OPTIMAL OVERLAY GENERATION FOR P2P VIDEO STREAMING APPLICATIONS

Livio Lima, Marco Dalai, Pierangelo Migliorati, Riccardo Leonardi

{*name.surname*}@ing.unibs.it
Department of Information Engineering, University of Brescia, Italy

ABSTRACT

Video streaming with Peer to Peer (P2P) architecture appears to be an interesting solution for efficient streaming in the heterogeneous scenario of Internet applications. One of the key aspects in a peer-to-peer video broadcast system is to organize the peers into a high quality overlay for disseminating the video stream in order to optimize the bandwidth capacity of the P2P system. This work proposes an innovative approach based on Integer Linear Programming for the optimal network overlay generation, suitable to push-based approaches for data dissemination. Preliminary results show the flexibility of the proposed model to generate an overlay following different constraints or operational requirements. The usability and the computational complexity of the proposed method when the overlay includes an high number of peers is analyzed.

***Index Terms*—** Video streaming, peer to peer, overlay optimization

1. INTRODUCTION

Recently, there has been significant interest in the use of peer-to-peer technologies for Internet video broadcast/streaming [1], especially in heterogeneous systems, encouraged by the tremendous popularity that P2P applications such as file download and voice over IP have gained. The potential benefits of P2P technology for video streaming are low costs for system development (no infrastructure requirements) and scaling ease with increasing demand (resource sharing).

The main problem in a P2P streaming system is to organize the peers into a high quality overlay (the logical network that connects the physical nodes) for disseminating the video stream, since this can greatly affect the system performance. Many factors are involved in the overlay generation, as the number of peers, the input ad output bandwidth of each node, the latency, the delay, the link packet loss rate and other system considerations. In particular a critical aspect is the bandwidth utilization, since the video requires high bandwidth and different peers have variabile upload and download capacity. Consequently, it is important to obtain an overlay design with the maximum throughput in which the resources of each peer (both upload and download) are exploited as much as possible. Hereafter we consider as network throughput the sum of the data rates that are delivered to all peers in a network.

Existing overlay generation approaches can be grouped in two categories: tree-based and data-driven. References to different overlay approaches in literature can be found in [1]. Tree-based are typically push-based approaches. Each node just "pushes" the received data in determined data routes, and a parent-child relationship between nodes is defined. One of the most famous three-based approach is End System Multicast [2]. Data-driven are pull-approaches, each node independently requests portions of the video from neighbors, and nodes are typically connected in a mesh. In both approaches, the logical connections between peers are built by the exchange of messages within a subset of peers using specific algorithms for parent selection or partnership management. One of the most famous data-driven approach is Coolstreaming [3].

This work is a part of a large scale project, that has the objective to investigate a new system for video streaming adopting an hybrid approach where the peers are typically connected in a mesh and the data are disseminated by a push approach. Furthermore, in our approach the logical connections between peers that define the overlay are generated by an external entity, as a network manager, that creates the overlay through an optimization process based on peers properties specified during the system access phase and the estimation of network features. In particular in our approach this optimization is performed through the solution of an Integer Linear Programming (ILP) problem that produces the overlay in terms of active connections between peers and the data-rate to be transmitted on each link. It will be shown as this approach enables high flexibility in overlay generation, since different topologies can be created just including some constraints in the ILP model.

One of the critical aspects of the proposed method is the computational complexity since, as it will be shown in the following sections, the ILP problem dimensionality can greatly increase in presence of an high number of peers. In this work we evaluate two aspects of our approach for overlay generation: the flexibility to generate different topologies and the

This work was partially supported by the PRIN 2008 project prot. 2008C59JNA founded by the Italian Ministry of University and Research (MIUR).

complexity in order to demonstrate that the method can be applied also with an high number of peers.

The rest of the paper is structured as follow. Section 2 provides an overview of the P2P streaming system under investigation. Section 3 describes the proposed solution for overlay generation problem. In Section 4 the method will be evaluated in terms of flexibility and complexity while conclusions will be drawn in Section 5.

