|| Volume 1 ||Issue 2 ||May 2016||ISSN (Online) 2456-0774
INTERNATIONAL JOURNAL OF ADVANCE SCIENTIFIC RESEARCH AND ENGINEERING TRENDS

APPLICATION OF NETWORK CODING FOR PEER TO PEER FILE SHARING

Ashwini Chaudhari

Student, Computer Science & Engineering, Everest College of Engineering & Technology, Aurangabad, India

***Abstract*— In the last several years, the internet has increased large number of web application which contain web based application video broadcasting and conferencing. Web based application is more popular because of its flexibility and user-friendliness. Many such web-applications contain one source (server) and number of destinations (receivers).in file sharing application web or file server hold the file which contain number of client. In peer to peer system, peer generally acts as end or last host. It is not possible for file sharing system to be reliable and flexible at the same time for gaining good throughput. Network coding aspires to improve parameter such as throughput and reliability. By using network coding the any node between source and destination are used to transfer the message. Network coding is scheme in which node can generate output message by encoding the message that received from the same node.**

Keywords:- Component, formatting, style, styling, insert.

I INTRODUCTION

Now days, the node between source and destination are used to store and transferred messages. Recently to improve the network throughput and reliability the development of network coding is carried. Network coding in which node is allow to build input message by encoding. Thus in contrast traditional routing approach intermix the information and each node forward and received messages

In the previous years, the Internet has witnessed great rising of variety of web-based applications which contain web-based content sharing, video broadcasting or conferencing. Web-based applications are of more interests because of flexibility and user-friendliness. Many such applications contain one source (server) and multiple destinations (receivers). However, because of lack of multicast support over the internet these applications usually shows scalability problem, which restrict the number of receivers contains. Peer-to-peer is a excellent technology that can use application layer to be multicast, and where receivers (peers) not getting only data, but also peer forward that data.

In web based application there is problem of scalability, which can removed peer to peer technology. And also the system performance like (throughput, latency, etc.) is also improved by this technology. For distributed application Peer-to-peer (P2P) systems provides good infrastructure which achieve good performance.

In this paper by applying peer-to-peer technology to the file sharing, in which server like web and file which holds a file that is requested many clients (receivers). In peer-to-peer network, peers may act as end users. In private computers

Resources like bandwidth, CPU time etc are limited or even be disturbed. It is critical for the file sharing system to be reliable and durable while achieving good throughput at the same time. Multicast Application layer can be incorporated in peer-to-peer technology. Using peer-to-peer networking. By using multicast network coding algorithm which can fully exploit the network capacity. Network coding refers to a scheme where a node is allowed to generate output messages by encoding (i.e., computing certain functions of) it's received messages. Thus, Intermixing of information allow in network coding where contrast to the conventional routing approach where each node transferred received messages. Using butterfly network in Fig. 1 to demonstrate the benefit of network coding for multicast.

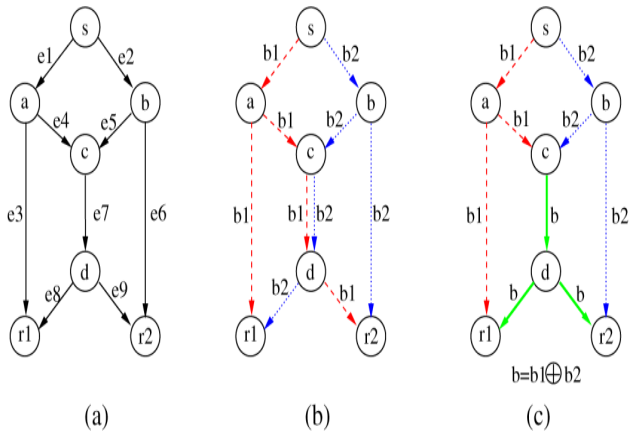

Figure 1 Benefit of network coding in multicast. (a) The butterfly network with DAG. (b) Multicast with lacking network coding. (c) Multicast with network coding [2].

