Cooperative Caching and Relaying Strategies for Peer-to-peer Content Delivery

György Dán
ACCESS Linnaeus Center, School of Electrical Engineering
KTH, Royal Institute of Technology, Stockholm, Sweden
E-mail: gyuri@ee.kth.se

***Abstract*— Peer-to-peer content distribution has become a major source of bandwidth costs for Internet service providers (ISPs). One way for ISPs to decrease these costs is to deploy caches for p2p traffic. To make efficient use of the caches, in this paper we propose a cooperative caching and relaying scheme that is compatible with the existing business relations between ISPs. We formulate the problem of cooperative caches as a resource allocation problem, and show that it is related to the problem of r-configuration studied in graph theory. We propose a distributed algorithm to solve the resource allocation problem, and show that cooperation leads to significant gains compared to non-cooperative caching.**

I. Introduction

Peer-to-peer content distribution offers content providers the promise of data delivery to a large population of users without the need for big investments in server capacity in terms of processing power and access network capacity. The costs of the data delivery are shared among the spectators - the end nodes - and their Internet service providers (ISPs).

The delegation of costs from the content provider to the ISPs led to frictions in the case of off-line content distribution: with 50-70 percent of their network load due to peer-to-peer file sharing applications, some ISPs started to throttle the bandwidth of p2p file sharing applications, which in turn introduced encryption to avoid throttling (e.g., Bittorent and E-mule).

Peer-to-peer streaming, though still in its infancy, may also turn into a major source of costs for ISPs once it becomes popular worldwide. Large p2p streaming providers serve more than 100 thousand peers simultaneously every day [1], [2], located mostly in China. This corresponds to around 40 Gbps aggregate download bitrate, but larger broadcasts with nearly 1 million participating peers were also recorded [2]. Without any collaboration between streaming providers and ISPs, one can expect ISPs to start throttling streaming p2p traffic.

Locality aware peer selection could decrease the inter-ISP traffic generated by p2p overlays, but despite its potential benefits [3], measurement studies indicate that such techniques are not widely used in commercial p2p systems [1], [2]. Alternatively, ISPs can decrease their inter-ISP traffic by employing caches for p2p traffic, such as PeerCache for FastTrack [4]. ISPs do not infringe copyright by caching and are not liable for caching illegal content in the US and in the EU. The deployment and the maintenance of p2p caches incurs however costs, hence ISPs are interested in making efficient use of these resources.

In this paper we propose an architecture that enables ISPs to use their p2p caches more efficiently. The key elements of the architecture are *collaborating relay nodes*, i.e. caches, deployed by peering ISPs. The caches of the ISPs cooperate to serve each others' subscribers and hence decrease the amount of IP transit traffic. The proposed cooperation scheme can be applied to caches for p2p file sharing systems, but we describe the scheme in the context of p2p streaming systems (and hence we use the term *relay* instead of *cache* from now on).

The proposed collaboration scheme offers several incentives for both ISPs and streaming providers. First, the IP transit traffic of an ISP can be decreased by its own relay nodes and by those of its peering ISPs at the cost of increasing the peering traffic. Second, collaboration between ISPs and p2p content providers is cheaper on the long term than throttling p2p traffic. Third, the relay nodes can improve the quality experienced by the users, hence the architecture gives an advantage to early adopters.

The rest of the paper is organized as follows. In Section II we describe our model of peering relations between ISPs. We describe the proposed application layer peering scheme in Section III, and present resource allocation

This work was in part supported by the Swedish Foundation for Strategic Research through the project Winternet.

strategies for the peering scheme in Section IV. We evaluate the performance of the proposed cooperative relaying strategy in Section V and describe the related work in Section VI. We conclude our work in Section VII.

