

## ABSTRACT

The bulk of Internet traffic is generated by content distribution, particularly by P2P content sharing applications in recent years. One of the advantages of P2P content distribution is that the overlay networks can be extended without increase in the bandwidth capacity of server. In other words, the server workload can be significantly reduced due to balanced network flows. P2P large scale multimedia content distribution systems, allow peers to work together so that a large scale multimedia content can be distributed to a large group of users, without the support of dedicated servers. It implements a simple design philosophy such as the large scale content is divided into number of blocks, and peers connect with one another in a random mesh topology and exchange these blocks with random gossiping. In gossiping, each peer transmits a subset of the blocks it has received to a subset of its neighbours that are selected by using randomized algorithms. Such random gossiping on random mesh topologies is easy to implement. It is resilient to the level of unpredictability caused by peer churn, and has been shown to attain good performance from a theoretical viewpoint. P2P content distribution systems, including Gnutella, KaZaA, Napster, and BitTorrent, have widely been adopted to quickly distribute content on the Internet. P2P model has been installed over the Internet recently to offer on-demand or live media streaming services.

Network coding has been used as a promising technique to improve the performance of P2P content distribution in recent years. It not only minimizes the average download time, but also enhances the performance of network even in presence of peer churn which has been shown as the benefits introduced by network coding. Network coding also makes content distribution more effective and easier in P2P content distribution network and reduces the burden of the original seeder. The randomization introduced by network coding

makes each packet equally important and it resolves the problem of locating the rarest block. Further, it reduces network traffic in the network.

Network coding has received a substantial amount of research attention in the networking community, since its inception in information theory. It has been proved that, with network coding, the cut-set bound of information flow rates in both P2P and multicast communication can be achieved. The core idea of network coding is a paradigm shift to allow coding at intermediate network nodes between the source and destination in a communication session. With the ability to code at intermediate nodes, there is a possibility to store, code and forward incoming packets in the communication session. This capability is in contrast to traditional commodity flows, where intermediate node only forwards the incoming flow.

The main objective of this thesis is to maximize the innovative information flow in the network code based P2P content distribution network at the same time minimize the download time and coding complexity. This thesis proposes three different novel mechanisms to achieve the above stated objectives.

First a new framework, called Intelligent-peer selection and Incremental- Network Coding (I2NC) is proposed to eliminate the unnecessary flooding of non innovative coded packets and thereby improve the performance of network coding in P2P content distribution further. This proposed system consists of seven different modules. They are namely Bandwidth manager, Buffer map exchanger, Intelligent neighbourhood selector, Incremental encoding module, Concurrent data request sender, Innovative information checker and Decoding module. In this framework, the intelligent neighbourhood selector and incremental coding module play an important role to reduce the network traffic and concurrent data request sender aims to maximize the throughput. This system presents a realistic implementation of network code based content

distribution in P2P network. A comparative study and analysis of the proposed system is made through various related implementations. Experimental results show that 10–15% of traffic is reduced and it improves the average and maximum download time by reducing original seeder's workload.

Random Linear Network Coding (RLNC) has proved to be a promising technique to cope with network topology changes since it does not require entire network topological information. Hence it can be used in P2P network because of its dynamic nature and difficulty to obtain entire topological information. In these RLNC mechanisms, the coding resource minimization is not a major concern and each node has to exchange its coding weights throughout the entire network. This exchange leads to under utilization of network capacity and so it is necessary to investigate how to enhance the performance of network coding based on the entire network topology and at the same time the advantages of RLNC need to be acknowledged. The problem of finding the minimal set of node to perform the coding operation and links to carry the coded data is called a Network Code Minimization Problem (NCMP). It is proved to be NP hard problem. It can be solved using Genetic Algorithm (GA) because GA can be used to solve the diverse NP hard problem. Second in this thesis, a GA based solution for solving a proposed new NCMP model is proposed which considers both minimization of network coding resources and dynamic changes in network topology due to disconnection. The simulation results prove that the proposed system outperforms the random selection and coding at all possible node mechanisms in terms of both download time and system throughput. It proves that the proposed system's throughput and download time is improved by 15% to 20% when compared to existing system.

Third, the dynamic network coding model proposed in this thesis considers the robustness to some extent when compared to previous works. Most of the existing P2P content distribution systems implement a random or rarest

piece first transmission procedure to avoid duplicate transmission of same pieces of data and rare pieces of data occurring in the network. Network coding can be used to solve this problem in a simple manner. The proposed mechanism codes content into several messages and split interested peers into several groups with each group responsible for spreading one of the messages. The coding scheme is designed to assure the property that any subset of the messages can be utilized to decode the original content as long as the size of the subset is suitably large. To meet this condition, a dynamic smart network coding scheme is defined which assures preferred property. Next, connect peers in the same group to send the corresponding message, as well as connect peers in different groups to disseminate messages for carrying out decoding operation. Moreover, the proposed system can be readily expanded to support topology change to get better system performance in terms of reliability, throughput and link stress. The simulation results prove that the proposed system can attain 20%–25% of higher throughput than existing systems. It also further achieves good reliability and robustness to peer churn and link failure.

This dissertation is aimed to propose an efficient mechanism for a large scale P2P multimedia content distribution by using network coding. The performances of the proposed systems are evaluated and analyzed based on the simulations that are carried out in NS2 and Peersim simulators and are compared with traditional systems.