2. P2P SYSTEM OVERVIEW

As previously introduced, this work is a part of a P2P streaming system with an overlay based on a mesh topology in which each peer distributes data in a chunckless push-based fashion. Such approach potentially combines the benefits of tree-based and data-driven solutions. Typical push-based approaches show reduced overhead since a node receives data only from one parent and limited delay since the information a node receives is immediately replicated to its children. Data-driven approaches have more robustness, since each node receives data from multiple sources, avoid data duplication and enable better peers resource utilization. With the proposed approach the reduced overhead is obtained adopting a chunckless approach for data dissemination, while the optimization of peers resources is obtained with the generation of an optimized overlay. Such overlay includes the description of active links between peers and the data-rate to send on each link since each peer adopts a push approach for data distribution but it receives data from multiple sources (and send to multiple peers as well).

Typically, in most of the approaches in literature, this links description is obtained by the exchange of messages within a subset of peers using specific algorithms for parent selection and partnership management [1]. When a peer joints to the network it receives a list of active peers and dynamically exchanging peer status information and active peer lists the connections between peers are established and periodically re-evaluated. This approach has the advantage of limited complexity, since the overlay is built in a distributed fashion. In this paper we build the overlay as a solution of an optimization problem. When a peer joins to the network it provides its status information to a network manager that is in charge to collect peers information, optimize the network and then distribute routing information to peers. This optimization is performed periodically, when changes in peers or network status happen. The proposed mechanism for such optimization is the Integer Linear Programming (ILP) modeling, that offers several advantages as the high flexibility to obtain a rich problem description and the computational advantages of Linear Programming. More details about ILP optimization will be provided in Section 3. It has to be noted that in other previous works as [4], ILP has been proposed for overlay generation, but only to describe the active links, since a pull-based approach was adopted.

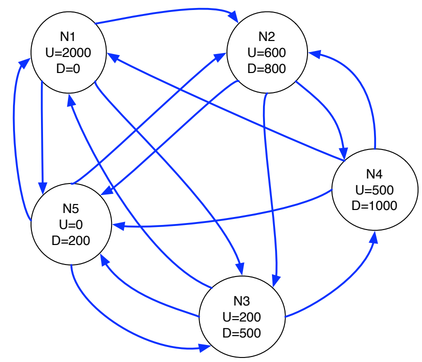

Fig. 1. An example of a mesh with N=5 peers and M=3 connections per peer.

The adoption of a multi-source push-based chunkless approach introduces the problem of robust dissemination of data in order to prevent duplication. One possible solutions is to use network coding based approaches as [5, 6], that are recently gaining popularity for P2P applications. In such approach each peer first decodes and combines the data received from multiple-sources, then re-encode the data generating different data descriptions to distribute to its children following the bandwidth utilization map provided by the overlay generation mechanism. The main drawback using network coding approaches is the possible increased delay due to the data decoding/encoding process.

3. OVERLAY GENERATION

In this section we describe how the Integer Linear Programming can be used for the optimal overlay generation with network throughput allocation. We adopt the ILP approach since it is a common approach in other disciplines as the Operations Research to solve optimization problems since it can offer high flexibility in problem modeling and computational advantages due to the linearity.

In this work we consider a simplified model of the overlay generation process, stated as follows. Let us consider N peers as represented in Figure 1, each one defined by its upload (U) and download (D) capacity and potentially connected to M other peers, where obviously $M < N$. The objective of the optimization is to determine which links have to be maintained active and the data-rate transmitted on each link, in order to minimize or maximize a particular target function. In this work we consider the maximization of the network throughput, expressed as the sum of the data rates that are delivered to all peers in a network.