II. Background and definitions

ISPs ensure global reachability through buying IP transit services and through maintaining bilateral or multilateral settlement-free peering agreements. Settlement-free peering agreements enable ISPs of similar size and geographic coverage to exchange IP traffic freely for mutual benefit. Hence, the costs of peering are insensitive to the short term fluctuations of the amount of exchanged traffic (as long as the traffic does not cause congestion, in which case a port congestion charge might apply at public peering exchanges, e.g., at Lynx). IP transit traffic is however usually charged according to the 95 percent rule (i.e., the client pays for the 95 percentile traffic calculated over a month), and hence increased traffic leads to an increased cost even on the short term. Consequently, a caching or relaying scheme should strive to minimize the IP transit traffic but may increase the peering traffic as long as it is kept balanced.

We model the network of ISPs with a graph $\mathcal{G} = \{\mathcal{I}, E\}$. Each vertex of $\mathcal{G}$ corresponds to an ISP and there is an edge $\{i, i'\}$ between vertexes i and i' if the corresponding ISPs have a settlement-free peering agreement. We denote the minimum node degree in $\mathcal{G}$ by δ, and use the notation $\mathcal{P}(i)$ for the set of neighbors of i. We do not model multiple links connecting two ISPs and we assume that peering capacities are sufficient to carry all relayed traffic, hence we do not consider link capacity constraints.

Let us denote the set of available p2p streaming channels by $\mathcal{H}$, $|\mathcal{H}| = H$. A channel $h \in \mathcal{H}$ is characterized by its bitrate B_h, by its popularity (the number of peers watching the channel) N_h, and by its popularity in ISP i (the number of peers in ISP i watching the channel) N_h^i. We denote the IP transit traffic generated by channel h in ISP i at time t without relaying by $S_h^i(t) = f(N_h^i(t), N_h(t), B_h(t))$, and we make the assumption that it is proportional to $N_h^i(t)$.

III. Application layer settlement-free peering scheme

The architecture we propose is built on relay nodes deployed at the ISPs. Each relay node has a number of relay resources, and each resource can be used to relay one streaming channel to an arbitrary number of peers in a p2p fashion. The relay node monitors the number of peers that watch selected streaming channels in its ISP's network, and joins a channel with one of its relay resources as a regular peer once the channel becomes popular enough. At this time, the relay resource starts to act as the source of a relayed p2p streaming overlay for the channel it just joined. The members of the relayed overlay will be the peers that reside in the local ISP's network and possibly in some peering ISPs' networks, as shown in Fig. 1. We discuss the problem of deciding when a relay node should join a channel in Section IV. When the relative popularity of the channel drops, the relay node gradually delegates the peers back to the global streaming overlay and finally leaves the channel.

In the following we discuss three key technical requirements of the proposed scheme and argue that the scheme is technically feasible.

A. Monitoring of the popularity of channels

Commercial p2p streaming systems use (eventually multiple) trackers for every channel, and the addresses of the relay nodes can be known to these trackers. A relay node acts as an additional tracker for the peers belonging to the local ISP: it maintains a list of the channels and the peers participating in the channels. Peers periodically update their membership information just like they do with the tracker of the streaming provider. Since the number of ISPs (and hence relay nodes) is fairly low, this solution is feasible.

B. Prediction of the popularity of channels

Caching strategies for p2p file sharing were discussed, for example, in [5], and the techniques used there can be used for the prediction of content popularity for p2p file sharing. We are not aware of any work on predicting live streaming channel popularity, but we believe that the diurnal variation [2] of channel popularities and viewing statistics from content providers can help in the prediction.

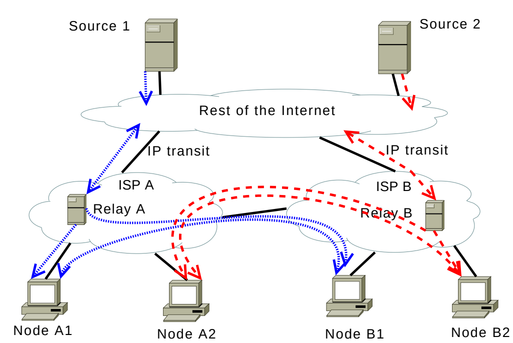

Fig. 1. Application layer peering architecture. Peers in both ISPs are served by relay node A of ISP A for Channel 1 and by relay node B of ISP B for Channel 2 respectively. The relay nodes are members of the respective global p2p streaming overlays. Both ISPs save on IP transit traffic.

