Modeling and performance analysis of pull-based live streaming schemes in Peer-to-Peer network

Jianwei Zhang [a], Wei Xing [a,*], Yongchao Wang [b], Dongming Lu [a]

[a] College of Computer Science and Technology, Zhejiang University, Hangzhou 310027, China
[b] Library and Information Center, Zhejiang University, Hangzhou 310027, China

A B S T R A C T

Recent years mesh-based Peer-to-Peer live streaming has become a promising way for service providers to offer high-quality live video streaming service to Internet users. In this paper, we make a detailed study on modeling and performance analysis of the pull-based P2P streaming systems. We establish the analytical framework for the pull-based streaming schemes in P2P network, give accurate models of the chunk selection and peer selection strategies, and organize them into three categories, i.e., the chunk first scheme, the peer first scheme and the epidemic scheme. Through numerical performance evaluation, the impacts of some important parameters, such as size of neighbor set, reply number, buffer size and so on are investigated. For the peer first and chunk first scheme, we show that the pull-based schemes do not perform as well as the push-based schemes when peers are limited to reply only one request in each time slot. When the reply number increases, the pull-based streaming schemes will reach close to optimal playout probability. As to the pull-based epidemic scheme, we find it has unexpected poor performance, which is significantly different from the push-based epidemic scheme. Therefore we propose a simple, efficient and easily deployed push–pull scheme which can significantly improve the playout probability.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Recent works for Peer-to-Peer (P2P) streaming can be broadly classified into two categories: multi-tree based streaming and mesh-based streaming [1,2]. The multi-tree based systems organize all peers into one or more multicast trees, and diffuse different substreams along these trees by using application layer multicast routing techniques. Due to the good scalability, low server infrastructure cost and et al., mesh-based streaming has become one of the most important means of P2P live video streaming.

Mesh-based P2P systems can be approximately divided into two categories, *push-based* systems and *pull-based* systems, which are dependent on whether the local peer or the remote peer decides to upload a chunk. In particular, in order to improve the effectiveness of each transmission for a selected chunk, peers in both the push-based and pull-based systems may maintain a list of neighboring peers, and periodically exchange buffer maps with their neighbors.

In push-based systems, there exists the chance that two or more peers try to push the same chunk to the same peer. Usually this problem is already taken into account in most related works. In

pull-based (also known as *data-driven*) systems, there exists the chance that two or more peers try to send the same request to the same peer. To address this problem, most literatures assume that the overlay size is large enough to avoid such collisions, or peers have large enough upload bandwidth to deal with all requests from other peers in one time slot. In this paper, we demonstrate this assumption is unreasonable, and confine each peer's ability to process fixed number of requests per time slot. Moreover, we take into account both the random peer selection and random useful peer selection, and classify chunk selection strategies into three categories to facilitate understanding the streaming schemes in pull-based network.

More specifically, the contribution of this paper lies in the following four aspects.

- We establish the analytical framework for the pull-based streaming schemes in P2P network, and evaluate the peer first scheme, the chunk first scheme and the epidemic scheme in terms of the playout probability and the playout delay. The impacts of some important parameters, such as size of neighbor set, reply number, buffer size and so on are investigated in detail. Our analytical framework and evaluation results reveal the essential distinctions between pull-based streaming schemes and push-based streaming schemes.

- Unlike previous work towards the pull-based live streaming, we study both the chunk selection and the peer selection strategies analytically. We give accurate iterative equations of three most common chunk selection strategies, i.e., the latest first selection, the greedy selection and the random selection, and two most common peer selection strategies, i.e., the random peer selection and random useful peer selection.
- More importantly, we also investigate the impact of reply number, i.e., the maximum number of replies a peer can make in one time slot, on all considered streaming schemes.
- Aiming at enhancing the poor performance of the pull-based epidemic streaming scheme, we propose a simple, efficient and easily deployed push–pull scheme, which will not suffer from the transmission delay of the buffer maps.

The rest of this paper is organized as follows. Section 2 reviews related work. Section 3 describes our analytical framework for pull-based P2P streaming systems. Section 4 gives the analytical models of three categories of streaming schemes. The performance evaluation of all considered streaming schemes is given in Section 5, and we make the conclusion in Section 6.

2. Related work

As to the multi-tree based P2P systems, whether or not network coding is adopted has a huge influence on most important conclusions.

