Simulation Modelling Practice and Theory 31 (2013) 96–115

Contents lists available at SciVerse ScienceDirect

Simulation Modelling Practice and Theory

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

Modeling Video on Demand services taking into account statistical dependences in user behavior

Ana Lobo *, Roberto García, Xabiel G. Pañeda, David Melendi, Sergio Cabrero

Department of Computer Science, University of Oviedo, Campus de Viesques, sn, 33204 Gijón-Asturias, Spain

ARTICLE INFO

Article history:
Received 28 November 2011
Received in revised form 2 October 2012
Accepted 8 October 2012
Available online 20 December 2012

Keywords:
Video on Demand
Multimedia systems
Workload characterization
User modeling
HD videos

ABSTRACT

The popularity of video services on the Internet, and moreover High Definition (HD) videos, has increased continuously in recent years. This growing demand for high quality video services can cause problems in current communication networks. For this reason, the characterization and modeling of video workloads have become essential factors in evaluating the performance of these services. This paper presents the implementation of a simulation model of a Video on Demand (VoD) service. The model takes into account both the behavior of the users and the server workload. While the former is analyzed through a real VoD service, the latter is analyzed through lab experiments. Nevertheless, the main contribution of our model is that it also considers the dependences among user interactions. The model has been validated using real data and evaluated in different situations. Results show that performance is influenced both by massive access to the contents and by the dependence structure of user interactions.

1. Introduction

Until recently, computer networks were used primarily for large data transfers, such as digital files. However, in recent years we have seen a spectacular growth of real time streaming services. Audio and video streaming is one of the most resource-consuming services of current networks and Internet servers. Furthermore, we may foresee an increase of the volume of traffic generated by these services due to the current trend towards high quality video delivery and the introduction of High Definition (HD) technologies. Therefore, the characterization and modeling of streaming workloads have become essential tools to evaluate the performance of these services and their impact on other services running on the same networks.

In this paper, we present a simulation model of a Video on Demand (VoD) service with HD video. The main novelty of this model is that it takes into account the dependences between user interactions. To design the model, we have previously analyzed the user behavior of one of these services, as well as the workload generated by a server that is offering high quality video.

User behavior was analyzed from the real VoD service LNE TV (http://videos.lne.es), which is the multimedia section of the online newspaper La Nueva España Digital (http://www.lne.es). This is one of the most successful digital news services in Spain and its VoD section (LNE TV) has several interesting features, such as a wide variety of subjects and a wide range of content lengths. As this is a real commercial and entertainment service, its user behavior is not conditioned by any educational or research environment. The details of the analysis were published in García et al. [10]. Apart from finding the most suitable distributions fitted to the empirical data, this analysis also considered correlation coefficients to study the

* Corresponding author. Tel.: +34 985182378; fax: +34 985181986.
E-mail addresses: loboana@uniovi.es (A. Lobo), garciaroberto@uniovi.es (R. García), xabiel@uniovi.es (X.G. Pañeda), melendi@uniovi.es (D. Melendi), cabrerosergio@uniovi.es (S. Cabrero).

1569-190X/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.simpat.2012.10.005

dependences between user interactions. All these results, with special emphasis on user interaction dependences, are included in our simulation model. The inclusion of these dependences makes our model more precise than previous work on workload analysis, in which only a single distribution function for each of the modeled features is usually considered. A good example of how critical dependences might be is when we need to model the length of user requests. This characteristic depends on the length of videos, but if the model considers both characteristics independently, the generated workload can produce an underestimation of the results.

Moreover, to characterize the workload generated by the server we have carried out several laboratory experiments. In these experiments, we delivered HD videos from a *Darwin Streaming Server* (DSS), generating streams between 2 and 10 Mbps per user request. By capturing the traffic between client and server we can characterize multimedia traffic, as well as the way standard multimedia protocols work.