C. Handover between overlays

Relaying requires that nodes be handed over from one (possibly relayed) p2p streaming overlay to another relayed overlay and vice versa. The handover has to be smooth, but based on the flash crowd scenario with over 800 thousand peers recorded in [2], we estimate that an arrival or departure rate of up to hundreds of peers per second can easily be handled with current commercial systems. Consequently, the smooth handover of thousands of peers can be done in a few minutes.

IV. Relaying strategies

In this section we are interested in what relaying strategies the ISPs should follow. Let us denote the number of relay resources installed at ISP i by K_i. We describe the relaying strategy of ISP i at time t by a real-valued function $r^h_{i,i'}(t) : \mathcal{H} \times I^2 \times \mathbb{R} \to [0,1]$. $r^h_{i,i'}(t) = 1$ corresponds to channel h being relayed to i' at time t and $r^h_{i,i'}(t) = 0$ to it not being relayed. $0 < r^h_{i,i'}(t) < 1$ corresponds to building up or tearing down the relaying of the channel. The speed at which relaying can be built up or torn down is application specific and is a function of the number of users involved in relaying in the case of peer-to-peer streaming (it would be a function of the file size in the case of file sharing). We will denote the maximum speed (in terms of number of peers per time unit) of building up relaying and tearing down relaying by dr^+ and dr^- respectively. We make the natural assumption that a relaying resource can only be used to relay one channel at a time, even when relaying is built up or torn down.

Furthermore, we define the set of relayed channels in ISP i at time t as $\rho_i(t) = \{h | \sum_{i' \in I} \lfloor r^h_{i,i'}(t) \rfloor + \lfloor r^h_{i',i}(t) \rfloor > 0\}$, i.e., the channels that are relayed by ISP i itself and the ones that are relayed to it by one of its peering ISPs. We define the channel ranks $R^i_h(t)$ to reflect the ordering of the channels with respect to $S^i_h(t)$ within ISP i, e.g., $R^i_h(t) = 1$ for the channel with highest $S^i_h(t)$. Without loss of generality we use $t = 0$ to denote the time instance when a relaying decision has to be made.

Using these definitions we can formalize the problem of cooperative relaying (CRP) as a constraint optimization problem

$$
\begin{aligned}
\text{min.} \quad & \sum_{i \in I} \sum_{h \in \mathcal{H}} \int_0^\infty S^i_h(t)(1 - r^h_{i,i}(t) - \sum_{i' \in \mathcal{P}(i)} r^h_{i',i}(t))dt \\
\text{s.t.:} \quad & \sum_{h \in \mathcal{H}} \lceil r^h_{i,i}(t) \rceil \le K_i \qquad i \in I \\
& r^h_{i,i}(t) + \sum_{i' \in \mathcal{P}(i)} r^h_{i',i}(t) \le 1 \qquad i \in I,\ h \in \mathcal{H} \\
& dr^- < N^h_{i'}(t) \frac{\partial}{\partial t} r^h_{i,i'}(t) < dr^+ \qquad i, i' \in I,\ h \in \mathcal{H} \\
& \lceil r^h_{i,i}(t) \rceil \ge \lceil r^h_{i,i'}(t) \rceil \qquad i \in I,\ i' \in \mathcal{P}(i),\ h \in \mathcal{H}
\end{aligned}
$$

The relaying strategies have to be chosen based on the current relaying strategies $r^h_{i,i'}(0)$, the expected evolution of the channel popularities $N^i_h(t)$ and the resulting transit traffic $S^i_h(t)$.

