Peer-to-Peer Live Streaming: Optimality Results and Open Problems

Laurent Massoulié
Thomson,
Paris Research Lab, France

I. Introduction

Many peer-to-peer systems have been deployed on a large scale to provide users with "live streaming", that is Internet-delivered real-time multimedia content, much as the traditional television service (to name a few, CoolStreaming [9], TVants [8] and PPLive [4], UUSee [1] are examples in case). All such systems are organized in an *unstructured* manner, with *epidemic-style* information exchanges: users decide who to interact with in an adaptive manner, on the basis of past experience, and decide which data blocks to exchange with their "logical neighbours" by relying on simple, local rules.

Despite their success, the performance of such unstructured systems is still poorly understood, especially in comparison of structured systems, such as SplitStream [3], the performance of which is relatively well understood in symmetric scenarios. Nevertheless, because of their operational simplicity and their relative robustness to changing user populations (also known as user churn), unstructured systems seem to be winning against their structured counterparts, which require a significant overhead to reconfigure their logical topology upon user arrivals or departures.

It is therefore important to understand the performance limitations of unstructured approaches: if such designs can ensure live data diffusion at rates and delays that are competitive with those of more sophisticated approaches, then these must be the winning designs. In this paper we describe recent results which establish rate and delay performance optimality results for specific dissemination rules. We also formulate open problems concerning scenarios in which the efficiency of unstructured live streaming systems is still poorly understood.

II. Symmetric networks: rate and delay optimality

We first consider a symmetric network, comprising N users, including a traffic source, where each user has an upload capacity of one data chunk per time unit. We assume time is slotted, and all $N-1$ non-source users are to receive a stream of chunks initially injected at the source node.

In this context, we shall consider the so-called *random peer / latest useful chunk* push policy: in each time slot, each peer selects uniformly at random one peer to send data to; it then asks the destination peer which packets it has not yet received. The sending peer then forwards the packet with the latest time stamp, that is the latest that appeared at the source. Conflicts are resolved as follows. If several peers aim to send to the same receiver during one time slot, they are ordered in some arbitrary fashion. They then determine sequentially in this order what latest chunk to forward to the receiver, thereby avoiding to send several times the same chunk to the same receiver.

Clearly in such a context, since the overall uplink bandwidth is N chunks per time unit, and there are $N-1$ receivers, one cannot hope to disseminate a stream with rate above $N/(N-1)$ to all receivers. For large N, this limit is essentially 1. Moreover, a chunk present at m users at some time slot t can be present at no more than $2m$ users at time slot $t+1$. Thus the delay between the time when a chunk is injected at the source and when it reaches a random peer is of order no less than $\log_2(N)$ (see [2] for a more precise statement).

We now state a joint rate and delay optimality result. Here we assume that at each time slot, with probability $\lambda \in [0,1]$ a new chunk is injected at the source. λ is referred to as the injection rate.

Theorem 1: Let $\lambda < 1$ be fixed. Consider the *random peer / latest useful chunk* policy. Then there exists a constant $a > 0$ such that,for any $m \geq 0$, any arbitrary peer and any arbitrary chunk, the time delay D between injection of the chunk at the source and its reception at the peer verifies:

$$\mathbb{P}(D \geq \log_2(N) + m) \leq \frac{a}{m}, \quad (1)$$

for large enough N.
Thus the scheduling policy is such that for any injection rate λ strictly less than 1, for any peer, most packets get delivered in time $\log_2(N) + O(1)$, for large enough system size N. In this sense, the proposed policy is rate- and delay-optimal: a bound of 1 on the streaming rate and of $\log_2(N)$ on the transmission delay cannot be improved upon by any scheme, no matter how subtle it is in coordinating the operation of all participating peers.

The proof of this result, as well as extensive numerical evaluations of variants can be found in [2]. This builds on the results in [7].

III. Asymmetric systems

A. Complete graphs and access constraints

We now relax the assumption of symmetry of uplink bandwidth limitations, and assume that each peer u can upload at peer-specific speed c_u. We still assume that all peers can upload to any other peer, which is what we mean by the "complete graph" assumption. In this setup, we consider the following mechanism.