Liu et al. [3] theoretically investigate the minimum delay bounds for the single-tree, multi-tree and the proposed snow-ball P2P live streaming algorithm, which can approach the minimum delay bound. They also show the bandwidth heterogeneity among peers can be exploited to significantly improve the delay performance of all peers. Bianchi et al. [4] derive a performance bound for the stream diffusion metric in a homogeneous scenario, and show how this bound relates to the available upload bandwidth and the number of neighbors. Kumar et al. [5] develop a simple stochastic fluid model that accounts for the peers' real-time demand for content, peer churn, bandwidth heterogeneity and playout delay. The model provides closed-form expressions which can be used to shed insight on the fundamental behavior of P2P streaming systems. Liu et al. [6] study the tradeoffs between minimum server load, maximum streaming rate, and minimum tree depth under unconstrained peer selection, single peer selection and constrained peer selection. Through simulations, Magharei et al. [7] show that the inferior performance of the tree-based approach compared with the mesh-based approach should owe to the static mapping of content to a particular overlay tree and diverse placement of peers in different overlay trees.

As to the mesh-based P2P systems, scholars research the push-based systems mostly, but the analytical study of the pull-based systems is very insufficient.

Massoulie et al. [8] give the distributed algorithm that optimally solves the broadcast problem in edge-capacitated networks, and present the decentralized algorithm for solving the node-capacitated broadcast problem, and show some of its optimality properties analytically and through simulation. Sanghavi et al. [9] investigate one-sided push-based and pull-based chunk selection strategies, derive several very fundamental and important performance upper bounds based on the random gossip models, and propose a hybrid push–pull protocol INTERLEAVE.

Bonald et al. [10] prove that the random peer, latest useful chunk mechanism can achieve dissemination at an optimal rate and within an optimal delay, up to an additive constant term. They use mean-field approximations to derive recursive formulas for the

diffusion function of the latest blind chunk, random peer and latest blind chunk, random useful peer schemes. They give simulation results of various practically interesting diffusion schemes in terms of delay, rate, and control overhead. Some schemes can achieve near-optimal performance trade-offs. Fodor et al. [11] extend it by relaxing the complete-graph overlay assumption. They propose an analytical framework that allows the evaluation of scheduling algorithms. They consider the random and nearest strategies, and evaluate the playout delay, playout probability and the scalability of the two strategies. They conclude that the random strategy has fairly good performance for push-based P2P streaming systems. Furthermore, Chatzidrossos et al. [12] extend the previous work by considering the effect of the outdated buffer map information, node churn and a larger set of scheduling schemes.

Zhou et al. [13] study two chunk selection strategies, rarest first and greedy, and propose a mixed strategy that combines both of the two strategies. Furthermore, they study the tradeoffs between overlay size, buffer size and playout probability [14]. Zhao et al. [15] present a general and unified mathematical framework to study the large design space of priority-based chunk selection strategies. The analytical framework is asymptotically exact when the overlay size is large. For a given playout probability, the structure of the optimal strategy is fixed as a concatenation of a policy independent of buffer size and the rarest first strategy. Shakkottai et al. [16] study the minimal buffer size needed of the rarest first, greedy and the mixed strategies for a given playout probability, and validate that the mixed strategy can achieve order optimal performance.

Zhang et al. [17] propose an unstructured protocol with a push–pull mechanism to greatly reduce the latency, and evaluate this approach on PlanetLab. Moreover, they [18,19] model scheduling problem in data-driven streaming system as a classical min-cost network flow problem, and then propose both the global optimal scheduling scheme and distributed heuristic algorithm to optimize the system throughput.

Massoulie et al. [20] show the optimality of the random-useful packet forwarding algorithm for edge-capacitated networks, and also show the optimality of the most-deprived neighbor selection scheme combined with random useful packet selection for node-capacitated networks. Abeni et al. [21] present the formal proof that there exists a distributed scheduling strategy which is able to distribute every chunk to all N peers in exactly $\lceil \log_2 N \rceil + 1$ steps. Feng et al. [22] show that there exists a significant performance gap that separates the actual and optimal performance of pull-based mesh protocols. Moreover, periodic buffer map exchanges account for most of this performance gap. Dan et al. [23] analyze the performance and stability of P2P live streaming systems from the perspective of the packet loss rate and node churn. Compared with the authors' another publication [12], they further divide each streaming chunk into several packets.

This paper can be thought of as an extension of [13,14] and a complement of [10,12]. Especially, the method of modeling and numerical analysis in [12] is used as reference for our work. In [13,14], the authors assume that the selected peer's uploading bandwidth is large enough to satisfy all the requests in the same time slot. They also simplify the peer selection strategy, i.e., they assume each peer randomly selects another peer to download. In this paper, we relax above assumptions to achieve more important realistic meaning. From another perspective, our work also reveals the difference between the pull-based system and the push-based system studied in [10,12] in both modeling methods and their performance.

3. System modeling

As shown in Fig. 1, in a pull-based system neighbor relations between the peers and the forwarding of chunks are determined by