Local relaying: If the relay nodes are restricted to feeding peers within the ISP then the optimal relaying strategy of ISP i is the solution to the optimization problem CRP with the additional constraint

$$r^h_{i,i'}(t) \quad = \quad 0 \qquad\qquad i' \ne i,\ h \in \mathcal{H}.$$

It is easy to see that in the static case, i.e., if the channel ranks $R^i_h(t)$ do not change over time, the *optimal local relaying strategy* (OLR) that minimizes CRP is the one that relays the K_i channels that generate the most transit traffic at any point in time.

OLR strategy: Set $r^h_{i,i}(t) = 1$ for $h \in \rho^{OLR}_i(t) = \{h | R^i_h(t) \le K_i\}$.

The number of relay resources K_i is chosen by ISP i based on the channel popularity distribution such that the marginal saving by installing $K_i + 1$ relay resources would be less than its marginal cost.

Cooperative relaying: Typically, channels have similar ranks at geographically nearby ISPs, so that the relay nodes of nearby ISPs would possibly relay the same channels. Cooperative relaying can make the use of relaying resources more efficient by allowing relaying between ISPs with a bilateral peering agreement between them on the network layer.

Let us call the solution to CRP the *globally optimal cooperative relaying* strategy (OCR). If the ranking of the channels is the same in all ISPs then finding OCR is closely related to finding minimum dominant subsets of I, a well-studied problem in graph theory. In particular, for $K_i = 1$ ($i \in I$) finding OCR is related to finding the domatic number $D_1(\mathcal{G})$ of graph $\mathcal{G}$, i.e., the maximum number of dominant subsets of I [6]. For $K_i = K \ge 1$ ($i \in I$) the problem is known as finding the r-configuration of $\mathcal{G}$, $D_r(\mathcal{G})$, and was studied in [7]. Finding the domatic number and the r-configuration are NP-complete problems in general, and even approximating the domatic number within a factor $ln|I|$ would imply that $NP \subseteq DTIME(|I|^{O(loglog|I|)})$ [6].

There are three key differences between finding the r-configuration of a graph and the resource allocation problem considered in this paper. First, the number of resources K_i does not have to be equal in all ISPs. Second, channel ranks might differ in the ISPs. Third, it is not necessary to relay the same channels to every ISP for optimality even if the channel ranks are the same.

For small graphs and few relay resources per ISP the exact solution of CRP can be feasible, a centralized solution is however not suitable for the scenario considered in this paper as there is no authority that could enforce relaying strategies to ISPs. Several distributed algorithms have been proposed to approximate $D_1(\mathcal{G})$ [6], [8], [9], but they require information about the two-hop neighbors of the nodes and they do not generalize to arbitrary inhomogeneous relaying capacities and reconfigurations.

Instead, we are looking for an algorithm that (i) respects the confidentiality of the peering agreements between ISPs, (ii) gives incentives to selfish ISPs for cooperation, (iii) makes few changes to the relaying strategies when the channels' ranks change, (iv) and maintains relaying balanced between ISPs.

First, we define the actual relay balance between ISPs i and i' as

$$\gamma_{i,i'} = \sum_{h\in\mathcal{H}} B_h \int_{-\infty}^{0} r^h_{i,i'}(t)N^h_{i'}(t) - r^h_{i',i}(t)N^h_i(t)dt. \quad (1)$$

Let us call a relaying resource of ISP i *allocable* if (i) it is not relaying any channel, (ii) it only relays within ISP i (iii) it relays a channel h for which $R^i_h(t) > |\rho_i(t)|$. We define the interest group of ISP i for channel h as the sum of $N^h_{i'}$ for all ISPs $i' \in \{i \cup \mathcal{P}(i)\}$ and $r^h_{i',i''}(t) = 0$, $r^h_{i'',i'}(t) = 0$ for $i'' \neq i'$.

The greedy cooperative relaying (GCR) strategy that we propose in the following, works by always allocating available relaying resources to the biggest possible group of peers, and can be implemented in a distributed fashion using a voting scheme.

