Taking Advantage of Social Network Relationships in P2P Streaming Overlays

M.L. Merani, D. Saladino and M. Capetta Department of Information Engineering University of Modena and Reggio Emilia, Italy e-mail: marialuisa.merani@unimore.it, marco.capetta.85@gmail.com, daniela.saladino@unimore.it

Abstract—In recent years, classical Internet applications have been accompanied by the surging of a great variety of new services and exciting possibilities. Among such broad range, two particular phenomena are highly successful: Online Social Networks (OSNs) and Peer-to-Peer (P2P) services.

This paper merges these distinct worlds, via the proposal of a P2P streaming system that takes advantage of the friendship relationships of an underlying OSN, to better distribute videos among the overlay peers that are also friends within the OSN. A category of privileged users is therefore created, that is guaranteed a satisfying viewing experience when the P2P overlay operates in critical conditions, i.e, when bandwidth availability is scarce. We show that the help of direct friends, two-hops away friends and, in the limit, of the entire OSN community brings in considerable advantages to the peers that are OSN members. In particular: the number of those among them that are able to download the entire video significantly increases; the number of video portions they can obtain consistently raises; as desired, when the P2P system is operating in underloaded conditions, a proper functioning is guaranteed to all of its nodes, regardless of their being members of the OSN or plain P2P users.

I. INTRODUCTION AND RELATED WORK

Online Social Networks and P2P are both tools contributing to the way people are using and approaching the Internet today. The first pursues the target to bring, inside the network, social relationships like friendship or professional acquaintances, that constitute a significant fraction of our everyday life. In this manner people, regardless of being close or far away, may stay in touch and keep alive their social connections. The second, instead, aims at effectively providing networked services – such as file distribution, video streaming – via resource sharing, where resource means bandwidth, processing power, memory of the network users.

Although OSN and P2P were born and have evolved independently, very recently a few proposals have arisen in academia, that aim at merging some of their features, taking advantage of the strengths of both. Indeed, if P2P allows to better spread contents, reducing the server stress, OSNs can, e.g., greatly ease the search for content, taking advantage of the similarities in the personal taste of connected users.

In [1] Graffi et al. have put forth an OSN where the contents shared among users (comments, tags, photos) are not saved on a single server, but stored in a distributed manner, in a style that reminds the P2P approach. Other works have analyzed

various setup policies for the resources that the peers share, based upon the peers' relationships within an OSN as security assurance. In particular, in [2] Ross et al. propose a new P2P trading paradigm for OSN that, using a credit based exchange among friend users, solves the problems of asynchronicity over time and over nodes of traditional synchronous bilateral P2P trading. In [3] Antoniadis and La Grand encourage the contribution of low-level resources using incentives generated at higher (social) layer to provide more cooperation among peers. The same approach is followed by Hales and Arteconi in [4] and by Lin et al. in [5]. TRIBLER, the social-based P2P system proposed by Pouwelse et al. in [6], uses the OSN relationships as the base layer of a P2P system, not only for content discovery or recommendation, but also to improve download performance. The improvement is achieved thanks to the cooperative downloading implemented by the users that join the same OSN groups, where members who trust each other cooperate.

However, most of the previous works focus on P2P for file-sharing and there is very little work that merges the OSN friendship concept with P2P streaming. Only the concise contribution of [7] represents a first attempt in this direction. Our work moves along a similar path: the idea is to propose a P2P streaming architecture that exploits the social relationships within an OSN as the preferred links to retrieve content. The goal is to implement a delivery mechanism that privileges OSN peers whenever critical conditions start building up within the P2P overlay, i.e., when the overall bandwidth is scarce. In detail, when the system operates in an overloaded condition, an OSN peer requesting video contents and not finding any, asks for the help of direct friends, that discard non-OSN peers currently served to make room for their mate. It is indeed quite reasonable to expect that friends within the social network are willing to help each other in the process of gathering the desired content. Several variants to this priority concept implemented in favor of OSN friend peers are examined. Priority is first extended to friends of friends, a choice motivated by the observation that mutual friendship is a strong bond, that can lead to personal information sharing, as shown in [8], [9] and [10]. As an extreme case, the strategy where priority is granted to all other OSN members, regardless of direct friendship, is also investigated.