In the above figure, s is the node source and r1 and r2 is nodes of receivers, all edges in network having power 1, which means edge can only transmit 1 unit of data (bit) per unit time (second), and b1 and b2 are two source of s to multicast to both r1and r2.Firstly we use conventional multicast which is not having network coding as shown in Fig. 1(b). Lacking of loss of majority, we used dashed line with red color bit b1, and another dotted line with blue color stand for bit b2 and the bold line with green color to stand for both bits b1 and b2. R1 reaches to Bit b1 in seconds, and bit b2 can reach r2 in seconds. Both b1, b2 bits is coming towards node c, it move forwards them in

sequence. Suppose it forwards bit b1 first. Then r1 and r2 get both bits in 4 seconds and 5 seconds sequentially. Now consider Fig. 1(c). When node c receives r1 and r2, it first mixes then using exclusive OR (XOR) operation. Then it sends the mixed bit b to d node. When nodes r1 or r2 receive the mixed bit, it can recover the exclusive bits b1 and b2 by XO-Ring the mixed bit and the other received bit. This process is done in 4 second. The definition of link c to d throughput (γ) from node c to d is given by

γ = number of node transferred from node c to d inspection duration / Observation duration

Some insufficiency in conventional network that removed in Peer-to-peer (overlay) networks as follows: 1. The topology used in peer to peer network is not permanent 2. Every node is the end host in Peer-to-peer network, so it is easier by using network coding we can easily used the complex operation such as encoding and decoding than storing and forwarding the message.

There are three types in network coding schema useful for routing i.e. uncast, multicast and broadcast, which allow a more proficient data transmission.

II LITERATURE REVIEW

Many network coding ideas given in earlier period on multicast networking as given below:

- 2003: Important steps taken toward realistic implementation. Li, Yeung and CAI gives the mathematical model in which mathematical operation such as addition multiplication is given which overcome the problem of complexity.
- 2005-2006: Important design algorithms published. Sidharth Jaggi, then at Caltech, with Peter Sanders of the University of Karlsruhe in Germany, published less-complexity algorithms for manipulative the functions used by each node in a multicast network. The first paper they give stracturised approach for implementing design, the second showed that choosing functions subjectively and separately for each node should work just as well[6][7].
- 2005:One of the major issue in the network cost for the connection which solve by linear program in polynomial time in decentralized way[9].
- 2006: Uses for wireless networks explored. D.M.Chiu, R.W.Yeung, J.Huang, and B.Fan, gibes the wireless network whose approach is prominently useful[10].
- 2014: In peer to peer file sharing liner network coding is develop in which file sharing and good performance is carried out.; Peer-to-Peer File Sharing Based on Network Coding (PPFEED) utilizes grouping networks as its cover topology prototype. But in further linear network shows less performance than grouping network. As a result, PPFEED inherits its great performance when applied network coding and presents its authority compared to other existing peer-to-peer file sharing systems. Besides, it achieves higher consistency and resiliency [1].

III PROPOSED SYSTEM

To overcome the drawbacks of Existing system as written above we will go through proposed system, In recent survey, network coding has emerged as promising information theoretically move toward to improve the performance of peer-to-peer (P2P) networks [4]. It has been widely received and accredited that network coding can theoretically progress the network throughput of multicast network.

The proposed system is explained as following:

1. Architecture defines two modules, one at server side, and another at peer side (Client side). Unlike other P2P distribution networks the server is not source of files in this system.
2. Peers are source of files, which can share files with any peer of another peer
3. Background progress running at peer module participates in P2P sharing activity, which helps other to encoding messages
4. Peer will contact to server for file availability
5. The network cover in a directed graph, in which nodes are peers in the network, and edge, is bandwidth between them.

Peer join/leave is decide by server.

6. Server has database of files, each file is recognized by its checksum. File checksum is the result of a hash function over the file content. This will avoid redundant entry of same file into the database.
7. The server decides routing of message. It will detect sub graphs which resembles with butterfly network [2] into network overlay graph, and inform intermediate nodes to participate. Hence server assists the peers to download the file.

The main aim of this work is to reduce the total number of transmission in multicast network and also to reduce the bandwidth consumption in multicast network using Bit Torrent file sharing protocol, by implementing network coding algorithm. Here the abstract diagram of file sharing between peer-to-peer using network coding algorithms as shown in figure below.

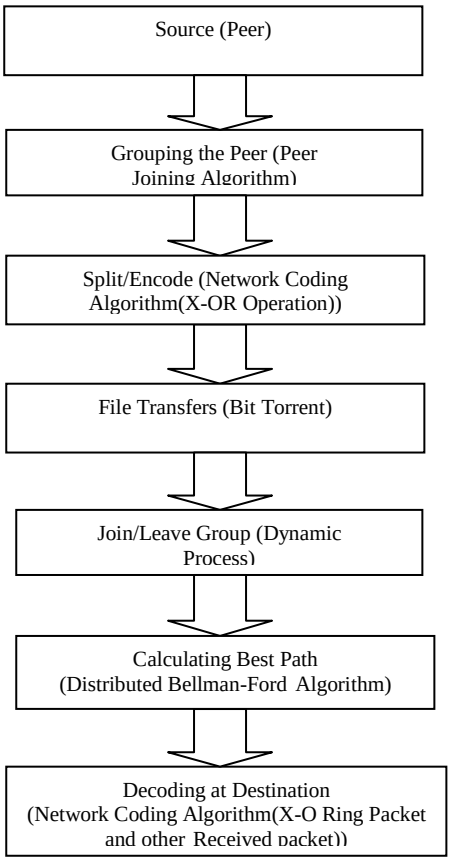

Figure 2 Abstract Diagram of File Sharing Between Peer to Peer Using Network Coding Algorithm