GCR strategy: For a channel $h \notin \rho_i(t)$ and $R^i_h(t) \leq |\rho_i(t)| + K_i - \sum_{h\in\mathcal{H}} r^h_{i,i}(t)$ ISP i looks for ISP $i' \in \{i \cup \mathcal{P}(i)\}$ with an allocable resource and with a maximal interest group. If such an ISP exists (could be i as well), it is chosen as the relaying ISP (in the case of a tie the ISP with higher relay balance is chosen), otherwise the channel will not be relayed.

An ISP i does not start to relay any channel to any ISP i' for which $\gamma_{i,i'}$ is above its threshold.

In order to be able to follow the GCR strategy, each ISP has to be informed about the number of peers watching the channels in its neighboring ISPs.

V. Performance evaluation

In the following we evaluate the gain of cooperative relaying compared to local relaying. We quantify the gain of cooperation in terms of the number of relayed channels instead of in terms of transit traffic. We believe that this metric makes the results easier to interpret. Without loss of generality we limit ourselves to the evaluation of relaying strategies on a set of ISPs I connected by peering agreements, i.e., $\mathcal{G}$ is a connected graph. We define the peering gain for ISP i as $PG_i = |\rho_i|/K_i$, and the mean peering gain as $PG = \frac{1}{|I|}\sum_{i\in I} PG_i$. We focus on the case when the channel ranks $R^i_h(t) = R^{i'}_h(t)$ for all $i, i' \in I$. We argue that this assumption is likely to be valid for ISPs with settlement-free peering agreements as they are typically within the same country or region.

The number of channels that can be relayed within any ISP satisfies

$$|\rho_i(t)| \leq K_i + \sum_{i'\in\mathcal{P}(i)} K_{i'}. \quad (2)$$

Hence, the higher the degree of an ISP, the higher the peering gain it can expect. The mean peering gain can be bounded based on (2) by $\overline{PG}$,

$$\overline{PG} = 1 + \frac{1}{|I|}\sum_{i\in I}\frac{\sum_{i'\in\mathcal{P}(i)} K_{i'}}{K_i} \geq \frac{1}{|I|}\sum_{i\in I} PG_i = PG. \quad (3)$$

If the number of relaying resources is equal in all ISPs ($K_i = K$) then (2) can be used to give a bound on the number of channels that can be relayed in *every* ISP.

Lemma 1: For an arbitrary connected graph $\mathcal{G}$ the number of channels that are relayed in every ISP is bounded from above

$$d_K(\mathcal{G}) = |\cap_{i\in I}\rho_i| \leq K(\delta + 1). \quad (4)$$

Proof: The number of channels that can be relayed in every ISP is the r-configuration of $\mathcal{G}$. The lemma is then a direct consequence of $D_1(\mathcal{G}) \leq \delta + 1$ [7]. ■

Note however that $\cap_{i\in I}\rho_i \subseteq \cup_{i\in I}\rho_i$, that is, not all channels have to be relayed to all ISPs in the solution to CRP.

We can also obtain a lower bound on the efficiency of cooperative relaying under the OCR strategy.

Lemma 2: For an arbitrary connected graph $\mathcal{G}$ and equal number of relaying resources in the ISPs, the peering gain of every ISP is bounded from below by

$$PG_i \geq 2. \quad (5)$$

Proof: For the r-configuration of a graph $D_r(\mathcal{G}) \geq rD_1(\mathcal{G})$ [7]. Furthermore, for any connected graph $D_1(\mathcal{G}) \geq 2$. The proof of the lemma then follows from the definition of PG_i. ■

Consequently, ISPs can at least double the number of relayed channels and hence eventually halve the IP transit traffic through cooperative relaying compared to local relaying if all of them deploy the same number of relay resources. Alternatively, it is enough for them to install

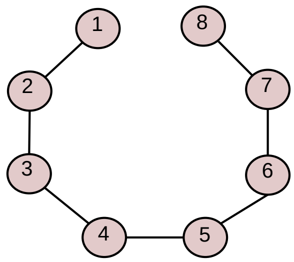

Fig. 2. Graph (1): $\delta = 1$, $D_1(\mathcal{G}) = 2$, $|\rho_i| \leq 3$.