Each peer maintains a shortlist of neighbours. When it has finished sending a chunk, it measures how many chunks it could usefully send to all of its current neighbours; it then chooses to send to one of the neighbours to which it could send the largest number of chunks. We refer to this rule as the "most deprived peer" selection strategy. We then use the following chunk selection strategy: the source node sends preferentially chunks it has never sent before; if there are no such chunks, or for a non-source sending node, one selects uniformly at random, among chunks present at the sender and not at the receiver, which chunk to send. This is the so-called "random useful chunk" selection rule.

Finally, we assume that at some random instants, any given peer picks a new candidate peer at random, according to some sampling mechanism. It then adds the sampled peer to its neighborhood, and finally removes from its neighborhood one of the least deprived peers in it. Thus the neighborhood size remains constant. We refer to this as the dynamic neighborhood mechanism.

In this context, if we denote by s the source node, and by λ the injection rate at the source, to relay a stream of rate λ to all receivers it is necessary that

$$\lambda \leq \min\left(c_s, \frac{1}{N-1}\sum_u c_u\right). \tag{2}$$

Indeed, the source capacity c_s must be larger than λ, and the overall uplink bandwidth $\sum_u c_u$ must be larger than the total required download rate $(N-1)\lambda$. For a continuous time model, in which packet injection times at the source form a Poisson process with rate λ, and packet transfer times from some peer u are Exponentially distributed with mean $1/c_u$, there is a Markov process describing the system state, and for which we have the following result:

Theorem 2: When condition (2) holds with strict inequality, under the *most deprived peer / random useful chunk / dynamic neighborhood* policy, the Markov process describing the state of the system is ergodic. It therefore reaches a steady state, in which the time for a packet to be received by all peers is bounded in probability.

A proof of this theorem is given in [5] (see also [6]). It thus states that the proposed scheme is rate-optimal, since the constraint (2) on the injection rate is sharp; it says nothing on the magnitude of delays.

So far we discussed only single stream dissemination; we now introduce a multi-stream scenario and give a related result. Assume that there are several source nodes $s \in \mathcal{S}$, and that a set V_s of receiver nodes is associated with each such source s. By convention we let $s \in V_s$; these sets can overlap; however we do not allow a source s to be a receiver of another source $s' \neq s$, that is $s \notin V_{s'}$. We also restrict the exchanges that can take place in the following manner. A node u can only receive (and send) data relative to streams that it is a receiver for.

Under these assumptions, denoting by λ_s the injection rate at source s, it is not hard to see that necessary conditions for successful delivery are as follows:

$$\lambda_s \leq c_s, \quad s \in \mathcal{S}, \tag{3}$$

$$\sum_{s\in J}(|V_s|-1)\lambda_s \leq \sum_{u\in\cup_{s\in J}V_s} c_u, J \subset \mathcal{S}. \tag{4}$$

Indeed, the first equation guarantees that each source s is able to send out the stream injected into it; Condition (4) states that for each set of sources J, the total required download bandwidth is no larger than the total uplink bandwidth that can be used to deliver traffic from this set of sources.

In this context, we adapt the previous mechanism as follows: when a node measures how many chunks it can send to a neighbour, it does so over all commodities from all sources, and again chooses to send to the most deprived, according to this measure; chunk selection is done again uniformly at random among useful chunks, not distinguishing between distinct commodities. Sources still prioritize transmission of fresh packets (those never forwarded yet), if any. We call the resulting scheme the *bundled most deprived peer / random useful chunk / dynamic neighborhood* mechanism. The following is proven in [5]:

Theorem 3: When the conditions (3-4) hold with strict inequality, the Markov process describing the system state under the *bundled most deprived peer / random useful chunk / dynamic neighborhood* mechanism is ergodic. It therefore reaches a steady state, in which the time for a chunk from one source to be received by all peers in its receiver set is bounded in probability.