Nevertheless this model includes several limitations. First we assume that only peer status information, as the upload and download capacity, are involved in the overlay generation process. Information about the physical network, as delay and the link packet loss rate are not included at the moment, but ongoing investigation is currently devoted to extend the model in order to include such information. Another limitation is that we assume links with very high capacity dedicated to video streaming data, even if this condition is not true in general due to traffic congestion. Consequently, in our model, the data rate transmitted on each link is limited only by the upload and download characteristics of involved peers. Additionally, typically P2P applications suffer of peer churning problem, i.e. the peers join and leave the network with high frequency. Consequently the overlay has to be updated each time that a node join the network and, if we want to limit the performance degradation, also when a peer leaves the network. If the overlay generation requires high computational complexity, the approach could be unusable. In this latter case a possible solution could be to perform the optimal overlay generation periodically, and applying sub-optimal strategies in the middle.

The general formulation of an ILP optimization problem is given by:

$$Z = \min \mathbf{cx}$$

$$\text{subject to} \begin{cases} A\mathbf{x} \geq b \\ B\mathbf{x} \geq d \\ \mathbf{x} \geq 0 \quad \text{integer} \end{cases} \tag{1}$$

where $\mathbf{x}$ is the vector that represents the problem unknowns, Z is the function that has to be optimized while the problem characteristics are described by the inequalities. For the specific problem of optimal overlay optimization, the vector $\mathbf{x}$ can be expressed as $\mathbf{x} = \{y_1, ..., y_T, w_1, ...w_T\}$, where $y_i \in [0, R]$ represents the data rate transmitted on link i and $w_i \in \{0, 1\}$ is a binary variable indicating if the link i is active. Consequently each one of the $T = MN$ links of Figure 1 has associated a couple $\{y_i, w_i\}$. As previously introduced, in our problem the function Z to optimize is represented by the overall throughput, that is represented by $\sum_{i=1}^{T} w_i y_i$.

The problem description is obtained defining a set of constraints, generally represented as $A\mathbf{x} \geq b$ and $B\mathbf{x} \geq d$ in model (1), that are problem-specific. For the overlay generation problem we have identified several constraints, described in the following.

For each peer k let us assume that U_k and D_k are respectively the maximum upload and download capacities, $\mathcal{I}_k$ and $\mathcal{O}_k$ are the set of indexes related respectively to the ingoing and outgoing links. The basic set of constraints are related to the ingoing and outgoing bandwidth capacity of each peer:

- the ingoing rate is limited by the download capacity, i.e. $\sum_{i \in \mathcal{I}_k} y_i \leq D_k$
- the ingoing rate of non-source peers has to be non-negative, i.e. $\sum_{i \in \mathcal{I}_k} y_i > 0$
- the outgoing rate is limited by the upload capacity, i.e. $\sum_{i \in \mathcal{O}_k} y_i \leq U_k$

Configuration	Constraints
C_1	basic + (c)
C_2	basic + (c) + (d)
C_3	basic + (a) + (b) + (c) + (d)

Table 1. Configurations used in the experiments of Section 4.

- each peer can redistribute only the data received, i.e. $y_j \leq \sum_{i \in \mathcal{I}_k} y_i, \forall j \in \mathcal{O}_k$

Additionally, depending on the particular overlay topology and mechanisms for data dissemination, further constraints could be introduced. More details on the adoption of this additional constraints are provided in Section 4. In particular we identify the following set of additional constraints:

(a) Limit the number of ingoing connection for each peer, i.e. $\sum_{i \in \mathcal{I}_k} w_i \leq I$

(b) Limit the number of outgoing connection for each peer, i.e. $\sum_{i \in \mathcal{O}_k} w_i \leq O$

(c) Avoid data dissemination loops

(d) Limit the number of hops, i.e. the number of peers crossed by the same data flow.

The constraint on maximum number of hops could be particularly useful to limit the delay in data dissemination. The constraints on maximum incoming and outgoing connections can also be imposed to only a restricted number of peers, or differentiated by groups of peers. Avoid data loop is a mandatory constraint in case we want to obtain a tree configuration.

