Optimizing Technique to Reduce Traffic Load in Peer-to-Peer Systems

G. Thejonath Reddy **Dr. S. Vasundra**

Abstract— Peer-to-peer applications have gained remarkable increase over the years which can pose high pressure to internet service providers (ISP's) with high data volume traversing the internet. Recent studies show that these peer-to-peer applications generate 60% of the internet traffic worldwide, which raises intolerable ISP traffic. Hence the measurement studies and traffic engineering (TE) which can reduce the peer-to-peer traffic got importance. Many models which estimate the general traffic fail to capture the diverse peer-to-peer system characteristics. In this paper first measure the traffic between each peer nodes by considering certain factors like distance between the nodes, number of peers and response rate of each peer system. Then the greedy heuristic algorithm for deploying the cache devices can be proposed to deploy the cache devices between the peer nodes. Hence by using this technique the internet traffic by peer-to-peer application can be reduced.

***Index Terms*—Traffic Engineering, Measurement Studies, ISP traffic**

I. INTRODUCTION

Peer–to-peer computing has attracted large number of users across the world in recent years with its diverse uses, so this can make interested to research on this topic for the major number of organizations. These peer-to-peer systems are computers can form as network in order to share the data among themselves which is a better alternative to the present client-server model. This technology can be used in different areas like software organizations, scientific research centers, and multimedia organizations for efficient and fast data transfer.

In peer-to-peer network each node can act as both client as well as server. They can form in an overlay network [6] which is directed graph where a pair of systems connected by link can aware of each other IP address.

Peer–to-peer systems are having different applications like the systems can share and distribute the file example Bit Torrent [4] application, Napster which can reduce the transfer time of a system and load on the network. Media streaming is another application which can distribute the media to the systems with strict timing requirements example is PPLIVE [5]. Another application of peer-to-peer systems is making video and audio calls on internet which is most popular example SKYPE.

Though the peer-to-peer systems having the diverse uses and popularity they become target for attacks. Hence it has some challenges to overcome in order to provide the efficiency like data protection, membership control, managing the peer-to-peer systems.

Due to increasing popularity of peer-to-peer applications, the number of users using these applications has been increasing making a huge traffic in the internet. This can brings new challenges to the traffic engineering and network management group. More measurement studies [7] show that these peer-to-peer applications generating 60% of the internet traffic which can pose high pressure on internet service providers. This increase of load can make network more congestion and failure.

More models are proposed earlier for the traffic estimation gravity model [8], independent connection model for the general traffic fail to estimate the peer-to-peer traffic due to dynamic traffic characteristics of the peer-to-peer applications.

In this paper a model [1] is proposed to estimate peer-to-peer traffic which can make the deep analysis on the

ISSN: 2278 – 1323

different peer-to-peer traffic characteristics like distance between the peers, request rate, response rate, and flow throughput of the peer systems by using the real world datasets from file sharing systems and media applications we evaluate the model. Then find optimality to reduce the traffic in the peer-to-peer network.

II. Related Work

A. *Peer-to-peer systems:*

Peer-to-peer systems are a network of PC's can share the data among themselves which is an alternative to the present client-server model. Recent studies states that the peer-to-peer applications are getting high popularity in different areas of real world.

These peer-to-peer systems can broadly classify into two categories: measurement and improvement systems. Under measurement category we came to know how various peer-to-peer applications works which include distributing and sharing file systems [2], streaming media systems [3]. By these application systems we can transfer any type of data like files, multimedia. Hence it can generate huge amount of internet traffic worldwide which gives new challenges to the network management group.

The other category of peer-to-peer systems is improvement systems which deal with the improving of effectiveness and friendliness among the peer systems. There is large number of ways in order to improve the effectiveness like Bit Torrent application which divide the file into blocks and make the peer can easily serve the block which it can had to the requested peer, it applies (LRF) policy among the nodes.

Another one is locality awareness feature [10] which can make peer nodes available locally and can reduce the transfer time of data to other nodes and make network free from congestion.

B. *Models to Estimate Traffic:*

There are different models to estimate the traffic volume in a network. Some of them are briefly described below.

Gravity Model:

This model says that the traffic entering in the network at a particular node can exit the network at a node proportional to total traffic exiting in the network. Here in this model the starting node and ending node are independent.

IC Model:

In this model [9] the system not assuming the packets coming from the nodes are independent but the starting and exiting nodes of a connection is independent.

The above models can be well worked in connection oriented network but they fail to capture the characteristics of peer-to-peer systems.

