

User Behavior Anticipation in P2P Live Video Streaming Systems Through a Bayesian Network

Ihsan Ullah, Guillaume Doyen, Grégory Bonnet, Dominique Gaïti

► To cite this version:

Ihsan Ullah, Guillaume Doyen, Grégory Bonnet, Dominique Gaïti. User Behavior Anticipation in P2P Live Video Streaming Systems Through a Bayesian Network. Proceedings of the IFIP/IEEE International Symposium on Integrated Network Management, May 2011, France. pp.337-344, 2011. <hal-00951778>

HAL Id: hal-00951778

<https://hal.archives-ouvertes.fr/hal-00951778>

Submitted on 25 Feb 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

User Behavior Anticipation in P2P Live Video Streaming Systems through a Bayesian Network

Ihsan Ullah and Guillaume Doyen and Grégory Bonnet and Dominique Gaiiti

ERA / Institut Charles Delaunay – UMR 6279

Université de Technologie de Troyes

12 rue Marie Curie – 10000 TROYES – France

Email: ihsan.ullah,doyen,bonnet,gaiiti@utt.fr

Abstract—For a few years, Peer-to-Peer (P2P) architectures have emerged as a scalable, low cost and easily deployable solution for live video streaming applications. In these systems, the load of video transmission is distributed over end-hosts by enabling them to relay the content to each other. Since end-hosts are controlled by users, their behavior directly impact the performance of the system. To understand it, massive measurement campaigns covering large-scale systems and long time periods have been performed. In this paper, we gathered and synthesized results obtained through these measurements and propose a Bayesian network that captures and integrates all of them into a synthetic model. We apply this model to the anticipation of peer departures which is an important challenge toward the performance improvement of these systems and especially churn resilience. The validation of our proposal is performed through intensive simulations that consider a streaming system composed of thousand users over two hundred days. We especially study two deployment scenarios: a system-scale one and a local one. We also compare our proposal with two standard estimators and we show under which conditions an estimator outperforms the others.

I. INTRODUCTION

During the past few years P2P video streaming has attracted great attention in both academia and industry. On one hand it does not require any major change in the current Internet infrastructure unlike IP multicast. On the other hand, it pushes the load of content delivery to the end-hosts through sharing their upload bandwidth. It makes these systems scalable as compared to centralized Client/Server (C/S) architecture which requires an increase in the servers' capacities with an increase in the number of users. Moreover, a P2P system is easy to deploy with a low cost.

P2P approach enables end-hosts to self-organize into a virtual network where a peer not only receives the content but also transmits it to other peers. Since these peers are controlled by users, their behavior directly impacts the virtual network and consequently the performance of the system. Therefore, management of P2P streaming networks with an awareness of user behavior becomes critical.

To do so, first an understanding of user behavior and then its modeling are required. For an understanding of user behavior we have gathered, analyzed and synthesized the measurement studies found in the literature over user behavior in video streaming systems. We extracted from them the components of a user behavior, environmental factors that impact these com-

ponents and the network parameters that are impacted by these components. On this basis, we model the relationships among all these elements through a Bayesian network that allows the estimation of a user activity. This Bayesian network enables thus the management of the P2P network with an awareness of user activities and eventually improves the performance. Applications of such a network cover a large spectrum, from the anticipation of system size and churn or the estimation of streaming quality. In this paper, we focus on the estimation of the departure time of peers. The latter is crucial to avoid service disruption in the content delivery and the use of large buffer that induces a significant delay in real-time programs.

In order to validate our model, we simulate two scenarios. In the first scenario, the whole population of users is considered together requiring thus only one Bayesian network for their behavior estimation. In the second scenario, we estimate the activities of individual users which require one Bayesian network for each user present in the P2P system. We intensively analyze the performance of our Bayesian network over these two scenarios. We also compare the performance of the Bayesian network for departure time estimation with two others estimators. The first relies on an exponential moving average and the second on a Bayesian estimation restricted to historical session times. These two alternative estimators do not integrate any contextual information and we show under which conditions the use of a Bayesian network is relevant.

The rest of the paper is organized as follows. In section II, we discuss works related to user behavior modeling. Section III first gives a synthesis of user behavior measurements for identification of user behavior metrics and their relationships. It is followed by the presentation of the user behavior model. Section IV provides the simulation details and results for the validation of our network. Finally, section V gives conclusions and future directions.

