Computer Networks 79 (2015) 203–215

Contents lists available at ScienceDirect

Computer Networks

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

Lyapunov stability and performance of user-assisted Video-on-Demand services

Pablo Romero *, Franco Robledo, Pablo Rodríguez-Bocca, Claudia Rostagnol

Departamento de Investigación Operativa, Facultad de Ingeniería, Universidad de la República, Julio Herrera y Reissig 565, PC 11300, Montevideo, Uruguay

ARTICLE INFO

Article history:
Received 7 July 2014
Received in revised form 30 December 2014
Accepted 31 December 2014
Available online 17 January 2015

Keywords:
Video on-demand
Fluid model
Combinatorial optimization problem
GRASP
Path-relinking
Lyapunov stability

ABSTRACT

Currently, most video on-demand services offered over the Internet do not exploit the idle resources available from end-users, including YouTube. We present a taxonomic analysis of user-assistance in video on-demand systems, where users are both clients and servers, helping with the task of video distribution. From a theoretical perspective, we develop a deterministic fluid model suitable for sequential systems. We mathematically prove the Peer-to-Peer Sequential Fluid Model is globally stable in the Lyapunov sense, no matter the network parameters of the cooperative system. We theoretically prove that cooperative systems always outperform non-cooperative solutions. From a practical point of view, a caching problem is proposed and discussed in order to tackle technological concerns to massively distribute popular videos on-demand. The goal is to distribute video items into repositories minimizing the waiting times of end-users. The caching problem is inside the class of NP-Complete computational problems, and heuristically solved with a GRASP methodology enriched with a path-relinking technique. Predictions inspired in a statistical analysis of real-life YouTube traces suggest the introduction of cooperation is both robust and economically attractive. These results highlight the harmony between our theoretical development and practice.

1. Introduction

In today's Internet, we are witnessing an explosive development of video distribution, where Google videos, YouTube [1], NetFlix and PPTV [2] are valuable examples. Paradoxically, the most successful on-demand service in the world is expected to work far away from optimal. Established in 2005, YouTube has become the most successful Internet site providing a new generation of short video sharing service, comprising approximately nearly 10% of all traffic on the Internet [3]. However, the network access is yet working with a client–server architecture, and the operator (Google Inc.) must afford more than one million dollars per day just for bandwidth requirements, which is a clear motivation to exploit idle uploading resources from YouTube's users [4]. Experimental works also converge to the fact that user-assistance offloads the server and provides high scalability to video on-demand systems [5,6]. The deployment of video on-demand services must cope with several challenges, including asymmetric playback (users connect when they wish) interactivity (fast forward, rewind and pause options) and high playback quality over a bandwidth-sensitive best-effort infrastructure. Peer-to-peer networks represent an emerging promising alternative, where users self organize in an overlay topology developed at the application layer, and peers act as both client and servers. Intuitively a hybrid distribution including user-assistance could

* Corresponding author.
E-mail addresses: promero@fing.edu.uy (P. Romero), frobledo@fing.edu.uy (F. Robledo), prbocca@fing.edu.uy (P. Rodríguez-Bocca), crostag@fing.edu.uy (C. Rostagnol).

http://dx.doi.org/10.1016/j.comnet.2014.12.018
1389-1286/

offload the server of a raw client–server architecture. The most popular video services in Asia use P2P technology with high success, including PPTV [7] that in 2011 receives the largest investment in a video company since Googles acquisition of YouTube [8].

However, the introduction of the peer-to-peer philosophy comes at a cost. There are several scheduling policies related to peer cooperation: resources in the systems are highly dynamic, peers are heterogeneous (they have different broadband connections) and there are non-altruistic peers called free-riders (i.e. peers who wish to exploit the resources of the system without contributing with it). A thorough exposition of these challenges in a real-life peer-to-peer system is included by researchers from the university of Hong Kong [9]. They achieved millions of concurrent users enjoying a peer-to-peer on-demand service with some interactivity, acceptable quality of experience and reduction in the burden of servers.