III. Proposed System

Due to high popularity of P2P applications and large number of users the traffic increases. Hence traffic measurement and finding optimal solution in P2P networks got importance. The proposed system aims to provide greedy heuristic algorithm for deploying the cache devices which can store the recent requested objects by the clients. The model can reduce the traffic load of P2P systems. The system can be better useful in concrete application areas.

The system consists of the following stages in finding the optimality to reduce the traffic in peer-to-peer systems.

A. *P2P nodes:*

Peer-to-Peer network having definition "Decentralized communications model in which each node can act both as client and server". Here node means any PC's which have the capacity of storage and processing. This system of network can be arranged in different topologies but having one common feature i.e., sharing the resources and services by exchanging directly between the systems.

In peer-to-peer environment as shown in Fig 1, access rights are made by setting sharing permissions on individual machines.

ISSN: 2278 – 1323

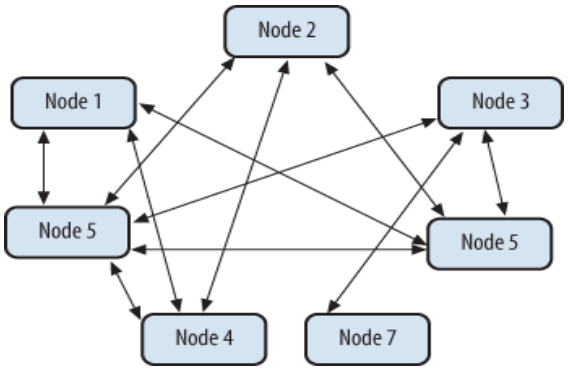

Fig 1: Peer-to-Peer network

B. Creation of clusters

There are more numbers of peers in peer-to-peer systems Therefore; system employs an aggregation process of user groups.

- Aggregation level: how many rounds individual peers are aggregated.
- Peer cluster: set of peers with certain common features at the aggregation level.

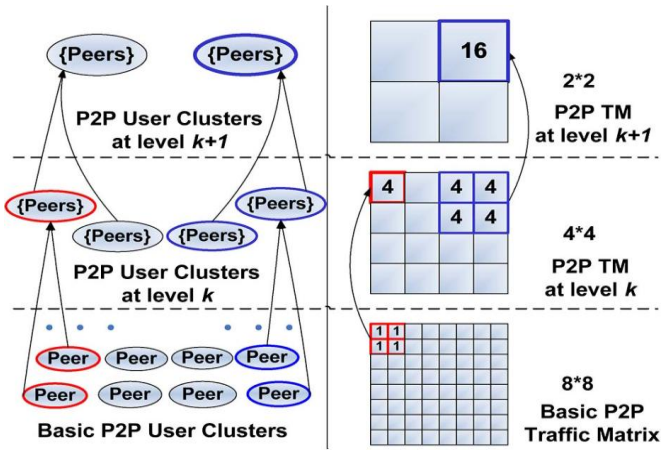

Fig 2: Peer-to-Peer clusters

C. Analyzing the traffic characteristics of each P2P node

Peer to peer systems having diverse behaviors hence it is difficult to find the traffic between them. System assumes that Peers remain stable within a certain time interval t. In order to analyze the traffic in peer-to-peer systems, system need to analyze the traffic characteristics of each peer nodes like selection of neighbor, requesting the data, response rate, and transmission of data.

In selection of neighbor phase a new peer node is register in server named tracker and retrieves the list of peers which have interested in same file. Here network distance between peer nodes has been measured.

$$P^s \propto |H_{ij}| / (d_{ij})^s \quad (1)$$

The probability of peer in list P^s is proportional to the population ratios $|H_{ij}|$ and inversely proportional to the network distance $(d_{ij})^s$ between them.

In requesting data phase the peer who is going to download the data can request to its neighbors for the data on the list. Then connections have been setup between the neighbors who accepted the request then we transmit the data.

$$T_i \propto D_i \quad (2)$$

$$P^r \propto U_i \quad (3)$$

Here the request rate T_i is directly proportional to the total downloading capacity D_i and the response probability P^r is proportional to the uploading capacity U_i.

By calculating all these traffic characteristics we designed a model to estimate the traffic Refer to "(1)." In the peer-to-peer network is illustrated in Fig 3.

$$X_{ij} = P^s{}_{ji} T_{ji} P^r{}_{ij} B_{ij}. \quad (4)$$

$P^s{}_{ji}$ = probability of peer containing in list.