After implementing and validating the simulation model, we have used it to evaluate the performance of a VoD service in different situations. We have mainly increased the number of VoD clients until we have reached what may be considered a massive access and we have also changed the dependences between user interactions. Under these different conditions, we have evaluated the influence of the number of clients or the dependence structure on critical parameters for a video service, such as the packet transfer delay or the amount of traffic generated in the network. Moreover, as previous studies have shown that current Internet traffic is usually self-similar, we have also checked the self-similarity of the generated traffic, as well as the influence of the different conditions on this property.

Our main contributions have been to develop a model of a VoD service, taking into account the real dependences between user interactions and considering the delivery of HD videos. We have performed the whole process, from the analysis of user behavior and HD video workload, to the implementation, validation and evaluation of the model.

The rest of the paper is organized as follows: in Section 2, we review the related work in this area. A description of the followed methodology is provided in Section 3. The analysis of the behavior of the users is summarized in Section 4 and the analysis of the workload of the server is presented in Section 5. Section 6 is dedicated to describe the simulation model and its validation, and Section 7 presents the results obtained during the evaluation of service performance. Finally, our conclusions are indicated in Section 8.

2. Related work

The growth of Internet VoD systems in the last few years has led to an increase in the research work in this field. In this paper we review the aspects directly related to user behavior, server workload, streaming control protocols, model implementing and the HD videos used in the experiments.

Almeida et al. [1] analyze user sessions and client interactions in an educational server. In a similar way, Brampton et al. [4] work in the sports and music world and also with new special interactions such as *book marking* (direct links to the most interesting points in the video). The statistical distributions for user behavior derived from the analysis made in [4] are also used in [27] to generate a synthetic P2P VoD dataset. The main drawback of these works [1,4,27] is that the special interactions and the type of contents offered clearly condition the behavior of the users, and perhaps the results obtained cannot be extrapolated to general environments. Nevertheless, the results used in our paper are obtained from the analysis of a real service that provides a great variety of contents and video lengths. Thus, user behavior is not conditioned by the contents and reflects the real interactions in entertainment VoD services.

Loguinov and Radha [18] analyze some interesting aspects about real time video streamed over the Internet, but working with low quality videos, and Kuang and Williamson [11] analyze *Real Media* traffic and the effect of streaming this traffic on an IEEE 802.11b wireless LAN. Nevertheless, the work presented in our paper analyzes high quality videos (more than 2 Mbps), anticipating most of the future contents on the Internet, sent by a streaming server using the standard protocols RTP (*Realtime Transport Protocol*) and RTCP (*RTP Control Protocol*).

When choosing the HD videos, the main aspects that condition the frame sizes, and thus bandwidth consumption, are movements and picture details. In fact, the size of each frame depends on the capacity of the encoder to compress information, as it is exploited in the encoders presented in [12,13]. This video compression and its subsequent transmission can distort the perceptual quality of digital video, so the goal of other works [2,24] is to analyze this distortion in HD videos. All these works obtain the test video sequences to check their encoders or analyze video qualities from the same sources [17,26]. We also use several HD video sequences, with different movements and details, from those sources to carry out our lab experiments and analyze server workloads. There was another possibility to obtain test video sequences, like the one used in [3,16], but the source used in these papers only provides the list of frames of each video, with their sizes and types. We prefer the complete video, to play it and make a characterization, not only of the frame sizes but also other features of the video transmission, to implement a more precise simulation model.

In the reviewed work on workload analysis in multimedia systems the dependence structure of the involved variables is not taken into account. The authors of these papers [1,4,9,11,18,25,27] only consider a single distribution function for each modeled feature. This simple approach can affect the results of simulation scenarios and lead to imprecise results. Nevertheless, in [10] it is shown that there is a dependence structure between user interactions. This dependence structure will be implemented in the simulation model presented in this paper, using the concept of copulas [20]. A copula is a function that links univariate marginal to their full multivariate distribution. Using a copula, it is possible to construct a multivariate