4. PRELIMINARY EXPERIMENTS

In this section we present an experimental analysis of the proposed approach for overlay generation. In particular our solution is evaluated from two point of view: the flexibility to generate different overlay topologies and the complexity of the proposed approach in particular when the optimization is performed on a large group of peers. The complexity is evaluated as the computational time required to solve the ILP problem described in section 3. ILP problems can be solved using solvers for mixed integer linear programming problems; in our experiments we used CPLEX solver [7]. In both the set of experiments we create the network to optimize randomly extracting N peers within a set of five possible configurations: $(U = 1000, D = 1000)$, $(U = 200, D = 1000)$, $(U = 200, D = 500)$, $(U = 600, D = 800)$, $(U = 500, D = 1000)$. The networks generated include also one source with upload bandwidth equal to 2000, and a sink node with download bandwidth equal to 200.

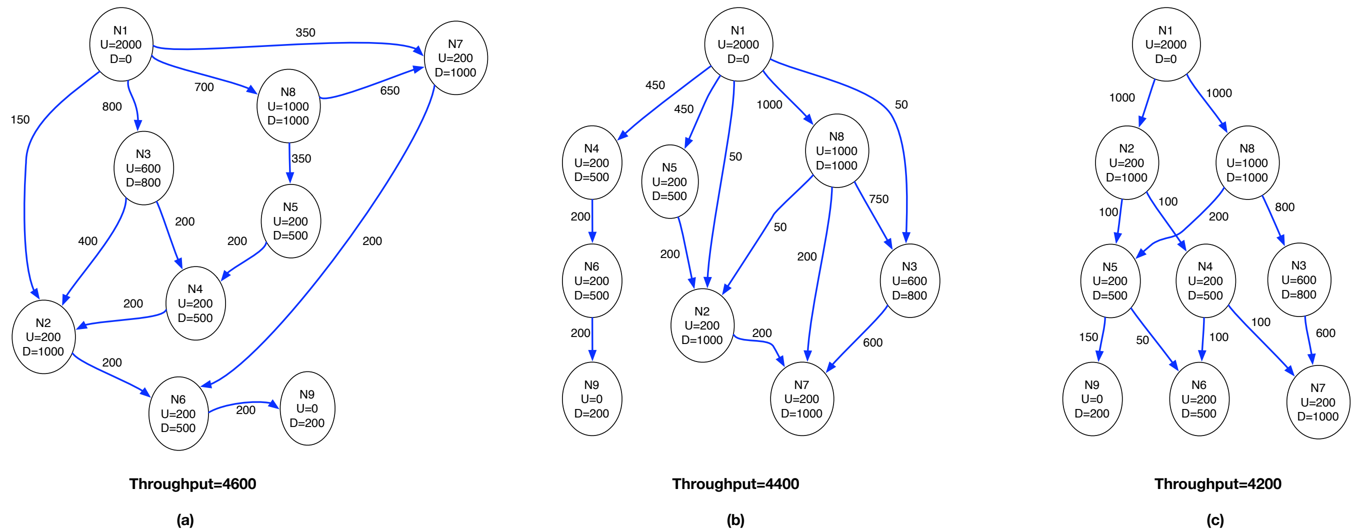

Fig. 2. Optimal mesh for the configurations specified in Table 1: (a) C_1, (b) C_2 and (c) C_3

4.1. Analysis of flexibility

In order to evaluate the flexibility of the proposed approach, we present the solutions of the overlay generation problem described in Section 3 obtained under different constraints configurations, as reported in Table 1. Figure 2 shows the overlays generated for configurations C_1, C_2 and C_3 for 9 peers with different features in terms of upload and download bandwidth. In configuration C_2 and C_3 the number of hops is set to 4, while in configuration C_3 we set $I = O = 2$. For each configuration is reported the value of the obtained throughput.

