ABSTRACT

P2P systems are a form of distributed networks constituting autonomous nodes organized by cooperation of all participating nodes in the process of improving the performance of a particular task. These networks are comprehended as logical application level overlays formed by equal peers over the underlying physical network of Internet by using resources voluntarily shared by the node. The resources may include network bandwidth, processing power or disk storage. Any node in the internet can participate in a P2P network by installing appropriate P2P software. Key applications incorporating the P2P networks include file sharing systems like Gnutella and BitTorrent, distributed computing systems like SETI@home, streaming systems like P2PTV and multicasting systems like Peer Cast.

P2P overlays are either structured or unstructured based on the organization of the overlays.Structured overlays use Distributed Hash Tables (DHT’s) to store the details of the data items stored by nodes and it uses identifier based search schema to find the data items or objects. Maintaining these DHT tables involves lot of overhead on the network. In unstructured networks, node does not have the knowledge of the data items and hence it uses a search based mechanism.

Gnutella is an unstructured P2P network which uses flooding based object search mechanism to find the data items. The arbitrary joining and leaving of peers and also the formation of overlays by randomly choosing the logical neighbors without the knowledge of underlying physical topology leads
to topology mismatch problem, which attributes to considerable amount of redundant traffic in the network there by limiting the performance gains from various search and routing mechanisms. The topology mismatch problem influences the performance of Gnutella network in its operations such as searching, forwarding and replaying the queries. Studies revealed that more than half of the existing P2P systems are affected by the topology mismatch, which hinders the growth of the P2P networks.

The research is focused on solving mismatch problem by taking into account the underlay network properties such as latency, bandwidth and location of nodes to improve the performance of the Gnutella network in terms of message overhead and query response time without sacrificing the convergent speed of the network. The overlay formation techniques include the use of virtual coordinate systems, bandwidth of the peer, RTT (Round Trip Time) between nodes and structure of the IP addresses. All these techniques will help in finding the distance between nodes in underlay and this information is useful in forming the overlays by optimizing the logical neighbors of the peers.

The number of messages, query delay, average hops length and query success rate are some of the metrics considered to estimate the effectiveness of the proposed techniques. The proposed methods are simulated using a general purpose java based simulator considering various scenarios and complexities. The simulation results support performance enhancement of the P2P systems.