The first contribution of our study is to identify and properly amend a model to generate the graph representing the OSN members and their mutual relationships. The second consists in demonstrating that the proposed strategies achieve a clear separation in performance between users belonging to the OSN, that will always be guaranteed a good probability to correctly fetch the desired video or a significant fraction of it, and those outside the OSN, that will to some greater or lesser extent be penalized.

The organization of the paper reflects the points delineated above. Section I illustrates the proposal; Section II describes: (i) the model used to capture the most salient features of the OSN for the sharing of pictures and videos that we took as a reference; (ii) the amendments introduced to it, to better fit the experimental data; Section III details the way the newly proposed P2P system operates; Section IV numerically quantifies the performance achieved and Section V reports our concluding remarks.

II. THE PROPOSAL

The majority of current P2P systems treat their users as anonymous and uncorrelated entities, neglecting any social connections among them. To fill this gap, this work introduces a novel, OSN-based P2P streaming overlay, where social relationships are exploited to develop a privileged video content distribution mechanism among peers that are also OSN members.

The focus is on a specific operating condition of the P2P overlay, namely, that of overloading, where the resource index σ [11] the system displays is lower than one: if N active nodes are present within the system, divided into M classes, each with a specific value of upload capacity C_j and an occurrence probability p_j , j = 1, 2..., M, given the video server exhibits an upload capacity S of its own and the desired streaming rate is d, then the resource index σ is defined as the ratio between the overall upload capacity and the overall rate that system nodes would require to successfully download the video, i.e., as

$$\sigma = \frac{\sum_{i=1}^{M} C_j + S}{N \cdot d} \,. \tag{1}$$

When $\sigma \leq 1$, a critical condition that can often occur, an ordinary P2P overlay cannot provide all of its members an adequate viewing experience. In this circumstance, the functioning of the new system is simple: peers that do not belong to the OSN exploit the video distribution mechanism in the usual manner, hence are subject to a degraded service; on the contrary, peers that are OSN members have an advantage, as they can request their friends' help to retrieve the video.

In detail, in the first proposed strategy an OSN member that newly joins the P2P overlay and meets difficulty in finding portions of the video, is allowed to contact those among its OSN friends that fall in the list of the potential parent peers, and to ask for their help. Upon receiving such request, the contacted peers search among their children, looking for a peer that does not belong to the OSN, and if they find one, its connection is discarded, to make room for the mate in need for content.

For the second proposed strategy, the request for help of an OSN peer unable to fetch the desired video stream is extended to friends of friends within the OSN, i.e., to nodes that are up to two-hops away from it on the OSN graph: any of them, once contacted, tries to fulfill the received request applying the mechanism described above.

Finally and in the limit, the third strategy proposes all OSN members can help any other OSN node, when the latter runs into difficulty retrieving the video.

In essence, OSN members that are also active P2P nodes implement a preemptive priority mechanism in serving the new connection requests, favoring a specific subset of other OSN users.

The Numerical Results Section will illustrate what benefits different policies can achieve and what prices are paid by non privileged users. However, before commenting upon these points, it is mandatory to describe in detail the OSN model we resorted to and its interaction with the P2P overlay.

III. THE ONLINE SOCIAL NETWORK

A. Network Evolution Model

In order to create the OSN graph, whose vertices represent members of the OSN and edges the relationships between them, we took as our starting point the general model of network evolution developed in [12], that also reports data collected from four distinct OSNs: Flickr, LinkedIn, Delicious and Yahoo Answers. In essence, the model the authors put forth in [12] relies on three distinct processes:

- the node arrival process, that is, the process ruling the arrival of new users joining the social network;
- the process of initiating new edges within the OSN, that details the behavior of an active node, statistically specifying the times when the node creates new edges;
- the edge destination selection process, that indicates what node will be the destination of a newly created edge.

As our interest lies in videos, we concentrated on Flickr [13], a social network that was born to share pictures and that later on was extended to allow the sharing of videos: we therefore resorted to the specific statistics and parameter setting summarized below, that according to [12] tailor the OSN model to the features Flickr exhibits.

Namely, given N(t) indicates the number of nodes within the network at a generic time t of its evolution, its behavior was suggested to be exponential, with

$$N(t) = e^{0.25t} \,. \tag{2}$$

As for the active lifetime of a node, a, this obeys an exponential probability density function,

$$p_l(a) = \lambda \cdot e^{-\lambda a} \,, \tag{3}$$

with $\lambda = 0.0092$.