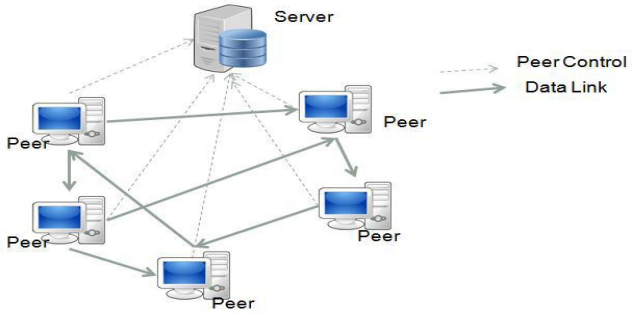

Figure 3 Peer To Peer File Sharing Network

3.1 Peer Joining:

Consider the server whose IP is known to all the peer with some address translation service called DSN (Domain Name System).When the peer want to receive hosted file then it send Join request to server, in this process server maintain the counter of the peer.for each group Gi server maintain record Gci to store peer group and gli store the cluster which contain group of peer. Besides, the server maintains a list of existing peers and their individual residue upload bandwidths and IP addresses for each and every group. As more and more peers join the system, it is source consuming to maintain a full list of peers for each group. The server keep record of limited list of peers with largest residue uploaded. Meanwhile, peers send to the server their updated residue upload bandwidths and it timely to update the limited list on the server.

Table-1: Peer Joining Algorithm

INPUT : joining peer v
OUTPUT : updated overlay network

BEGIN
//suppose the cluster analogous to peer v is Ci
if |cli| <k
Si=the set of groups not in cli;
else
Si=the set of groups in cli;
choose a group g_i S such that gl_i is the negligible;
if multiple groups have the same smallest gl, choose a group gi
with less gci;
Peer v is assigned to group gi.
END

After receiving the list of peers, the new peer will connect them and create cover links with them. These peers are called intra-neighbours of the new peer because they are within the same group. In contrast, the neighbours with different groups are called inter-neighbours. The new peer asks one of its intra-neighbours to provide a list of it's inter-neighbours. When preference the intra-neighbour, greater priority is given to the peer in the same cluster. The new peer then takes the list of peers as its inter-neighbours. The topology of the peer-to-peer network can be considered as a combination of number of unstructured peer-to-peer networks, each of which is composed of the peers within the same group. The topology within one group is randomly as long as it is connected. The only constraint is on the edges between different groups. It is required that each peer is connected to at least k-1 peers in k-1 different groups respectively. Here k is multicast capacity of network. When

More than k-1 peers are connected then the consistency of system in increased.

The pseudo-code for peer joining is listed as Table 1 shows.

Following are two types of peer leaving

1. Friendly

2. Abruptly (terminating or changing suddenly)

In friendly process, peer is leaving by sending leave request To intra- neighbours and inter-neighbours, so that as and when system required it make changes accordingly.

In opposite to the abruptly leaving, the peer which is leaving is not send any initially message to both inter and intra neighbours.This is because link crash or computer crash.