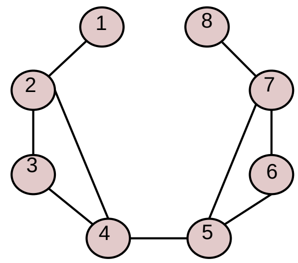

Fig. 3. Graph (2): $\delta = 1$, $D_1(\mathcal{G}) = 2$, $|\rho_i| \leq 4$.

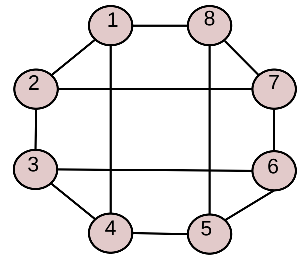

Fig. 4. Graph (3): $\delta = 3$, $D_1(\mathcal{G}) = 4$, $|\rho_i| \leq 4$.

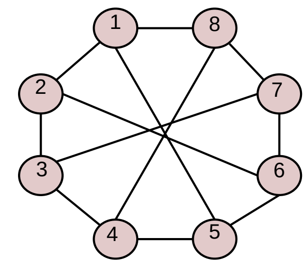

Fig. 5. Graph (4): $\delta = 3$, $D_1(\mathcal{G}) = 2$, $|\rho_i| \leq 3$.

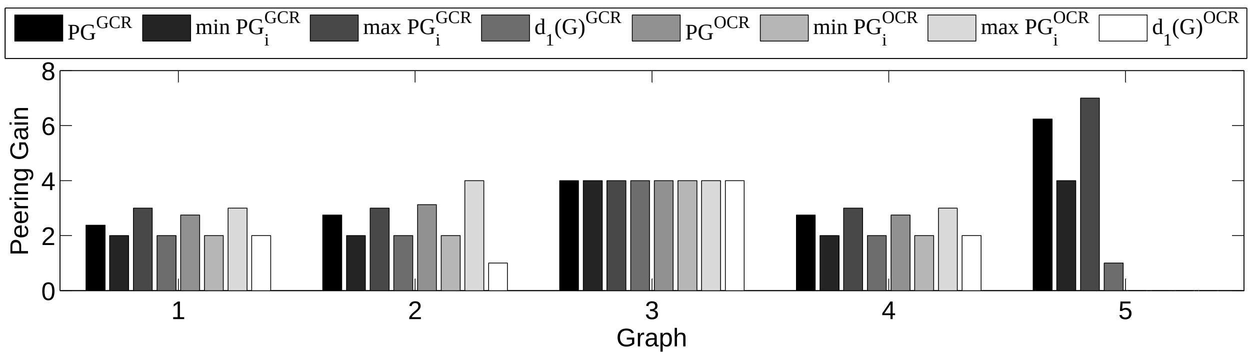

Fig. 6. Peering gain achieved by the GCR and the OCR strategies for $K_i = 1$ on graphs (1)-(5).

half as many relay resources as under the OLR strategy. We will use these bounds as a benchmark in evaluating the efficiency of the proposed GCR strategy.

We evaluate the performance of the proposed GCR strategy on five graphs. Graphs (1)-(4) are shown in Figs. 2 to 5. The number of ISPs is $|I| = 8$ in these graphs, but the graphs have loops of different lengths and differ in their domatic numbers. The 3-regular graph shown in Fig. 4 is domatically full. The 3-regular graph shown in Fig. 5 is however not domatically full, and shows why the constraint on relay balance has to be included in the distributed algorithm: since $D_1(\mathcal{G}) = 2$ but $|\rho_i| \leq 3$, the strategy of every ISP would be not to offer its relaying resource for the two most popular channels but to reserve it for the third most popular one. No ISP can however follow such a strategy for a long time: its neighbors will stop relaying to it due to its insufficient relay balance. The fifth considered graph is shown in [10] and represents the settlement-free peering agreements between the major autonomous systems in Northern Europe (Denmark, Finland, Norway and Sweden). The graph was obtained based on the RIPE *whois* database. It contains 17 nodes, the minimum node degree is $\delta = 4$, consequently, the domatic number of the graph is $D_1(\mathcal{G}) \leq 5$, and for $K = 1$ the maximum number of relayable channels is $max_i|\rho_i| = 16$.