II. RELATED WORK

Most of the works on user behavior in live streaming systems are dedicated to its measurements. There are a few others that also model some aspects of user behavior. Here, we only discuss the latter.

Tang *et al.* [14] measure the stability of users in a C/S and a P2P video streaming system and find a positive correlation between the elapsed time of a user and remaining time of a

Reference	Type	System	Period	Method	Metrics					
					OD	SD	CP	A/D	D/U	
[1]	P2P	PPLive	Feb. to Nov. 2008	Crawler					✓	
[2]			Apr. to Dec. 2006			✓	✓			
[3]			Nov. 2006 (about 28 hours)			✓			✓	
[4]			2006 – 07	Passive/ Crawler		✓	✓	✓	✓	
[5]			Jun. 2006	Passive		✓			✓	
[6]		PPStream	unknown	Crawler			✓	✓	✓	
[7]		Zattoo	Mar. 2008 (2 weeks)	Logs		✓			✓	
[8]		Cool- Streaming	Mar. 2005 (4 days)			✓	✓	✓		
[9]			Oct. 2006 (1 day)			✓	✓	✓		
[10]			Sep. 2006 (1 day)			✓	✓	✓	✓	
[11]		unknown	2006 (11 hours)				✓		✓	
[12]		GridMedia	Jan. 2006 (4.5 hours)			✓				
[13]		UUSee	May to Jun. 2008 (5 days)				✓		✓	
[14]		C/S	CCTV		Feb. 2005/Jan. 2006 (2 pop. events)		✓			
[14]					Oct. to Jan. 2004 – 05		✓			
[15]	Akamai		Oct. to Jan. 2003 – 04			✓	✓	✓		
[16]	unknown		2002 (90 days)			✓	✓	✓		
[17]	IPTV		Telco- Managed		Apr. 2008 (6 days)	✓	✓	✓	✓	
[18]			Jun. 2008 (1 month)				✓			
[19]			May to Oct. 2007			✓	✓	✓		

OD: Online Duration; SD: Session Duration; CP: Channel Popularity; A/D: Arrival/Departure rates; D/U: Download/Upload traffic

TABLE I
OVERVIEW OF MEASUREMENT STUDIES OVER LIVE STREAMING SYSTEMS

user watching a particular channel. They propose a model for the selection of an upstream peer in such a way that a peer with longer elapsed time is preferred over others. Wang *et al.* [3] use a similar method for the identification of stable peers. To minimize the impact of churn, they put those peers in the backbone. The method of stable peers identification in both of these approaches tends to assume all the peers that have recently joined the system as unstable which is not always the case.

The above approaches do not consider any external impacting factors on the stability of users. By contrast, Liu *et al.* [13] measure a P2P live streaming system and analyze some potential impacting factors on users' stability and bandwidth contribution ratios. They observe that some of them such as streaming quality impact the stability and bandwidth contribution ratios of peers. Based on this finding they present models for predicting the longevity and bandwidth contribution ratios of peers. However, their models are global for all user behaviors while the behavior of an individual can be different from the global one.

Horovitz *et al.* [20] propose a machine learning approach based on Support Vector Machines (SVM) for detecting actively the load in the uplink of provider peers. This approach only considers upload bandwidth contribution as a metric of the user behavior, while others such as stability of the provider peers are also important for the performance.

In our earlier work [21], we proposed peers' stability estimators based on history of a user's past sessions to estimate his current session duration. Estimators are based on Exponential Moving Average (EMA) and Bayes' rule. Moreover, a proactive mechanism of switching to a new upstream peer before the anticipated departure of the current upstream peer was proposed. The limitation of this work was that it did not take into account other external impacting factors. Therefore,

in [22], we presented a preliminary model of user behavior through a Bayesian network that takes into account all the known impacting factors. The limit of this model was the binomial variables which could not estimate states of metrics with a satisfying granularity.

III. UNDERSTANDING AND MODELING OF USER BEHAVIOR

In this section, we first give a brief synthesis of user behavior measurements that highlights the components of user behavior and their relationships with the external environment and network. Then we present our Bayesian network and explain its structure.