T_{ji} = data request rate

Pr_{ij} = probability of response to the request

B_{ij} = flow throughput

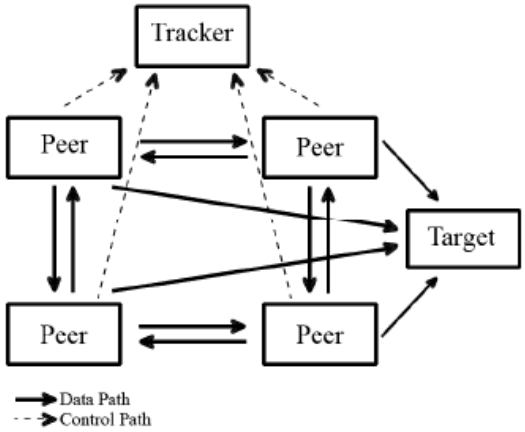

Fig 3: Working process of Bit Torrent.

ISSN: 2278 – 1323

D. Deploying the cache devices between the nodes:

The model can find the traffic flowing between peer nodes in the network by measuring all the peer-to-peer system traffic characteristics by following the phases specified in the previous section repetitively. Then by using the greedy heuristic algorithm the system deploys the cache devices by calculating the benefit link utility function which specifies where we want to install the cache by this we can reduce the traffic in overall network. By using this technique about 40% of the peer traffic can be reduced and the throughput of the system can be increased.

Algorithm: Greedy Heuristic algorithm

1. Dpy Set= {};
2. lftCosts=c;
3. nodeLks={ links of network}
4. While (true)
5. mxMargin=0; t=-1;
6. for i=1 to m do
7. if ci > LftCosts then
8. nodeLks= nodeLks-{li}
9. if(li ≠ nodeLks) continue;
10. m=((DpySet U {li})- (DpySet))/ci
11. if maxMargin < m
12. t=I; maxMargin=m;
13. end if
14. end for
15. If(t== -1)
16. Break;
17. nodeLks= nodeLks-{li}
18. lftCosts=lftCosts-ci
19. end while

In the above algorithm lftcosts(c) can be used to calculate the network traffic on each link, candidateLKs can represent the links between the peer-to-peer systems. By using the while loop iteratively calculate the optimal path between the source and destination.

IV. PERFORMANCE RESULTS

In performance analysis, we mainly focus on influence of different traffic characteristics of peer-to-peer systems on our proposed model and on its performance. We carry out a comparison of different existing models with the proposed optimizing technique by different traffic parameters on a set of randomly generated networks. In our simulation results we take 40 nodes uniformly spread over length of 200 units.

A. Aggregated error vs. Distance factor:

In order to evaluate the accuracy of the model which is developed to estimate the traffic of peer-to-peer systems we considered an aggregated error metric. The aggregated error can take the estimation errors of all origin and destination flows at all times. We also considered another parameter distance factor to find the estimation accuracy of model by varying the values of it. We taken these two parameters and draw a graph by simulating the system. This shows that the accuracy of our model outperforms than the previous models in estimating the traffic.

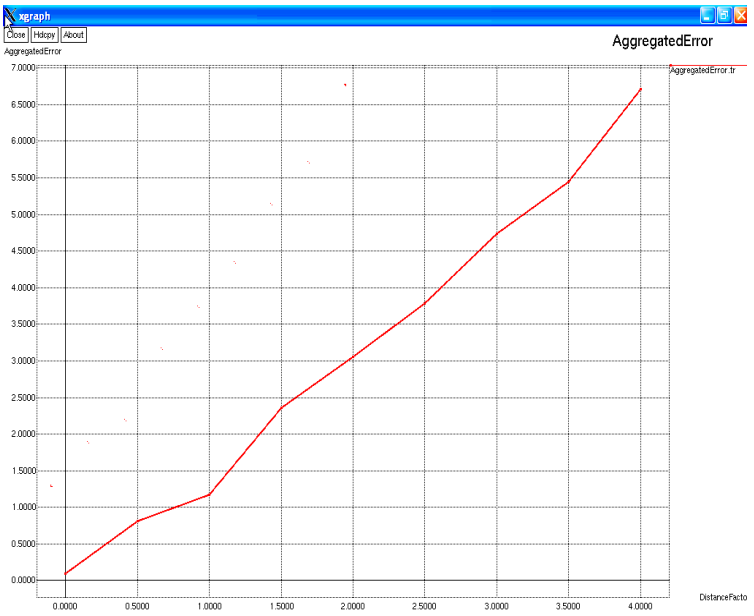

Fig 4: Aggregated error vs. Distance factor

B. Throughput of model:

Another parameter to check the performance of the system is throughput of the system. In the proposed system we deploy the cache devices by calculating the traffic across all

ISSN: 2278 – 1323

the links in the network. These devices can store the recent requested objects of the client. If the client request same object the cache device can serve the request. Hence the traffic and congestion can be reduced. Here we check number of packets transferred in certain time.