Fig. 6 shows the mean, the minimum and the maximum peering gain, and $d_1(\mathcal{G})$ obtainable with the OCR strategy and with the GCR strategy. For graphs (1)-(4) the number of subscribers is equal in all ISPs, for graph (5) we used an estimate of the number of subscribers of the ISPs [10]. We set $\sum_{h \in \mathcal{H}} N_h^i = 10^4$ and let the channel popularities N_i^h follow a Zipf distribution with parameter $\alpha = 0.7$ [11]. The particular values of the channel popularities do not influence our results, only their proportions. The figure shows that the greedy strategy achieves close to optimal performance in terms of mean, minimum and maximum peering gain on graphs (1)-(4). Note the difference between finding the r-configuration and the CRP on the results for graph (2): the OCR strategies are $\rho_1 = \rho_5 = \{4\}$, $\rho_2 = \rho_7 = \{1\}$, $\rho_3 = \rho_6 = \{2\}$, $\rho_4 = \rho_8 = \{3\}$. That is, only channel 1 is relayed to all ISPs ($d_1(\mathcal{G}) = 1$), even though the graph's domatic number is $D_1(\mathcal{G}) = 2$. Using the GCR strategy, both channels 1 and 2 are relayed to all ISPs, but the mean peering gain is lower than with OCR.

We were not able to calculate the results for the OCR strategy for graph (5) due to the complexity of the solution. We can calculate however the upper bound of the mean peering gain $\overline{PG} = 9.7$, which shows that the GCR strategy performs well: it leads to a gain of 4 to 7 in terms of number of relayed channels, even though there is only one channel (the most popular one) that can be relayed to all ISPs.

Fig. 7 shows the mean peering gain obtained with the GCR strategy for the considered graphs in a dynamically changing environment for one hundred iterations. To model dynamism, at each iteration we swap the ranks of two

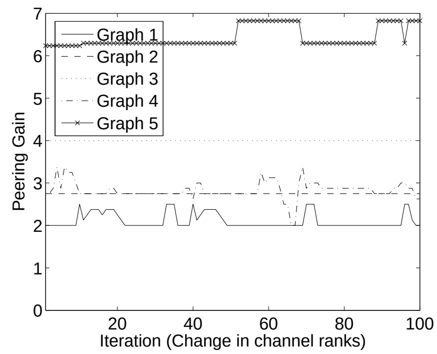

Fig. 7. Mean peering gain achieved by the GCR strategy as channel ranks change for $K_i = 1$ on graphs (1)-(5).

randomly chosen channels with neighboring ranks, and recalculate the relaying strategies using the GCR strategy starting from the previous relaying strategies. The mean peering gain shows modest fluctuations due to the changing channel ranks, and surprisingly, reconfigurations can sometimes even increase the mean peering gain. For the graph (3), for which $d_1(\mathcal{G}) = maxPG_i$, the mean peering gain is not affected by the changing channel ranks, as reconfigurations are not needed at all. It will be subject of future work to test our algorithms on measured traces of channel popularities.

VI. Related work

Cooperative content caching schemes were first considered for HTTP traffic. Hierarchical proxy caching strategies [12] were proposed and evaluated, but the performance metrics, i.e., cache hit ratio and speedup, are different from the ones in the case of p2p systems.