During its active lifetime, a node adds a new edge every δ time steps, where δ is a temporal gap described by a truncated exponential random variable, whose probability density function is

$$p_g(\delta) = \frac{1}{Z} \cdot \delta^{-\alpha} e^{-\beta d\delta} , \qquad (4)$$

where Z is the normalization constant,

$$Z = \int_0^\infty \delta^{-\alpha} e^{-\beta d\delta} d\delta = \frac{\Gamma(1-\alpha)}{(\beta d)^{1-\alpha}},$$
 (5)

 $\alpha = 0.84$ and $\beta = 0.002$.

Last, the authors of [12] found that a satisfying way to choose the destination edge (hence setting up a mutual relationship between two OSN members) was the simple *randomrandom triangle closing model*: based on it, during the active lifetime of a node, when δ expires, first the originating node *s* picks at random one of its neighbors, say *u*, next it randomly picks one of *u*'s neighboring nodes, say *v*, and finally a new edge is initiated between the origin node *s* and the last chosen node *v*, therefore closing – at random – the (s, u, v) triangle.

B. Amendments to the Model

We actually had to introduce a few amendments to the model, to have it satisfyingly fit the experimental data.

The first modification dealt with N(t) behavior: data show that Flickr would count up to approximately 580.000 nodes in a relatively short time span of observation, 25 months. However, it is impossible to reach such significant size exclusively relying upon (2) (even [12] indicates (2) fits well N(t) only in the long term). We therefore decided to rapidly and forcedly increase the number of OSN users at the beginning of the simulation, to later converge to the steady trend described by the exponential function given in (2). Only in this manner the artificial data would fit the experiments, as Fig.1 shows.



Fig. 1. Experimental and recreated data for N(t)

The exponentially distributed lifetime of the nodes has also been modified: experimental data highlight that there is a relatively high probability that a node, once invited to join Flickr, creates very few edges and then never returns, a condition that the exponential distribution cannot capture. We believe this difference is non negligible and according to the data set reported in [12], we put forth a possible solution to this problem, modifying the exponential p.d.f. in (3) into a mixed one: the latter displays a probability p(0) = 0.55 that a newly invited node has a null lifetime, a probability p(1) = 0.05 that the node lifetime is 1 day and a probability p(2) = 0.05 that the node lifetime is 2 days. As a consequence, (3) is modified into:

$$p_l(a) = p(0) \cdot u(y) + p(1) \cdot u(y-1) + p(2) \cdot u(y-2) + (1-p(0)-p(1)-p(2)) \cdot \lambda \cdot e^{-\lambda a}, \quad (6)$$

where u(y) is an impulse located at the origin. Fig.2 indicates that the recreated data are now very close to the real ones (and actually much closer than those obtained sticking to the suggestion of [12]).



Fig. 2. Experimental and recreated data of the node lifetimes

Last, the truncated exponential shape of (4) also had to be amended. When employing the α and β values suggested by the authors of [12], we found that the values of time gaps δ generated in accordance to (4) are much larger than in reality: this condition does not have to be underestimated, as the longer the node time gap, the lower the number of edges the node creates, hence the lower its degree, i.e., the number of edges incident to the vertex representing the node in the graph. After several tests, β was set equal to 0.02, a choice that guarantees a slope for the network degree distribution of -1.76, very close to the real measured value, -1.74.

IV. P2P STREAMING OVERLAY AND OSN FRIENDS

Now that all tools to generate the OSN graph have been properly set, it is appropriate to thoroughly describe the P2P streaming system functioning.

We begin by observing that the OSN and the P2P architecture lie on two distinct planes, evolving at a completely different time pace. It takes several days, if not months, to build an OSN with a fairly large size; besides, relationships among its members can be considered stable in the medium to long term. It is definitely not so within the P2P overlay, where peers join and leave much more frequently and unpredictably. We therefore choose to create one sufficiently large instance of the examined OSN, whose graph remains static throughout the entire P2P system evolution; on the contrary, the birth and death of peers within the streaming system will occur much more frequently, obeying two distinct exponential distributions. Then, whenever a new user asks to join the P2P overlay, with probability P it will be tagged as belonging to the OSN and randomly associated to a specific node of the OSN graph (whereas with probability 1 - P will exhibit no relationships within it).