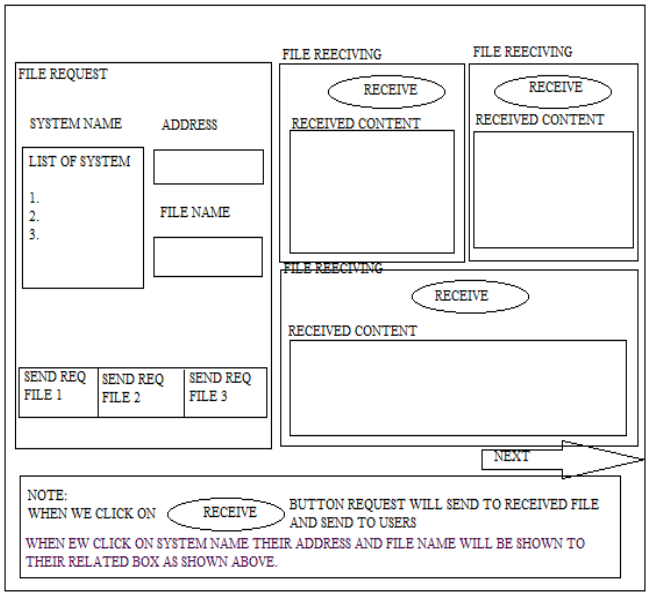

Figure 4 Peer details after peer joining and leaving

3.3 Network Coding Algorithm :

Network coding seen in different routing schema such as uncast, multicast and broadcast for more efficient data transfer. In network coding, any middle node is allowed to not only forward but also combine (code) data packets received from different incoming link. if necessary forwarding system in network coding is referred to as code-and-forward. There are two major streams of network coding. One stream investigates efficient encoding-and-decoding algorithms to increase the data transmission rate while reducing the computational cost and another emphasizes the applications of network coding

3.3.1Encoding :

Network coding in which multicast network where middle nodes perform simple linear operation on incoming packets. The pseudo-code for encoding algorithm is listed as Table 2 shows.

Table-2: Encoding Algorithm

Input : k (Original Packets $M_1, M_2, \ldots, M_k$)
Output: A_{new} (Encoded Packet)

1. Consider there are K original packets $M_1, M_2, \ldots$, Mk to be delivered from the source to one or more receivers.

2. Each packet contains encoding vector $A_i = (\alpha_{1i}, \ldots, \alpha_{ki})$ and information vector $X_i = \Sigma_{kk=1} \alpha_{ki} M_k$.

3. Assume there are m packets $(A_1, X1), \ldots, (A_m, X_m)$ that required to be linearly coded at intermediate node.

4. The node first takes a set of coefficients $(\beta 1, \ldots, \beta m)$ in GF $(2n)$.

5. To calculates the linear combination $X_{new} = \Sigma_{m\, i=1} \beta_i X_i$.

6. The new encoding vector A_{new} is calculated as $A_{new} = (\Sigma_{mi=1} \beta_i \alpha_{i1}, \Sigma_{mi=1} \beta_i \alpha_{i2}, \ldots, \Sigma_{mi=1} \beta_i \alpha_{in})$.

3.3.2 Decoding :

The receivers recover the original packets from the linearly combined packets, by solving a system of linear equations over a finite field. The pseudo-code for decoding algorithm is listed as Table 3 shows.

Table-3: Decoding Algorithm

Input : A_{new} (Encoded Packet)
Output: k (Original Packets $M_1, M_2, \ldots, M_k$)

1. Consider a receiver gets n packets: $(A_1, X_1), \ldots, (A_n, X_n)$

2. The node needs to solve the following n linear equations:

$X_1 = \Sigma_{kk=1} \alpha_{1k} M_k$
$X_2 = \Sigma_{kk=1} \alpha_{2k} M_k$
.
.
.
$X_n = \Sigma_{kk=1} \alpha_{nk} M_k$

3. To successfully recover the original data one needs to have: (1) $n \geq k$ i.e. the number of the received packets is no less than that of the original packets. (2) All equations are linearly independent

3.4 Distributed Bellman-Ford Algorithm :

Distributed Bellman-Ford Algorithm is also known as distance vector algorithm. It is simple and fast (i.e. do not require much extra process time) Distance vector algorithm Can be used for calculating best path (shortest path) in peer to peer network. Time complexity of distributed Bellman-Ford algorithm is O (V E) where V is the number of nodes and E is the number of link which is best complexity. We can find the shortest path from a given source node s to all other nodes [9].

IV ADVANTAGES OF PROPOSED SYSTEM

Following are some advantages for network coding algorithms to peer to peer file sharing.

4.1Throughput

Throughput is defined as the service given by peer-to-peer network provides per unit time. Here as let different peer-to-peer network send the same file, thus throughput can be simply represented by the time consumed by the transmission .Less the time consumed, the more the throughput. Let's start transmitting the file from time 0. Then the consumed time is the time when the peers stop receiving the file, denoted by finish time.

4.2 Reliability

This performance metric is used to assess the ability of the peer-to-peer network to overcome errors. Use the number of retransmissions to distinguish this ability. A peer-to-peer network with larger reliability will have a less number of retransmissions, and thus greater throughput. The redundant links can greatly improve the reliability of the peer-to-peer network with little overhead.

4.3 Link Stress

Link stress is defined as the number of copies of the same message transmitted through the same link. It is a performance metric that only used to an overlay network due to the mismatch between the overlay network and the physical network. Use it to estimate the effectiveness of the topology awareness improvement and the efficiency of the peer-to-peer network.