As it can be noted form Figure 2, additional constraints cause a decreasing of the effective throughput. Given the bandwidth features of the peers it is easy to show that the the maximum reachable throughput T is given by:

$$T = \min\left(\sum_k U_k, \sum_k D_k\right), \qquad (2)$$

that for the particular peers configuration of Figure 2 is equal to 4600. This value is obtained only in configuration C_1, while configurations C_2 and C_3 have lower throughput since the additional constraints preclude the possibility to reach the maximum value. In general, the more constraints are included the lower is the obtained throughput.

4.2. Analysis of complexity

The experiments of section 4.1 have been done configuring the overlay with a limited number of peers. Nevertheless, it is expected that, typically, P2P streaming applications involve hundreds of peers. The proposed method for overlay generation is based on the solution of a ILP problem, that generally has complexity increasing with the problem dimensionality. Consequently, it is expected that the proposed method could be unusable with a large number of peers due to the high computational complexity. In particular, analyzing the problem properties, the factors that strongly affect the model complexity are: the number M of possibly active links for each peer and the presence of constraints of type (c) and (d), where the later is, at the moment, still under investigation.

Figure 3 shows the complexity of the proposed method evaluated as computational time to solve the overlay generation problem. Here it is evaluated the effect of the parameter M (the number of active links per node) in case of configuration C_1 with the additional constraints (a) and (b) ($I = O = 4$) for 100 peers (Figure 3a) and 200 peers (Figure 3b). For each value of M it is indicated the value of the obtained throughput that, as can be noted by the values of overall download and upload bandwidth reported in Figure 3, is always lower than the maximum reachable value T given by equation (2).

Figure 4 shows the throughput obtained optimizing a tree overlay with the configuration C_1 with the additional constraints (a) and (b) ($O = 5$, $I = 1$) for 100 peers and different values of M, limiting the computational time to 15 seconds. For each value of M, Figure 4 shows also the time spent for the solution of overlay optimization problem.

From the analysis of Figures 3 and 4 it is clear as the constraint types (a) and (b) do not impact on the problem complexity. Furthermore, we have shown as also with low value of parameter M it is possible to obtain the maximum value of reachable network throughput. This enables limited complexity for the overlay generation problem also in presence of a large number of peers. In further experiments not reported here, we have evaluated that for a values of M lower than 10, it is possible to solve the overlay generation problem in