A. Understanding user behavior

To get an understanding of user behavior in video streaming systems, we collect all measurements from the literature and perform their synthesis. An overview of these measurements is shown in Table I. This table summarizes the reference of the measurement, the type and name of the system that was studied, the measurement time period, the methodology used for data collection and the observed metrics. It is obvious from this table that taking only one study into account is not sufficient to model user behavior accurately, because each measurement analyzes a few metrics under different conditions. Moreover, other information such as the dependency relationships of metrics with the environment and among themselves are not shown in the table which are split among different measurements and cannot be found in only one work. A global picture of these relationships is shown in Figure 1. An edge from one node to another shows the impact of one element on the other one. To explain these relationships, we discuss each impacted element with its impacting factors in the following.

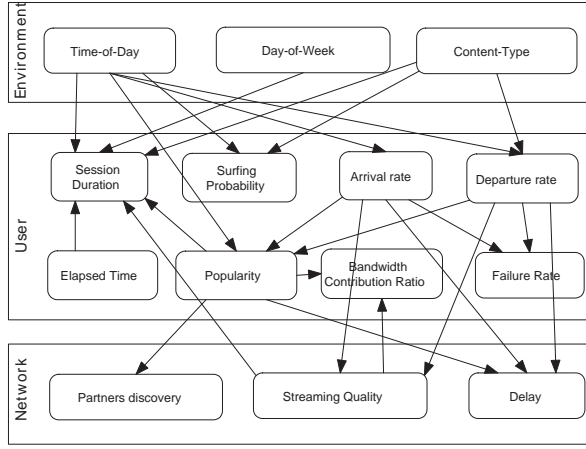


Fig. 1. Components of user behavior and related metrics

1) *Arrivals/departures*: Users arrivals/departures to/from a channel are impacted by the time-of-day and channel type. Studies [6], [4], [19], [15] agree on a higher arrival rate at the beginning of a program and higher departure rate at the end of the program as compared to the middle of a program. Channel switches during commercial breaks have been observed too. Arrival rates are smoother than departure rates because the number of users normally increases gradually at the start of a program, while the departures usually occur in batches towards the end of a program [19]. Since most of the TV programs and commercial breaks are scheduled on fixed times, time has an impact on arrival and departure rates of users onto a channel. Apart from time, content type also impacts the departure rate. Hei *et al.* [4] observe batch departures in a movie channel but they do not find this behavior in another (unspecified) popular type of channel.

2) *Channel popularity*: Since we are only interested in the user behavior at a particular time instance, we only consider the popularity of a channel at that time also called instantaneous popularity. Channel popularity is impacted by time-of-day, arrival rate and departure rate. Qiu *et al.* [17] find diurnal patterns in popularity dynamics. It means that popularity is time dependent and varies during a day. Similarly, it is evident that higher arrival rate than the departure rate increases the instantaneous popularity and vice versa. For BitTorrent system, Guo *et al.* [23] define the popularity of a torrent at a time as the peer arrival rate at that time.

3) *Session duration*: Holding time of a channel, called session duration, is impacted by the elapsed time in a session, streaming quality, popularity, type of content, time-of-day and day-of-week. Studies [14], [19] report a positive correlation between the elapsed time and remaining time of a session. Liu *et al.* [13], [24] find a strong correlation between the initial streaming quality and session duration of a peer. It states that a user receiving a good buffer level joining a channel is willing to stay longer. Popularity of a channel also impacts the session duration. Users stay longer while watching popular programs as compared to unpopular ones [13], [4], [24]. Similarly, the

time spent on each channel changes with the type of the channel. Cha *et al.* [19] observe shorter session durations for news and music channels than documentaries and kids channels. Finally, Liu *et al.* [13] reveal that session duration has a strong correlation with time-of-day but no correlation with day-of-week, but Veloso *et al.* [16] present a contrasting finding that time-of-day does not impact the session duration, while day-of-week has an impact on it.

4) *Other impacted parameters*: Here we discuss those parameters that are less studied than the above-mentioned ones.

- *Surfing probability*: Surfing (channel browsing) probability is impacted by time-of-day and type of content, which increases for less popular channels and specific genres like news and music [19], [17].
- *Failure rate*: The departure of a user from a channel before the video player becomes ready, called failure, is strongly correlated with the overall arrival/departure rates of users [9].
- *Delay*: Playback delay is increased with an increase in popularity and arrival/departure rates [6].
- *Streaming quality*: The quality of streaming is impacted by arrival and departure rates. It degrades under flash crowds and high departure rates [13].
- *Bandwidth contribution ratio*: User bandwidth contribution ratio is found to be strongly correlated with initial streaming quality a user receives and popularity of the channel [13].
- *Partners discovery*: A peer can find partner peers easily while watching a popular channel [8], [2].

