b*) is the geographical distance between two Hosts in *kilometers*, calculated using the *haversine formula*. The formula calculates distances on a sphere, by using the latitude and longitude of a node.

Based on these geographical latency values we introduce various constraints on the overlay formation methodology. The constraints are:

- The distance between client and SP must not be greater than λ.
- The distance between a SP and another SP must not be greater than λ+δ.

Here λ and δ are system configuration parameters which can be assigned according to the requirement of application. The first constraint will make the clients to exist at shorter distances to the SP's, while the second constraint will make the SP's to exist at a considerable distances.

4.5. Simulation

In the system each node is assigned a *value* metric based on the requirement of the application. The value we have considered is the capacity value in the range 1:20. Then the initial number of super peers is set along with the threshold values which again hold the application requirements. Each super peer in the network can have up to 32 clients and connections to other 30 other super peers. A threshold is set which states that a peer with more capacity value than 15 can become a super peer. To evaluate the performance of the system results of the proposed system are compared with the stable versions of Gnutella i.e. Gnutella 0.4 and 0.6.

The Fig 6 depicts the query delay in the three networks. Since Gnutella 0.4 uses the flooding approach it takes more time to get the query results compared Gnutella 0.6. The proposed technique uses the SP network for query traffic and replays the qeries faster than the others.
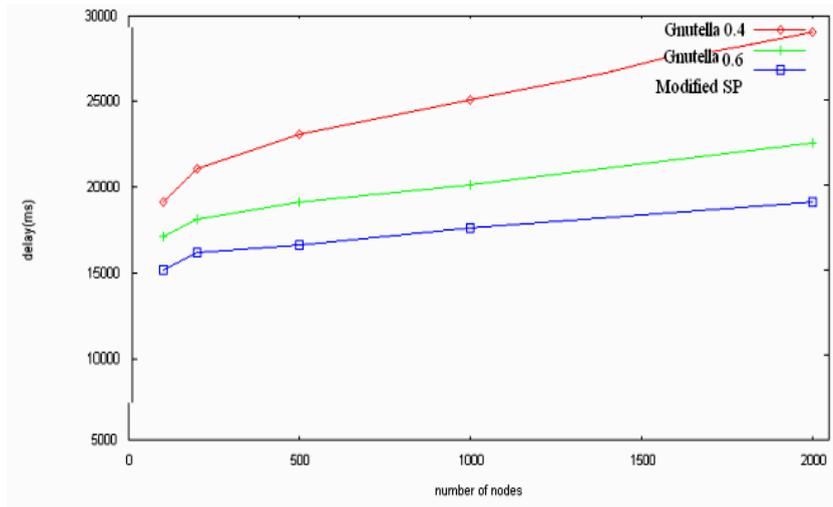
Fig.7. Query Delay

The Fig 8 describes the average traffic cost generated by the nodes. It shows that the average number of messages generated by the Gnutella networks gradually increases with the size of the network compared to the modified super peer network.
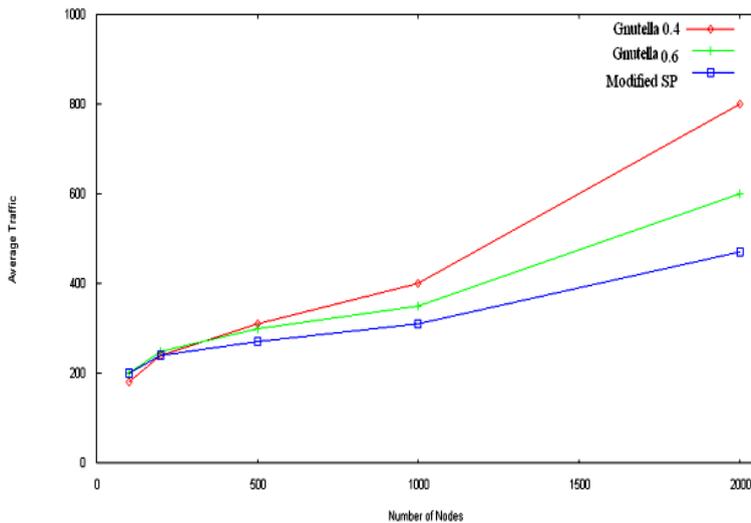

Fig.8. Average traffic

The proposed system implements a robust and scalable P2P overlay with improvements in the performance of the system and reducing the network overhead by using geographical latency measures. The presented system reduces the network overhead caused by the topology mismatch problem and message overhead caused by the heterogeneity using the modified super peer network. The system provides faster query responses with low message overhead and reasonable success rate.

