Simulation Modelling Practice and Theory 18 (2010) 1214–1224

Contents lists available at ScienceDirect

Simulation Modelling Practice and Theory

journal homepage: www.elsevier.com/locate/simpat

BitTorrent packet traffic features over IPv6 and IPv4

Cebrail Çiflikli [a], Ali Gezer [a,*], A. Tuncay Özşahin [a], Öznur Özkasap [b]

[a] Erciyes University, Kayseri Vocational High School, 38020 Kayseri, Turkey
[b] Koc University, Department of Computer Engineering, Istanbul, Turkey

ARTICLE INFO

Article history:
Received 10 December 2009
Received in revised form 12 April 2010
Accepted 26 April 2010
Available online 18 May 2010

Keywords:
BitTorrent
Peer-to-peer traffic
IPv6
IPv4
Traffic characteristics

ABSTRACT

At present, BitTorrent protocol packets constitute a large part of peer-to-peer application traffic on the Internet. Due to the increasing amount of BitTorrent traffic, it has become inevitable to take into account its effects on network management. Generally, studies on BitTorrent traffic measurement have involved analysis with packets transmitted via IPv4 protocol. However, with several facilities provided by IPv6 protocol, its traffic volume in operational networks is increasing day by day. New features of IPv6 enhance packet processing speeds over routers, switches and end systems. We consider that traffic features and packet traffic characteristics are likely to be affected with increasing amount of IPv6 protocol traffic. Therefore, it becomes significant to explore IPv6 packet traffic characteristics and application traffic features over IPv6. In this study, we investigate the IPv6 BitTorrent packet traffic characteristics in terms of autocorrelation, power spectral density and self similarity of packet size and packet interarrival time. We also perform distribution modeling for IPv4 and IPv6 BitTorrent packet traffic. With these models, efficient packet traffic traces are generated for network simulation studies. A detailed comparison is performed to determine differences between IPv4 and IPv6 BitTorrent packet traffic.

1. Introduction

Peer-to-peer (P2P) protocols, applications and their resulting network traffic have formed a significant part of the Internet [1–5]. A study performed in five different regions between August and September 2007 by IPOQUE [6] demonstrates that P2P file sharing produces majority of the Internet traffic, and its share varies between 48% (in the Middle East) and 80% (in Eastern Europe). P2P users cooperate through an overlay network that allows applications to share resources by acting as both clients and servers [7]. This approach was first popularized by the Napster system mostly for sharing music, video, and software files. Different download and file sharing strategies are adopted by several P2P applications. Some of them such as Napster use a centralized server to index files. Gnutella uses a fully distributed approach where queries are flooded to neighboring peers [8]. Some P2P applications such as BitTorrent, Gnutella 2, and Kazaa integrate centralized and distributed mechanisms. Supernodes in Kazaa, ultrapeers in Gnutella 2, and trackers in BitTorrent are responsible for handling index files for peers.

BitTorrent (BT) has been the most popular P2P file sharing protocol attracting millions of users since its introduction in 2001 [9–13]. BitTorrent system scales fairly well and is now widely used for various purposes, such as data distribution [15], media streaming, media on demand [16], and even to launch DDos attacks [17]. In recent years, there is a significant increase in P2P traffic amount, in particular, due to the popularity of BitTorrent protocol [13,14].

* Corresponding author. Tel.: +90 505 7795686.
E-mail addresses: cebrailc@erciyes.edu.tr (C. Çiflikli), aligezerr@hotmail.com (A. Gezer), atozsahin@erciyes.edu.tr (A.T. Özşahin), oozkasap@ku.edu.tr (Ö. Özkasap).

1569-190X/$ - see front matter
doi:10.1016/j.simpat.2010.04.009

For efficient resource utilization in networks, the nature of the traffic must be determined accurately. Due to the increasing amount of BT traffic on backbones, determination of its traffic characteristic would be invaluable. Most of the studies related to BT traffic are based on IPv4 protocol. With the increasing volume of IPv6 traffic in total, the characteristics of BT traffic via IPv6 must be determined elaborately. In this study, we investigate the characteristics of IPv6 BT packet traffic and the differences between IPv6 and IPv4 BT traffic in terms of spectral density, autocorrelation, distribution and self similarity of packet interarrival time and packet size.