4.4 Scalability

Files are distributed through a peer-to-peer network. With the increase of the network size, the total available bandwidth also increases. By using file sharing between peer-to-peer using networks coding algorithm scalability issue removes.

4.5 Efficiency

The network coding algorithm is deterministic and easy to implement. There is no requirement for peers to collaborate to construct the linear coding scheme on demand. All the peers need is the mapping between different group ID and the encoding function, and this mapping is not change with time. Compared to random network coding, the receiver can always recover the original messages after receiving more different messages and the data dissemination is more efficient as data messages are transmitted through the same overlay link at most once.

4.6 Resilience

Churn is a common issue in overlay networks. By adding redundant links, the negative effect of churn is eliminated.

4.7Heterogeneity Support

In case that links have different link capacities, Peer-to-Peer File Sharing Based on Network Coding can arrange the overlay topology to maximize the utilization of each peer's link capacity.

V FUTURE WORK

Network coding algorithm having some issue in which when there is two or more source at the same time in multicast network then it become more complex so that multisource multi-sink issue is demanding in network coding algorithms

VI CONCLUSION

In network routing, messages are generally transferred by routing during intermediate nodes between the source and the destination i.e. by having intermediate nodes store and forward message. The traditional technique for multicasting in a computer network commonly is not optimal. In network coding algorithm refers to as where a node is allowed to generate output messages by encoding received messages. Thus, network coding allows information to mix, in contrast to the traditional routing approach where each node only forwards received messages. Network coding algorithm can greatly improve the throughput, consistency, scalability and efficiency of a multicast network

ACKNOWLEDGEMENT

First and foremost I would like to thank my guide for their guidance and support. I would forever remain grateful for the constant support and guidance extended by guides, in making this paper. The invaluable discussion they had with me, the penetrating question they had put to me and constant motivation, has all led to development of this Paper.

REFERENCES

[1] D. Hughes, G. Coulson, and J. Walkerdine, "Free riding on gnutella revisited: the Bell Tolls?," IEEE Distributed Systems Online Journal, vol. 6, no. 6, June 2005.

[2] R. Ahlswede, N. CAI, S.-Y. R. Li, and R. W. Yeung, "Network information flow," IEEE Trans. Inf. Theory,vol. 46, no. 4,pp. 1204–1216, Jul. 2000.

[3] S.-Y. R. Li, R. W. Yeung, and N. Cai, "Linear network coding," IEEE Trans. Inf. Theory, vol. 49, no. 2, pp. 371–381, Feb. 2003.

[4] R. Koetter and M. Medard, "An algebraic approach to networkcoding,"IEEE/ACMTrans.Netw.,vol.11,no.5,pp.782–795,Oct.2003.

[5] T. Ho, M. Medard, J. Shi, M. Effros, and D. R. Karger, "On randomized network coding," in Proc. Annu. Allerton Conf. Commun. Control Comput. 2003, pp. 4413–4430.

[6] T. Ho, M. Medard, R. Koetter, D. Karger, M. Effros, J. Shi et al., "A random linear network coding approach to multicast," IEEE Trans.Inf. Theory, vol. 52, no. 10, pp. 4413–4430, Oct. 2006.

[7] D. S. Lun, N. Ratnakar, R. Koetter, M. Medard, E. Ahmed, and H. Lee, "Achieving minimum-cost multicast: A decentralized approach based on network coding," in Proc. IEEE INFOCOM'05 Mar. 2005, pp. 1607–1617.

[8] N.Cai and R.W.Yeung,"Secure network coding", Proc.-IEEE International symposium on information Theory, pp.323, 2002.

[9] K. Bhattad, N. Ratnakar, R. Koetter, and K. R.Narayanan, "Minimal network coding for multicast," in Proc. IEEE Int. Symp. Inf. Theory, Sep. 2005, pp. 1730–1734.

[10] D. M. Chiu, R. W.Yeung, J. Huang, and B. Fan, "Can network coding help in P2P networks?" in Proc. Int. Symp. Model. Optimiz. Mobile Ad Hoc Wireless Netw. 2006, pp. 1–5.

[11] Y. Zhu, B. C. Li, and J. Guo, "Multicast with network coding in application-layer overlay networks," IEEE J. Sel. Areas Commun.,vol. 22, no. 1, pp. 107–120, Sep. 2004.

[12] Rathod Vijay U and V.R.Chirchi International Journal of Computer Applications (0975 -8887)Volume 129 –No.9, November201524"File Sharing between Peer-to-Peer using Network Coding Algorithm .

BIOGRAPHIES

Mrs. Ashwini Choudhari I am pursuing Master degree in computer science & engineering from Everest Education society's college of Engineering and technology Aurangabad.