B. General topologies and edge constraints

We now relax the assumption that all peers can communicate directly to all other peers. Specifically we assume that there is a directed graph $G = (V, E)$, where V is the set of peers, and an edge (u, v) is present in the edge set E if and only if peer u can communicate directly to peer v. We also assume a different model of bandwidth constraints. Specifically, we assume that bandwidth resources are attached to edges: c_{uv} is the rate at which peer u can send directly to peer v, for any $(u, v) \in E$. In this context, there is no need to specify a peer selection mechanism, since there is no competition for bandwidth usage to send towards distinct neighbours.

Given a single source node $s \in V$, and an injection rate λ at s, a natural necessary condition for successful delivery at rate λ from source node s to all other nodes $u \in V$ is the following:

$$\forall W \subset V, s \in W,\ \lambda \leq \sum_{u\in W}\sum_{v\in V\setminus W} c_{uv}. \tag{5}$$

Indeed, the right-hand side is the total capacity at which one can send from W to $V \setminus W$; the above condition is clearly necessary for any node in $V \setminus W$ to receive data from s at rate λ. We then have the following:

Theorem 4: Assume that Condition (5) holds with strict inequality. Consider the *random useful chunk* selection policy along each edge of E. Then the Markov process keeping track of the system state is ergodic. It therefore converges to a steady state in which the time for a packet to be delivered to all receivers is bounded in probability.

IV. OPEN PROBLEMS

Many aspects need further investigation.

a) User churn: All the schemes described above can be implemented in the presence of user churn, without requiring any explicit steps to be taken to react to such churn. Since the schemes are rate-optimal, it seems intuitively that, so long as the churn ensures that there is enough bandwidth for successful delivery, performance is largely unaffected by churn. However this issue requires a more careful treatment, taking into account the time scale at which churn occurs.

b) Delay performance and topology: We have a clear understanding of delay performance only for highly symmetric systems. We may need to invent new unstructured mechanisms and analysis techniques to achieve joint rate- and delay-optimality for less symmetric environments, with particular network topologies.

c) Multiple streams: New results are needed to understand delay performance when there are multiple streams. Also, rate-optimality has been shown only in symmetric environments; it is open to determine rate-optimal schemes for multiple commodities under general models of network bandwidth limitation.

d) Relay nodes: All the results discussed here apply to the case where nodes only relay commodities for which they are receivers. It is again open to determine simple "unstructured schemes" which are rate- and delay- efficient, with or without network coding, under general network topologies and in the presence of relay nodes.

REFERENCES

[1] UUsee inc., http://www.uusee.com/.

[2] T. Bonald, L. Massoulié, F. Mathieu, D. Perino, and A. Twigg. Epidemic live streaming: Optimal performance trade-offs. In *Submitted to ACM Sigmetrics 08*, 2008.

[3] M. Castro, P. Druschel, A.-M. Kermarrec, A. Nandi, A. Rowstron, and A. Singh. Splitstream: High-bandwidth multicast in cooperative environments. In *Symposium on Operating System principles (SOSP 2003)*, Bolton Landing, NY, October 2003.

[4] X. Hei, C. Liang, J. Liang, Y. Liu, and K. W. Ross. Insights into PPLive: A measurement study of a large-scale P2P IPTV system. In *Proc. of IPTV Workshop, International World Wide Web Conference*, 2006.

[5] L. Massoulié and A. Twigg. Rate-optimal schemes for peer-to-peer live streaming. In *submitted to Performance Evaluation Journal*, 2008.

[6] L. Massoulié, A. Twigg, C. Gkantsidis, and P. Rodriguez. Randomized decentralized broadcasting algorithms. In *INFOCOM*, 2007.

[7] S. Sanghavi, B. Hajek, and L. Massoulié. Gossiping with multiple messages. In *IEEE Transactions on Information Theory*, pages 4640–4654, 2007.

[8] T. Silverston and O. Fourmaux. P2P IPTV measurement: A case study of TVants. In *Proc. of the 2nd Conference on Future Networking Technologies (CoNEXT 06)*, 2006.

[9] B. L. X. Zhang, J. Liu and T.-S. P. Yum. Coolstreaming/donet : A data-driven overlay network for live media streaming". In *INFOCOM*, 2005.