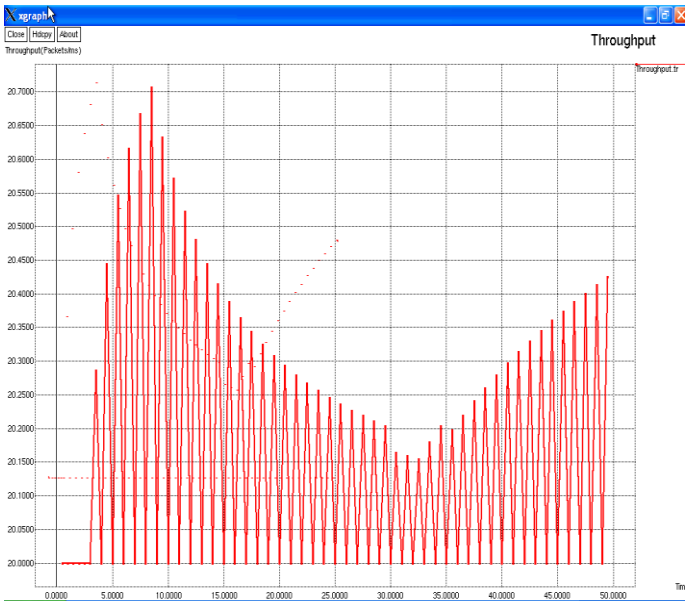

Fig 5: Throughput of system (packets/ms)

V. CONCLUSION

There is large number of existing models which fails to capture the traffic in peer-to-peer systems due to its diverse characteristics. Hence there is a need to design a model for estimating the peer-to-peer traffic features.

In the proposed system we created the peer-to-peer network then analyze the traffic characteristics of each peer-to-peer system which can reflect the better peer-to-peer network traffic. By considering these characteristics, the proposed system defines a model to estimate the peer to peer traffic then by using the greedy heuristic algorithm which can deploy the cache devices by checking the traffic on each link iteratively. The performance results on various factors like aggregated error with the distance factor and the throughput of the system shows that the proposed model can outperform than the existing model.

REFERENCES

[1] K. Xu,M. Shen, and M.Ye, ‘‘A Model Approach to Estimate Peer to-Peer Traffic Matrices,’’ in Proc. IEEE INFOCOM, Apr. 2011, pp. 676-684.

[2] S. Saroiu, P. Gummadi, and S. Gribble, ‘‘A Measurement Study of Peer-to-Peer File Sharing Systems Proc.Multimedia Comput. Netw., Jan. 2002, pp. 156-170.

[3] A. Mathur, H.Zhang. measurement of P2P live video streaming, in Recent Advances in P2P streaming 1st workshop, Aug2006

[4] Zhen xiao, Songqing chen and Xiaodong Zhang. Measurements, modeling, analysis of bittorrent system in IMC, 2005

[5] Hei, Liang, Liu. Insights into PPLIVE: A measurement study of large scale peer-to-peer IPTV systems, in 15th International world wide web conference, proc. IPTV workshop, may 2006

[6] O’ Mahony and Doval, “Overlay Networks A Scalable Alternative for peer-to-peer,” in proc. IEEE Internet Computing, vol 7, no. 4, pp 79-82, 2003.

[7] N. Traft, K.Salamatian and C.Diot. “Traffic matrix estimation existing techniques and new directions.” in proc. ACM SIGCOMM, pp 161-174, Aug 2002.

[8] M Roughan and N. Duffield, “Fast accurate computation of large scale IP traffic matrices from link loads.” in proc. ACM, pp 206-217, 2003

[9] M. Crovella and N. Taft, “An independent connection model for traffic matrices,” in proc. IMC, pp 251-256, 2006.

[10] Yi Cui, Yuan Xue, Bo Liu, “locality analysis of Bit Torrent Torrent like P2P systems,” proc. IEEE, 2010.

ISSN: 2278 – 1323

About the authors:

Dr S. VASUNDRA, presently working as Professor and Head of the Department CSE, JNTUA CEA. She completed her Ph.D from JNTUA university, anantapur, M.Tech from JNTUA and B.E from VTU. She is having 16 years of teaching experience and 5 years of research experience. Published 15 papers in various international journals and 3 in national journals. Her areas of interest include MANET's, Cloud Computing, Algorithms, Data Structures and Distributed Computing.

Mr. G. Thejonath Reddy pursuing M.Tech in software engineering from Jawaharlal Nehru technological university college of engineering, anantapur , Department of CSE. Got B.Tech degree from AITS college, Department of CSE. Areas of interest software methodologies, distributed computing, data structures, computer networks.

ISSN: 2278 – 1323