B. User behavior modeling

We require a model that can predict the next state of a user behavior metric. For example, whether the user will stay or leave during certain time period or/and what will be the bandwidth contribution ratio of a user during a next time period. This is an online analysis where prediction of future observations is needed given observations up to the present time. Moreover, the problem domain contains different variables with dependencies among them. Bayesian networks are well-suited in modeling such type of domains for a number of reasons. Firstly, they are expressive for modeling real-world problems. Secondly, prior knowledge can be incorporated into the network in the absence of any observation. Finally, they model well the relationships of causes and effects.

1) *Bayesian networks*: Bayesian networks (BN) are graphical models for problems of uncertain reasoning. They are made of two main components namely structure and parameters. The structure is an acyclic graph that encodes variables through nodes. The probabilistic conditional independence relationships among variables are encoded through arcs. A directed arc from one node to another makes them parent and child nodes respectively. A child node is dependent on its parent node. A node is conditionally independent from all other nodes given its parents. The joint probability distribution of variables in a Bayesian network can be written in the form shown in Eq. (1)

$$P(U) = P(X_1, \dots, X_n) = \prod_{i=1}^n p(X_i | \text{parents}(X_i)) \quad (1)$$

The network structure can be constructed by an expert or learned from data. Similarly, parameters can be assigned manually or learned from data. In our work, we construct the network manually and let it learn its parameters from the data.

2) *Proposed model*: Based on the synthesis of measurement studies, the Bayesian network we propose is depicted in Figure 2. It involves 10 nodes each of them representing a component of user behavior, an impacting factor of these components or a network performance metric derived from the graph given in Figure 1.

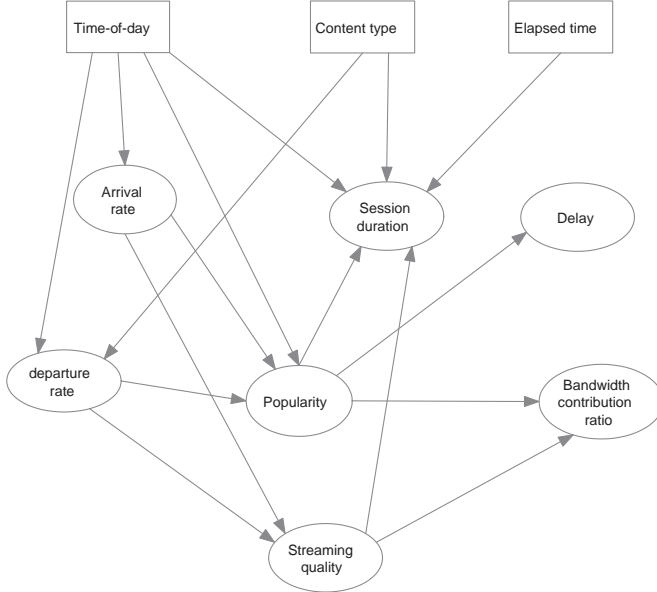


Fig. 2. Our proposed Bayesian network for user behavior

Day-of-week is omitted since it was observed only within two measurements with contradicting results. Surfing probability and failure rate are both merged into session duration because both of them represent short session durations and they can be represented by the session duration variable. Partners discovery has been excluded since it is required only in mesh-based systems and can be easily added if needed.

3) *Node types*: Our Bayesian network is a mixed model having both discrete and continuous variables. Discrete variables have been shown through rectangles and continuous ones through ellipses. In case of discrete variables, the number of states of a variable should also be chosen. Since our workload contains three content types, we choose three states for the content type variable. Similarly, we choose binary states for elapsed time of a user. Elapsed time can determine whether a user is in surfing mode or viewing mode. Cha *et al.* [19] observe that the channel holding time in surfing mode can last up to 2 minutes. Therefore, elapsed time less than 2 minutes makes the first state while the second state is assigned to

greater values. Concerning time-of-day, it can be discretized with several intervals and we discuss it in section IV-B1.

Since user behavior is highly dynamic and an accurate quantification of the dependencies among the variables is not possible, we let our network learn its parameters from data. In a real P2P network, it will be the case where the dependencies among different variables for each user are different and should be learned online.