As regards the examined P2P architecture, its features reproduce the macroscopic behavior that most current P2P prototypes display. The video to be distributed among peers is divided into m substreams, each with rate d/m, where we recall d is the streaming rate. All m substreams have to be received, in order to guarantee a proper reconstruction of the video. Upon joining the P2P overlay, the tracker server immediately passes the new peer a sufficiently wide list of neighbors, and it is among them that the peer randomly selects its potential parent peers: once these are contacted, if they possess the desired substream and have not exhausted their upload capacity, they start providing the newcomer with video chunks.

We additionally observe that the distribution scheme of the P2P overlay is push-pull: once a parent peer starts delivering video chunks to a child peer, it continues to do so until either the parent leaves the overlay or the child itself departs. Moreover, every peer is forced to provide each of its children with only one substream, to avoid the very likely disruption in video quality that its sudden departure would cause.

When the P2P overlay operates in normal conditions, i.e., it is underloaded so that its resource index σ is greater than 1, the behavior of P2P nodes that are also OSN members (and the quality of service they experience) is by no means different from that of ordinary peers. However, when the system happens to be overloaded ($\sigma < 1$), they become privileged users and one of the three strategies delineated in Section II comes into play.

Next Section will quantify the performance achieved by the different proposed schemes in terms of delivery effectiveness, for both OSN peers and for nodes of the P2P system that do not belong to the OSN.

V. NUMERICAL RESULTS

To evaluate the behavior of the proposed system, we have resorted to simulation. In order to realistically describe the dynamic behavior of the P2P overlay, we relied upon PeerSim, a Java based simulator, properly tailoring to our needs an additional protocol, named Overlay Streaming Distribution Protocol [14]: this mimics a hybrid push-pull, mesh-based streaming protocol devised for real-time content distribution.

Whereas the creation of the OSN graph is halted when its size reaches 50×10^4 nodes, the P2P overlay displays an average of $N = 10 \times 10^4$ peers: this 5 : 1 ratio will be maintained throughout all the numerical evaluations. The average peers' lifetime coincides with the distributed video duration, while their average interarrival time is set so as to guarantee that, after a short initial transient, the population size steadily takes on the $N = 10 \times 10^4$ value. Moreover, we initially fix the probability of a peer to belong to the OSN equal to P = 0.5. In the examined framework, the video to be distributed among the peers is divided into m = 8 substreams, each with rate d/8. The video stream is made of 10^4 chunks, the time unit the simulator adopts coincides with the time required to transmit a video chunk, and the simulation time is equal to the total video stream duration.

According to one of the scenarios suggested in [15], we have considered a heterogeneous population of nodes: 30% of the peers are residential, with an upload capacity of 0.5d, 60% are institutional, with an upload capacity of 1.2d, while the remaining 10% are free riders, i.e., their upload capacity is null. As for the streaming server, its upload bandwidth S is three times the streaming rate d. This configuration corresponds to an overloaded condition, where the system is stressed and the capacity to fulfill all peers' download requirement is scarce: indeed, it is in this condition that being an OSN member has to make a difference, introducing an advantage in retrieving the desired video stream.

A. Peers Receiving the Entire Video

The first simulation outcomes report the percentage of peers that receive all the *m* substreams, i.e., that are able to properly reconstruct and watch the full video, as a function of the time the video is distributed within the overlay, when the different priority strategies described earlier are considered. Each viewgraph displays two curves: the solid line refers to the performance experienced by peers that are also OSN members, that we have termed *inSN* peers; the dashed line refers to peers that do not belong to the OSN, the so-called *outSN* peers. To present more accurate results, every point of the two curves was obtained by averaging the results of 5 independent simulation runs.

When the help of direct friends is considered, only a slight performance differentiation between inSN and outSN peers occurs, as Fig.3(a) shows: this had to be expected, as (i) direct friends may not be that many; (ii) not all of them belong to the list of peers that can be contacted by an inSN newcomer looking for help; (iii) they do not necessarily possess the requested substream. The advantage in favor of inSN peers becomes more evident extending the preemptive priority concept to friends of friends: Fig.3(b) points out inSN peers enjoy a significantly higher probability to receive all substreams, whereas outSN peers are penalized, as they sometimes get discarded in favor of inSN peers. The gap in performance further increases when the preemptive priority mechanism that ISN peers adopt favors the totality of inSN members, as shown in Fig.3(c): now inSN nodes experience a very high probability of correctly receiving the entire video stream, at the expense of outSN users.