BT divides a single large file into small pieces. Peers query a centralized server to retrieve a peer list, and can simultaneously download different pieces of the file from multiple peers. Several studies have been performed to model BT networks, BT peers, and BT traffic. For instance, modeling of BT-like P2P networks was studied in [18], and an analytical model of a BT peer was proposed in [19]. Furthermore, many studies have been conducted to investigate and model BitTorrent traffic. In [20], BT session characteristics and message characteristics were modeled using distributions. The self-similarity of BT traffic was estimated using variance time, Rescaled Range (R/S), and a periodogram estimator [21]. According to that study, the packet size of BT traffic gives Hurst degrees of 0.82, 0.77, and 0.83, respectively. Generally, in measurement studies of BT traffic, packets are transmitted over the IPv4 internet protocol.

Due to the restrictions of IPv4, it was essential to design a new internet protocol. Therefore, IPv6 was proposed in the middle of the 1990s, providing solutions to a number of shortcomings. 128 bit IP address size of the IPv6 protocol has solved the problem of limited address space of IPv4. This variation brings an advantage of a much greater number of addressable nodes on the Internet. IPv6 provides a simpler and more extensible header structure. Some IPv4 header fields have been dropped or made optional to reduce the common-case processing cost of packet handling and to limit the bandwidth cost of IPv6 headers [22]. Flow labeling capability is added to address the packets to a particular traffic flow with the new internet protocol. Despite the many facilities provided with the IPv6 protocol, its widespread adoption has been delayed for many reasons. It also seems unlikely IPv4 will be abandoned in the very near future. Nowadays, the Internet is in transition from IPv4 to IPv6. Transition from one to the other is widely achieved with IPv6-over-IPv4 tunnels. IPv6 traffic load in total is increasing day by day with the provided facilities such as tunneling and pilot programs. Many internet service providers use IPv6 pilot programs to provide IPv6 connectivity to their customers. Free IPv6 tunnel brokers make it possible for anyone to obtain an IPv6 connection [23]. The growth in facilities to provide IPv6 internet connections results in a traffic load much greater than that in the near past.

In this study, our aim is to make a characterization of IPv6 BT packet traffic. Our work comprises analysis of self-similarity, autocorrelation, and power spectral density for packet interarrival time and packet size. Moreover, detailed analyses are performed to determine characteristic differences between IPv6 and IPv4 BT packet traffic in terms of the statistics mentioned. Network traffic exhibits non-deterministic behavior. Therefore, distributions are more suitable to characterize packet traffic of computer networks. We also perform distribution modeling for IPv6 and IPv4 BT packets it terms of packet interarrival time and packet size. With correct distribution models, efficient packet traffic would be generated for network simulation studies.

This paper is organized as follows. Background information about BT protocol, IPv4 and IPv6 are presented in Section 2. Measurement details of inspected Internet traces and analysis perspective are provided in the next section. Packet interarrival time, packet size comparisons of IPv6 and IPv4 traffic in terms of autocorrelation, power spectral density, and cumulative distribution function are described in Section 4 together with self-similarity analysis and distribution modeling of the obtained time series. Finally, Section 5 states our conclusions.

2. Background

2.1. BitTorrent protocol

BitTorrent is a P2P content distribution system comprised of a set of network protocols for realizing communication among the participating peers. The idea behind BT is organizing the peers in such a way that the load of peer is distributed to the entire system [24]. It organizes peers into an overlay network to distribute files. Peers can connect several other peers simultaneously and download different blocks of the file in parallel.

For sharing files via BT, at first a torrent file is created which specifies the tracker and describes how the file is partitioned into small blocks. The tracker is the server responsible for keeping track of the registered peers and helps peers find each other. The initial distributer posts the .torrent file to the tracker. Distributed files are divided into small pieces and pieces are subdivided into smaller data units called blocks, typically 16 KB in size. A block is a data exchange unit in BT. To download a file, a peer first downloads the .torrent file associated to the file of interest, then contacts the tracker to obtain the list of peers sharing the file. Peers use this subset to connect to other peers to exchange the pieces of the file. The peer then requests blocks of the file.

BT employs rarest first policy to select pieces for downloading. Each peer tries to download the least replicated pieces among its neighbors. Therefore, the distribution of the file among the peers could be realized faster. A peer can only upload to a limited number of peers due to the choking procedure. Choking is the temporary refusal of uploading to some neighbors. If permission is given, unchoke message is sent. Choked neighbors are chosen according to tit-for-tat policy. In particular a peer uploads to the peers providing the best download rates. BT employs a tit-for-tat policy to penalize free-riding [24].