IV. VALIDATION

In order to validate our proposal, we performed intensive simulations considering different scenarios. Results presented in this section have been obtained through Matlab and the Bayes Net Toolbox¹ extension. The duration of each simulation was about some hours up to thirty for some of them, leading to more than one hundred simulation hours for the whole presented results.

A. Simulation framework

1) *Simulation scenarios*: In order to evaluate the performance of our proposal, we considered two simulations scenarios standing for realistic concrete deployments. In the first one, we consider that the target P2P video streaming system presents a central entity owned by the service provider. This central entity could for example be used to authenticate peers as in the case of Skype. It would thus have a knowledge of all information depicted in our Bayesian network, from joining and leaving operations to the type of watched content and time-of-day. The estimation of session time of each peer is done considering information provided by the whole community.

In the second scenario, we consider a deployment of our Bayesian network directly into peers. In this case, only local information is directly known. These are the joining and leaving operations, the type of content the user watches and time-of-day. Nonetheless, external information, resulting from a community-scale aggregation, must be taken into account in order to complete the inference on the Bayesian network. The latter are the arrival and departure rates, and the population. We consider that such indicators can be easily obtained through the use of a decentralized aggregation protocol based on a tree structure or a gossiping approach and thus can be known by every peer with an acceptable accuracy.

For the two presented scenarios, simulation parameters are given in Table II.

2) *Workload generation model*: As a global model is inadequate to model a population of users with different behaviors, we inspired from [25] that defined different representative and fictional television watchers with different preferences and habits, called personas². Each single user is associated with a persona. It is controlled by a non-homogeneous semi-markovian process whose parameters are set according to the given persona. It means that the state of a user at a given time not only follows the markovian assumption (dependance on

¹<http://code.google.com/p/bnt/>

²This work is in the process of publication

Parameter	Value
Simulation duration	40 days (global scenario) 200 days (local scenario)
Total population	1000 peers
Content type	Reality, fiction and sport
Content duration	2 hours
Inference algorithm	Conditional Gaussian
Learning algorithm	Maximum likelihood learning
Time-of-day discretization interval	1 hour (see Section IV-B1)

TABLE II
SIMULATION PARAMETERS

the state at the previous time) but also on the time it spends in this state and on the global time of the process. Such kind of process fits well with the video streaming context because, as explained in Section III-A, the behavior of a user vary with respect to the time of the day (e.g. watching longer or more often in the evening than in the morning) and the time spent online impacts the time it will stay.

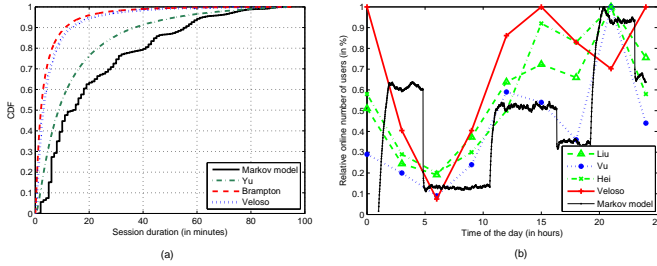


Fig. 3. User simulations: (a) CDF of session durations; (b) population evolution

Our Markov process owns the following features:

- 1) We only consider mono-channel video streaming applications. Consequently, we define two states $\{X_1, X_2\}$ where their semantic is respectively the user's online presence and its offline presence;
- 2) We consider that the user behavior is a cyclic behavior on each day as highlighted by [13], [19], [2]. Consequently, we consider the process global time $t \in \mathbb{N}^+$ as a day discretized in one-minute intervals, $\{t_1 \dots t_{1440}\}$;
- 3) In order to be consistent with the literature [26], [27] and as log-normal laws are commonly used to model slow fading phenomenon, we consider that the transition from state X_1 to state X_2 is controlled by a log-normal law. As an individual user watches longer in certain time-of-the day, the $\mu_t^i(d)$ and $\sigma_t^i(d)$ depend on the persona i , on the kind of content d , on the global time of the process t and on t_{X_1} the time spent in the state X_1 ;
- 4) As Poisson laws are widely used to model arrival processes, we consider that the transition from state X_2 to state X_1 is controlled by a Poisson law which is consistent with [26]. As an individual user has habits about his watching time, the parameter λ_t^i depends on the persona i and on the global time of the process t .