B. Statistics of the Received Stripes

Next results refer to the Cumulative Distribution Function (CDF) of the number of substreams that *inSN* and *outSN* peers



Fig. 3. Percentages of peers that receive all stripes

receive. The curves in Fig.4(a) refer to *inSN* peers, whereas the curves in Fig.4(b) refer to *outSN* users. In these figures there appear four CDFs, identified by different symbols: squares refer to the behavior without any help, circles refer to the system that grants priority to direct friends only, diamonds to the two-hops away friends help and triangles to the condition where the preemptive priority mechanism is implemented among all *inSN* users. Both graphs consider the system state as observed at the end of the simulation and as before, are



Fig. 4. CDF of the received number of stripes

obtained from the average of 5 distinct simulations.

Ranging from the original configuration where no help among OSN members is foreseen, until the last priority method, the effects of different priority mechanisms are evident. In the original system with no priority, every *inSN* and *outSN* peer, acting selfishly, receives very few substreams and is therefore subject to a poor viewing experience. The *inSN* members experience a steady increase in the number of received stripes the more OSN users can assist them, succeeding with high probability in receiving 7 or 8 stripes; on the contrary, *outSN* peers get fewer and fewer stripes, quite often fetching only 6, 5 or even 4 stripes. As desired, there are many *inSN* peers aiding friends, friends of friends or ultimately all OSN members, always at the expense of *outSN* nodes.

Last, it is worth pointing out that the presented results are subject to a further interpretation. If the *m* video substreams are obtained via an encoding technique such as Multiple Description Coding (MDC), so that the more substreams are retrieved, the higher the quality of the reconstructed video, then in an overloaded system *inSN* peers would be able to watch a good quality video, whereas *outSN* peers would experience a basic viewing quality, without being locked out.

C. Impact of OSN Members Percentage

We next investigate the effect that a different percentage of *inSN* peers over the total P2P overlay population size has on system performance, hence consider different values of the probability P introduced at the end of Section IV.

The following results refer to the strategy where the help with preemptive priority is implemented among peers that are either direct friends or lie two-hops away (friends of friends) within the OSN graph. Fig.5 reports the CDF for the number of stripes received by inSN peers: the five curves summarize system behavior for different P values, namely, P = 0, 0.2, 0.5, 0.8 and 1. Here too, the various CDFs are determined averaging the data collected at the end of 5 simulations. The Figure points out a modest degradation of *inSN* performance with the increase of the P value, from P = 0.2 to P = 0.8. This effect has to be ascribed to the growing fraction of inSN peers: if they represent a significant portion of the entire overlay population, then there is a decreasingly lower possibility that either a friend or a friend of a friend can discard *outSN* nodes. The CDF obtained for P = 1 actually corresponds to the original system, where no differentiation among inSN and outSN users is present (the totality of the peers are also OSN members), whereas for P = 0 there is no possibility to determine a CDF, inSN peers being absent.



Fig. 5. CDF of the received number of stripes by inSN peers for different P values

Although not reported in this paper, we also investigated the effects that different priority strategies have on the delay *inSN* and *outSN* peers experience in retrieving the video stream. Here too, the introduction of a priority mechanism is beneficial: the CDF of the delay faced by *inSN* peers has a lower average and a smaller variance than the CDF referring to *outSN* nodes, with an increasing difference the more OSN members are involved in providing their help. Finally, we verified that when the P2P system operates in underloaded conditions, the newly introduced mechanisms do not modify its behavior: as desired, no differentiation among *inSN* and *outSN* peers is present, since there is enough overall bandwidth to fulfill all peers' download requests.

VI. CONCLUDING REMARKS

This paper has proposed a new P2P architecture that leverages the existing relationships among peers that are also members of an OSN, in order to guarantee such users a better viewing experience when critical conditions start building up within the P2P overlay. Different strategies to help OSN members in retrieving the desired content have been investigated, and their achievable performance assessed, focusing on a P2P system that operates in an overloaded condition, due to the scarcity of the upload bandwidth its members make available to the community. When video streams are distributed in a preemptive priority fashion to OSN friends and friends of friends, it has been shown that OSN peers experience a very good probability of flawlessly receiving the entire video or a significant portion of it, achieving a clear service differentiation with respect to peers that do not belong to the OSN.