In this article, we develop a deterministic fluid model in order to have a major understanding of trades-off and performance in on-demand peer-to-peer video streaming. Additionally, we measure caching versus waiting times in a simplified caching problem. In the whole research process, accuracy is slightly compromised to gain simplicity. This is a natural fact inherent to the art of mathematical modeling, where the results give an overview of the system's behavior, and suggest hints for the network design.

The article is organized in the following manner. Section 2 describes qualitatively the main challenges during the design of an on-demand peer-to-peer streaming service. Section 3 presents the related background theory, of assisted and traditional sequential fluid models. It is worth to remark that the literature in mathematical modeling of peer-to-peer networks is huge. The mathematical machinery usually involves Markov processes, marginal probabilities, game theory, ordinary differential equations and fluid models, epidemic models, among many others. In our selection, we provide cites to related works mainly from fluid models since we want to measure an expected behavior of peer, and focused in on-demand services. Related theoretical results are summarized in Section 4. Remarkably, sequential systems are globally stable in the Lyapunov sense (a proof is included in the Appendix). In practice, this means that a rest point is always achieved. However, we would like to remark that this rest point strongly depends on the network parameters (uploading capacities, sharing efficiency, starting number of seeders), which is sometimes hard to measure.

A caching problem is presented in Section 5. The issue is to define the video items that should be stored in repositories to offer minimal waiting times to end-users. We prove this combinatorial problem is inside the class of NP-Complete computational problems. Therefore, we develop a GRASP technique enriched with a path-relinking post optimization stage in Section 6. Real-life scenarios based on YouTube traces are analyzed in Section 7. The interplay between theory and practice, and a discussion of applicability is included in Section 8. Finally, Section 9 contains the main conclusions and discusses several aspects for future research.

2. Background and challenges

In this section we present a background of mathematical analysis starting from inspirational file-sharing systems. Then, we move into fluid models for on-demand peer-to-peer video streaming.

BitTorrent represented a revolutionary paradigm in the field of peer-to-peer networks. Originally conceived for file sharing applications and created by Bram Cohen, BitTorrent is an unstructured network overlay designed for fast distribution and replication of media contents [10]. The new concept is inspired in incentives: "give to get". The tit-for-tat solution of a game theory problem (the Iterated Dilemma's Prisoner) was included in this new design philosophy, and promotes an altruistic behavior of players [11]. Peers are self-organized via swarms. Once they enter the network, they are included in a swarm. With a pull-based technique, peers can get the rarest pieces of the file content first, in order to maximize the file availability. It exploits the robustness of random topologies, and peers cooperate with each other as soon as possible (in previous networks peers could contribute only after complete downloading).

Several mathematical models and experiments confirm its resiliency and robustness [12–21]. The models developed by Laurent Massouli and Milan Vojnović are similar in spirit than the ones of Yang De Veciana [22], and Qiu-Srikant [15], via coupon replication systems, with rigorous mathematical results [23]. They propose a deterministic fluid model for a large scale limit, in both layered and flat systems. In a layered system, users exchange coupons only if they have the number of coupons, whereas peers are uniformly selected at random in the flat system. In order to understand the flash crowd, a closed system is considered, and Poissonian processes are used for a steady state analysis. The models assume that joining peers are awarded from the server one random coupon. A remarkable conclusion is that both flat and layered systems are stable even under random blind coupon selection, Rarest First, though intuitive does not represent the reason of robustness in BitTorrent. An additional analysis of the complete layer system with all but one coupon missing shows that the population geometrically decreases, strengthening the robustness of swarming systems that were previously predicted by Yang and de Veciana [22]. Other works are focused on swarming policies and techniques to cope with the missing piece syndrome. Ji Zhu et al. formalize the idea that inter-swarm interaction (or universal swarms) provides excellent stability properties in BitTorrent [24]. The authors show that the stability region is insensitive to uploading capacity, piece selection policy and number of swarms.

On the other hand, a minority of works point out drawbacks in BitTorrent. Ji Zhu and Bruce Hajek study the number of seeders in order to escape from the missing piece syndrome in file sharing systems, under random peer-contact and random useful piece selection [25]. The authors remark that the missing piece syndrome can be avoided by a minor seed contribution, and the result is insensitive with respect to the piece selection policy. Barlas Oguz