Several measurement studies [3], [5] considered the caching of content for p2p file sharing and showed its possible benefits in decreasing ISP traffic costs. There are also several commercial products that help ISPs to cache p2p traffic, such as PeerCache [4] and the Cache Discovery Protocol [13]. In [14] the authors proposed an application layer protocol that could use existing HTTP caches to decrease the inter-ISP p2p traffic. The scheme requires however fundamental changes in the application layer protocols and it does not consider cooperation between caches.

Finally, content distribution networks (CDNs) like Akamai and CacheLogic offer caching of all kind of content on a commercial basis, but their high costs render them infeasible for many content providers.

To the best of our knowledge, we are the first to propose and to give a mathematical formulation of a collaborative caching scheme that makes use of the settlement-free peering agreements between ISPs in order to minimize the peer-to-peer traffic costs.

VII. Conclusion

In this paper we proposed a cooperative caching scheme that helps ISPs to decrease their bandwidth costs caused by peer-to-peer content distribution systems. We gave a formal description of the resulting resource allocation problem, and gave bounds on the performance of co-operative caching based on results from graph theory. We proposed a greedy distributed algorithm to solve the resource allocation problem, and evaluated its performance on diverse ISP topologies. Our results show that cooperative caching can lead to a significant increase in cache efficiency. It will be subject of future work to evaluate the applicability of results from distributed algorithmic mechanism design [15], e.g., the use of a market-based distributed mechanism.

References

[1] X. Hei, C. Liang, J. Liang, Y. Liu, and K.W. Ross, "A measurement study of a large-scale P2P IPTV system," *IEEE Trans. Multimedia*, vol. 9, no. 8, pp. 1672–1687, 2007.

[2] C. Wu and B. Li, "Characterizing peer-to-peer streaming flows," *IEEE J. Select. Areas Commun.*, vol. 25, no. 9, pp. 1612–1626, 2007.

[3] T. Karagiannis, P. Rodriguez, and K. Papagiannaki, "Should internet service providers fear peer-assisted content distribution?," in *Proc. of Internet Measurement Conference*, 2005, pp. 63–76.

[4] PeerCache, "http://www.joltid.com/index.php/peercache," .

[5] O. Saleh and M. Hefeeda, "Modeling and caching of peer-to-peer traffic," in *Proc. of IEEE ICNP*, 2006, pp. 249–258.

[6] U. Feige, M. M. Halldórsson, and G. Kortsarz, "Approximating the domatic number," *SIAM Journal on Computing*, vol. 32, pp. 172–195, 2003.

[7] M. Fujita, S. amd Yamashita and T. Kameda, "A study on r-configurations - a resource assignment problem on graphs," *SIAM J. Discrete Math.*, vol. 13, pp. 227–254, 2000.

[8] F. Kuhn and R. Wattenhofer, "Constant-time distributed dominating set approximation," in *Proc. of ACM Symposium on Principles of Distributed Computing*, 2003, pp. 25–32.

[9] T. Moscibroda and R. Wattenhofer, "Maximizing the lifetime of dominating sets," in *Proc. of IPDPS*, April 2005.

[10] Gy. Dán, "Caching and relaying strategies for peer-to-peer content delivery," Tech. rep. TRITA-EE 2007:071, School of Electrical Engineering, KTH, November 2007.

[11] K. Sripanidkulchai, B. Maggs, and H. Zhang, "An analysis of live streaming workloads on the Internet," in *Proc. of ACM IMC*, 2004, pp. 41–54.

[12] S.G. Dykes and K.A. Robbins, "A viability analysis of cooperative proxy caching," in *Proc. of IEEE INFOCOM*, 2001, pp. 1205–1214.

[13] CacheLogic, "http://www.cachelogic.com," .

[14] G. Shen, Y. Wang, Y. Xiong, B. Y. Zhao, and Z. Zhang, "HPTP: Relieving the tension between ISPs and P2P," in *Proc. of IPTPS*, Feb. 2007.

[15] N. Nisan and A. Ronen, "Algorithmic mechanism design," *Games and Economic Behavior*, vol. 35, pp. 166–196, 2001.