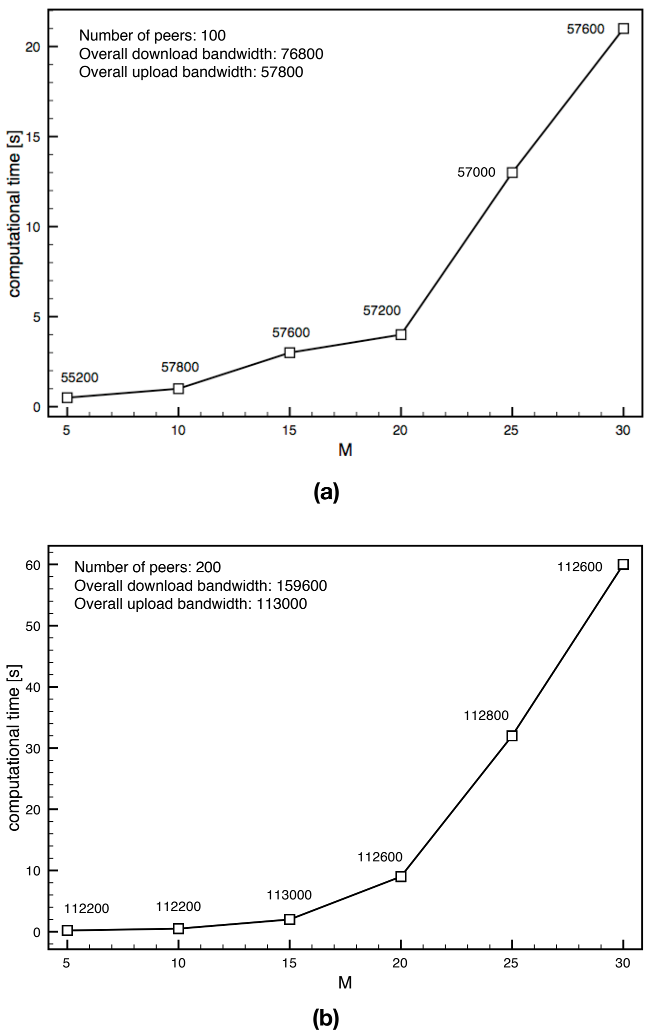

Fig. 3. Evaluation of the effect of the parameter M on the overlay generation complexity for a) 100 peers and b) 200 peers

less than 10 seconds considering up to 500 peers. In case of networks with more than 500 peers it is possible to define algorithms that split the peers in sub-groups and generate the overlay on each group.

5. CONCLUSIONS

This work describes a new approach for overlay generation in a P2P video streaming system adopting a Integer Linear Programming approach to model the problem. The ILP problem solution describes the overlay in terms of active connections between peers and data-rate to transmit on each link in order to maximize the network throughput.

The reported experiments show as the proposed method enables high flexibility to describe different overlay topologies, as tree, multi-tree and mesh. Furthermore, from the

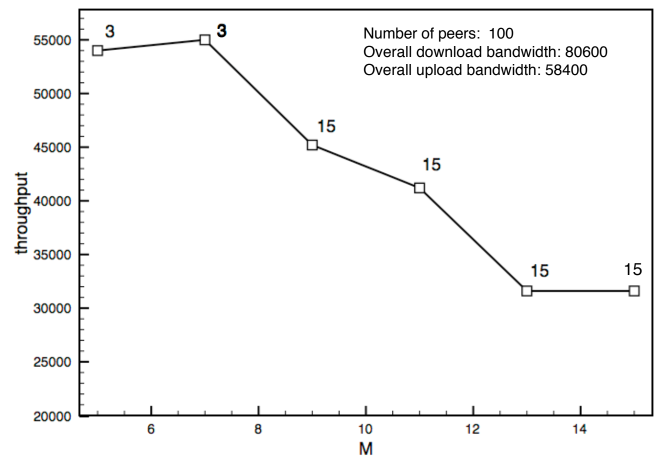

Fig. 4. Throughput analysis for tree overlay limiting the computational time to 15 seconds for different values of parameter M

complexity analysis we shown as, under an appropriate setting of model parameters, the computational time to solve the problem can be maintained limited, and consequently the proposed method is suitable to be included in a more realistic P2P streaming system.

6. REFERENCES

[1] J. Liu, S.G. Rao, B. Li, and H. Zhang, “Opportunities and challenges of peer-to-peer internet video broadcast,” *Proceedings of the IEEE, Special Issue on Recent Advances in Distributed Multimedia Communications*, vol. 96, no. 1, pp. 11–24, 2008.

[2] Y.H. Chu, S.G. Rao, B. Li, and H. Zhang, “A case for end system multicast,” in *Proc. ACM SIGMETRICS*, 2000.

[3] X. Zhang, J. Liu, B. Li, and T.-S. P. Yum, “Donet/coolstreaming: A data-driven overlay network for live media streaming,” in *Proc. IEEE INFOCOM*, 2005.

[4] R. Iqbal and S. Shirmohammadi, “An analytical approach to model adaptive video streaming and delivery,” in *Proc. Workshop on Advanced Video Streaming Techniques for Peer-To-Peer Networks and Social Networking*, 2010.

[5] C. Gkantsidis and P. Rodriguez, “Network coding for large scale content distribution,” in *Proc. IEEE INFOCOM*, 2005.

[6] R. Bernardini, R. Rinaldo, and A. Vitali, “A reliable chunkless peer-to-peer architecture for multimedia streaming,” in *Proc. Data Compression Conference*, 2008.

[7] ILOG INC., “ILOG CPLEX 12.1 reference manual,” Incline Village: ILOG Inc., CPLEX Div., 2009.