5. Conclusion

Both the presented techniques try to reduce the affects of the problems of topology mismatch and heterogeneity. The bandwidth-aided system solves the problem of mismatch by using the Record route option of the internet protocol and solves the problem of heterogeneity by distributing the load to high bandwidth nodes. The geographical latency based system will use the geographical latencies for solving topology mismatch and by segregating nodes into super peers and normal peers based on value metric will solve the problem of heterogeneity. Both the systems simulation results suggest that they have faster query response time with decent query success rate.

References

[1]   H. Schulze and K. Mochalski. Data about P2P, VoIP, Skype, file hosters like RapidShare and streaming services like YouTube. Ipoque, Internet Study 2007, November 2007.

[2]   The Gnutella Protocol Specification 0.4, http://rfc-gnutella.sourceforge.net/developer/stable/index.html.

[3]   S. Saroiu, P. K. Gummadi, and S. D. Gribble. "Measuring and analyzing the characteristics of napster and gnutella hosts.*Multimedia Systems*, 9(1):170-184, July 2003.

[4]   Li Xiao, Yunhao Liu, Xiaomei Liu, Lionel M. Ni & Xiaodong Zhang.(2005).Location Awareness in Unstructured Peer-to-Peer Systems, IEEE Transactions on Parallel and Distributed Systems, No. 2, Vol. 16.

[5]   Yunhao Liu, Li Xiao, and Lionel M. Ni, Building a Scalable Bipartite P2P Overlay Network, IEEE Transactions on Parallel and Distributed Systems, Vol.18, No. 9, September 2007, pp. 1296-1306.

[6]   Y. Liu, Z. Zhuang, L. Xiao, and L.M. Ni, AOTO: Adaptive Overlay Topology Optimization in Unstructured P2P Systems, Proceedings of IEEE GLOBECOM, 2003.

[7]   Dabek, F., Cox, R., Kaashoek, F. & Morris, R. (2004). Vivaldi: A Decentralized Network Coordinate System, Proc ACM SIGCOMM, 2004.

[8]   Tor Klingberg and Raphael Manfredi. Gnutella 0.6, 2002. http://rfc-gnutella.sourceforge.net/src/rfc-0_6-draft.html.

[9]   KaZaA, 2002. http://www.kazaa.com/.

[10]   R.Buyya (editor),"High Performance Cluster Computing", Prentice Hall, USA, 2000.

[11]   I. Foster, C. Kesselman (editors),"The Grid: Blueprint for a Future Computing Infrastructure", Morgan Kaufmann Publishers, USA, 1999.

[12]   R. Subramanian, B. Goodman (editors)", Peer-to-Peer Computing: Evolution of a Disruptive Technology", Idea Group Inc., Hershey, PA, USA, 2005.

[13]   G.Couloris, J. Dollimore, T. Kingberg,"Distributed Systems – Concepts and Design", 4th Edition, Addison-Wesley, Pearson Education, UK, 2001.

[14]   David R.Karger, Matthias Ruhl,"Simple efficient load balancing algorithms for peer-to-peer systems", In SPAA '04: Proceedings of the sixteenth annual ACM symposium on Parallelism in algorithms and architectures, ACM Press, 2004.

[15]   Yingwu Zhu, Yiming Hu, "Efficient, proximity-aware load balancing for structured p2p systems", In P2P '03: Proceedings of the 3rd International Conference on Peerto-Peer Computing, IEEE Computer Society, 2003.

Authors' Profiles

Dr. B. Lalitha received her B.Tech Degree from SITAMS, Chitoor in the year 2003. M.Tech degree from JNTU Anantapur in the year 2005 and received her Ph.D in CSE in the year 2015 from JNTU Anantapur. She is currently working as a Head of department and Assistant Professor in CSE department, JNTUA college of Engineering, Kalikiri. Her research interests are Distributed Computing, Peer to peer networks, Cloud computing and Service oriented architecture.