REFERENCES

- K. Graffi, S. Podrajanski, P. Mukherjee, A. Kovacevic, and R. Steinmetz, "A Distributed Platform for Multimedia Communities", Tenth IEEE International Symposium on Multimedia, ISM 2008, 15-17 Dec. 2008, pp.208-213.
 Z. Liu, H. Hu, Y. Liu, K.W. Ross, Y. Wang and M. Mobius, "P2P
- [2] Z. Liu, H. Hu, Y. Liu, K.W. Ross, Y. Wang and M. Mobius, "P2P Trading in Online Social Networks: The Value of Staying Connected", Proceedings of IEEE INFOCOM 2010.
- [3] P. Antoniadis and B. Le Grand, "Self-organized Virtual Communities: Bridging the Gap between Web-based Communities and P2P Systems", International Journal of Web Based Communities, Vol.5, Issue 2, March 2009, pp.179-194.
- [4] D. Hales and S. Arteconi, "Friends for Free: Self-Organizing Artificial Social Networks for Trust and Cooperation", in Algorithmic Aspects of Large and Complex Networks, 2005.
- [5] W.S.Lin, H.V. Zhao and K.J.R. Liu, "Incentive Cooperation Strategies for Peer-to- Peer Live Multimedia Streaming Social Networks", IEEE Transactions on Multimedia, Vol.11, No. 3, April 2009, pp.396-412.
- [6] J.A. Pouwelse, P. Garbacki, J. Wang, A. Bakker, J. Yang, A. Iosup, D.H.J. Epema, M. Reinders, M.R. van Steen and H.J. Sips, "Tribler: A social-based Peer-to-Peer System", in Recent Advances in Peer-to-Peer Systems and Security (IPTPS 2006), pp.127-138.
- [7] O. Abboud, T. Zinner, E. Lidanski, K. Pussep, R. Steinmetz, "Stream-Social: a P2P Streaming System with Social Incentives", Proceedings of IEEE International Symposium on a World of Wireless Mobile and MultiMedia Networks, Montreal, Canada, June 2010.
- [8] F. Nagle and L. Singh, "Can Friends Be Trusted? Exploring Privacy in Online Social Networks", ASONAM 09. International Conference on Advances in Social Network Analysis and Mining, 20-22 July 2009, pp.312-315.
- [9] A. Acquisiti and R. Gross, "Information Revelation and Privacy in Online Social Networks (the Facebook case)", in Proceedings of the 2005 ACM Workshop on Privacy in the Electronic Society (WPES), November 2005.
- [10] A. Acquisiti and R. Gross, "Imagined Communities: Awareness, Information Sharing, and Privacy on Facebook", 6th Workshop on Privacy Enhancing Technologies (PET), June 2006.
- [11] D. Wu, Y. Liu, K.W. Ross, "Queuing Network Models for Multi-Channel P2P Live Streaming Systems", Proc. of IEEE Infocom 2009, April 2009, pp.73 - 81.
- [12] J. Leskovec, L. Backstrom, R. Kumar, A. Tomkins, "Microscopic Evolution of Social Networks", Proc. of ACM SIGKDD, 2008 International Conference on Knowledge Discovery and Data Mining, August 2008, Las Vegas, Nevada, USA, pp.462-470.
- [13] Flickr, "Flickr Homepage", available at http://www.flickr.com/
- [14] D. Carra, R. Lo Cigno, E. Biersack, "Graph Based Analysis of Mesh Overlay Streaming Systems", IEEE Journal on Selected Areas in Communications, December 2007, Vol.25, no.9, pp.1667-1677.
 [15] Z. Liu, Y. Shen, K. W. Ross, S. S. Panwar, Y. Wang, "LayerP2P:
- [15] Z. Liu, Y. Shen, K. W. Ross, S. S. Panwar, Y. Wang, "LayerP2P: Using Layered Video Chunks in P2P Live Streaming", IEEE Trans. on MultiMedia, Vol.11, No.7, November 2009, pp.1340-1352.