In order to validate the global behavior of a population of users, we simulate three days of different content types with 10000 users dispatched among the six persona. Figure 3.a compares the cumulative distribution function (CDF) of the session duration of our model with the global models proposed by [26], [27], [16]. We can notice that our model fits well with the global model of [26] and has the same shape as the others. Figure 3.b represents the population evolution through the day. We can notice that our model is consistent with the traces presented in [2], [13], [16]. The sudden changes in the population with our model represents flash crowds which is a common phenomenon in P2P streaming systems. Thus, these results show that our model is consistent with results obtained through measurements on real systems.

B. Results analysis

1) *Network setup*: In Section III-B3, we presented and justified the nature of each network variable that can be binomial, multi-nomial or continuous. For each of them, its nature has been determined at the network design stage. Nonetheless, for the *Time-of-day* one, this determination is not possible. Although it is obvious that it has to be a discrete one, the discretization period cannot be chosen directly. Indeed, the choice of a discretization period has a strong impact on the network performance since it induces a tradeoff between the accuracy of session time estimations and the convergence time of the learning process: a shorter discretization period will induce a better accuracy while increasing the learning time and vice versa.

In order to choose the most appropriate value of the discretization interval, we perform tests with different intervals. We use the same dataset for all networks. All the observations are provided to the network one by one and for each one an inference test is carried out following by a learning step. In the inference test, variable session duration is kept hidden and all the other observations are given to the network as evidences. The network estimates the session duration under the given conditions. Then the complete observation is provided to the network for updating it beliefs. At the end, we compare actual session durations with estimated ones measure: (1) the Root Mean Square Error (RMSE) of the session time estimations once the learning process is completed and (2) the learning time. The latter is determined by measuring the end of the learning process, given by the time at which the variation of RMSE does not go over a given threshold, set at one minute in our case.

The results we obtained are different for the two deployment scenarios. In case of a global deployment, it appears that the large number of sessions due to the considered population induces an extremely short learning time, hardly measurable, while the RMSE only varies between 17.33 and 17.69 minutes making the discretization period invariant regarding the global network performance. By contrast, in the case of a local deployment, the discretization period presents an important impact. Figure 4 depicts the obtained results for the learning time and RMSE considering different values of

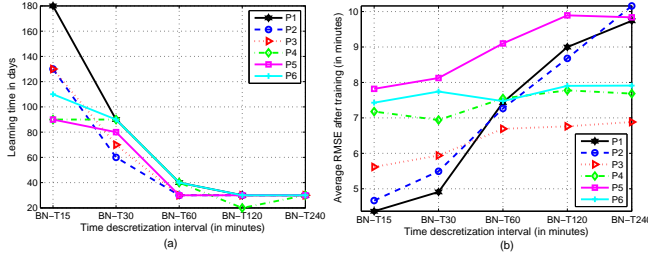


Fig. 4. Learning time (a) and error (b) generated by networks with different discretization intervals. The number after BN-T indicates the discretization interval in minutes and 'P' represents persona.

the discretization period. Whatever the considered persona, results are consistent since they exhibit similar shapes. The best tradeoff value appears as being 1 hour, giving thus an average learning time of about 35 days and an RMSE of about 7.6 minutes. We use this 1 minute interval in all simulations presented hereafter.

2) *Global case*: In order to understand the accuracy of the Bayesian network, we depict in Figure 5.a estimated session durations versus the actual ones over a ten days period. For actual sessions comprised between 0 and 40 minutes, one can see that the Bayesian network over- and under-estimates the same amount of sessions. Moreover, estimated sessions uniformly occupy the $[0 - 40] \times [0 - 40]$ square indicating a poor accuracy. Then, for actual sessions comprised between 40 and 100 minutes, the Bayesian network clearly under-estimates sessions. Although under-estimating a session time does not prevent from anticipating the user behavior, it makes it totally inaccurate, thus inducing a strong overhead for any networking mechanism that rely on it. Such results can be explained by the fact that the shuffle of all individual behaviors leads to an average behavior in which it is hardly feasible to accurately distinguish the elements that induce larger or shorter session durations.

3) *Local case*: We evaluate our network over six types of individual behaviors represented through six personas. To compare the network performance over these personas, we show the averaged RMSE over ten days period for each of them in Figure 5.b. It is obvious from the plots that our network performs consistently over different types of behavior. In the beginning, the error is larger and it is reduced with time through learning. Due to space constraints, in the following we only present simulation results for persona 6 since the RMSE for this persona remains at an average of all the six personas. Obtained results for the five other personas are consistent with those presented here.

