Chapter XXVI
Broadband User Behavior Characterization

Humberto T. Marques Neto
Federal University of Minas Gerais (UFMG)

Leonardo C.D. Rocha
Federal University of Minas Gerais (UFMG)

Pedro H.C. Guerra
Federal University of Minas Gerais (UFMG)

Jussara M. Almeida
Federal University of Minas Gerais (UFMG)

Wagner Meira Jr.
Federal University of Minas Gerais (UFMG)

Virgilio A.F. Almeida
Federal University of Minas Gerais (UFMG)

ABSTRACT

This chapter presents a broadband user behavior characterization from an Internet service provider standpoint. Understanding these user behavior patterns is important to the development of more efficient applications for broadband users. Our characterization divides the users into two categories: residential and small-office/home-office (SOHO). It employs four characterization criteria: session arrival process, session duration, number of bytes transferred within a session, and user request patterns. Our results show that both residential and SOHO session interarrival times are exponentially distributed, and point out that a typical SOHO user session is longer and transfers a larger volume of data. Our analysis also uncovers two main groups of session request patterns within each user category: (i) sessions that comprise traditional Internet services, such as WWW services, e-mail, and instant messenger, and (ii) sessions that comprise peer-to-peer file sharing applications, basically. We also analyzed and classified the e-business services most commonly accessed by users, which did not vary significantly across the user categories.

INTRODUCTION

The use of household Internet through broadband connections has grown over past recent years (Fukuda, Cho, & Esaki, 2005; OECD, 2006). This growth improves the use of typical applications, such as e-mail, WWW, and instant messengers, and, moreover, creates a good environment to foster others. Videoconferencing, interactive video and television, collaborative gaming, peer-to-peer (P2P) applications, grid-oriented computing, network-based backups, data-capable wireless network with portable networked gadgets that will used them, and voice over Internet protocol (VoIP) are some potential services which have been used in residential broadband network (MIT, 2005).

Peer-to-peer applications deserve special attention. Some studies analyze the growth of peer-to-peer usage in broadband networks (Gummadi, Dunn, Saroiu, Gribble, Levy, & Zahorjan, 2004; Hamada, Chujo, Chujo, & Yang, 2004; Lakshminarayanan & Padmanabhan, 2003). An environment with a good availability of Internet resources and with a good quality of service, compared with dial-up connections, encourages some users to search, share, download, and upload large files through the Internet. With the increasing demand on resources, the Internet service providers (ISPs) should create mechanisms to avoid the congestion of their backbones, because the growth in residential user-to-user traffic and in popularization of P2P systems has taken up a significant portion of backbone networks (Cho, Fukuda, Esaki, & Kato, 2006).

Understanding the nature and characteristics of broadband user behavior is a crucial step to improve the quality of service offered to users in broadband environments. Broadband user behavior characterization can lead to a better understanding of the interaction between users and service providers. It can also help the design of systems with better QoS metrics, such as performance, availability, security, and cost.

Studies of broadband user behavior are scarce in literature, mainly because of the difficulty in obtaining actual logs from Internet service providers. Most of the service providers on the Internet consider logs as very sensitive data. Existing studies, such as the one performed by Pew Internet & American Life (Pew, 2004), concentrate on qualitative analysis based on surveys. The Pew report shows how Americans' online behavior changes with high speed connections at home. The study also shows that broadband services allow users to distinguish themselves from dial-up counterparts in the following ways: (i) broadband users engage in multiple Internet activities on a daily basis, (ii) high speed users become creators and managers of different types of online content, and (iii) broadband users perform a large variety of queries for information. Yet, in another report of Pew Internet & American Life (Horrigan, 2005), the author points out growth in broadband adoption; however, he also shows that the growth rate is decreasing if compared with early years. In spite of Pew's reports, quantitative studies of broadband user behavior are still lacking.

This chapter intends to fill this gap reviewing the work presented in Marques-Neto, Rocha, Guerra, Almeida, Meira, & Almeida (2004). To understand the broadband user behavior, we present a characterization from a broadband ISP (a TV cable company that provides broadband services to its users), which classifies their users into two major categories: residential and small-office/home-office (SOHO). For each category, we identify user sessions, which are defined as the period during which a user is connected to the broadband network. Basically, the behavior of users is defined as a function of the way users arrive at the ISP, how long they remain online, the number of bytes they transfer and what they do while connected, that is, the request pattern within a session. Thus, the characterization process is performed along four criteria: (i) session arrival process, (ii) session duration, (iii) number of bytes transferred within a session, and (iv) user request pattern. The broadband user behavior characterization is based on logs collected on an authentication server and by Netflow (Netflow, 2004) running in a border router. The data collecting architecture implemented in the ISP allows us to identify the services used by each user category.