Similarly to the global case, we depict in Figure 5.c the estimated time of sessions versus their actual time over the 200 days period for 10 users of persona 6. Here, the estimation is more accurate than in the global case with a slight tendency to average the user behavior through over-estimated short

Estimator	Global case		Local case	
	Success in %	RMSE in minutes	Success in %	RMSE in minutes
BN	40.18	17.55	46.42	6.66
EMA	46.17	19.14	48.62	15.47
BR	80.6	21.3	76.95	15.99

TABLE III
SUCCESSFUL ESTIMATIONS AND AVERAGE ERRORS OF ALL APPROACHES

sessions and under-estimated long ones.

C. Comparative analysis

In order to evaluate the benefits of the use of a Bayesian network (BN), we compare it with two standard estimators that only rely on the past behavior of the system. They do not integrate any contextual information.

1) *Concurrent estimators*: The first estimator (EMA) is a statistical technique which estimates an average from a set of values by giving exponentially decreasing weights to older values. As given in (2), ES_t is the current session duration, S_{t-1} is the actual duration of last session, ES_{t-1} is the length of the last estimated session and α is a weighting factor in $[0, 1]$. The chosen optimal value of α is 0.7 suggesting to give more weights to the recent session durations.

$$ES_t = \alpha \times S_{t-1} + (1 - \alpha) \times ES_{t-1} \quad (2)$$

The second estimator (BR) is a probabilistic approach based on a Bayesian inference, which enables the estimation of peers' session duration in the presence or absence of past sessions' history. We assume that the prior probability distribution is a uniform one, avoiding thus any strong assumption on the user behavior. Due to space constraints we cannot add details of estimators any more. For further information, the reader can refer to [21].

2) *Comparison metrics*: We compare the three estimators (EMA, BN and BR) according to the following metrics. Firstly, we measure *success* which stands for the number of times our estimated session duration is less than or equal to the actual session duration which is important to anticipate the departure of a node. Secondly, we measure the *early reaction time* (ERT) which is the difference between the actual session duration and estimated duration when the latter is less than the former one. This measurement shows how optimal the anticipation is. Thirdly, we measure the *error* that is the difference between the actual session duration and estimated session duration whatever the success. In the global case, we give the error for the 20th simulation day. In the local case, we give RMSE since the error is averaged over all peers. Figure 6 depicts obtained results in both global and local case and table III gives synthetic results.

3) *Global case*: Success rates given in Table III clearly show that BR outperforms the two other estimators while presenting a slightly larger RMSE. Moreover, Figure 6.a reinforces the analysis we gave in section IV-B2 by showing that errors of both BN and BR are mainly positive, showing

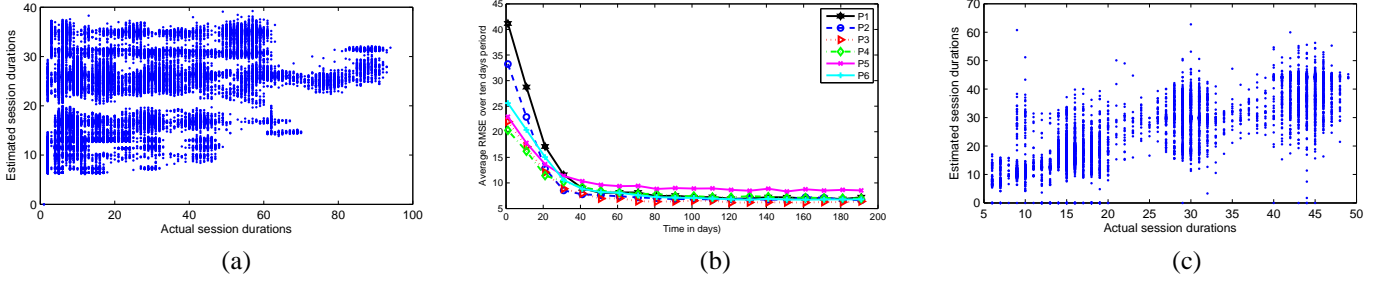


Fig. 5. (a) Actual sessions vs. estimated ones in global case; (b) RMSE averaged over 10 days period for the six personas; (c) Actual sessions vs estimated ones in local case

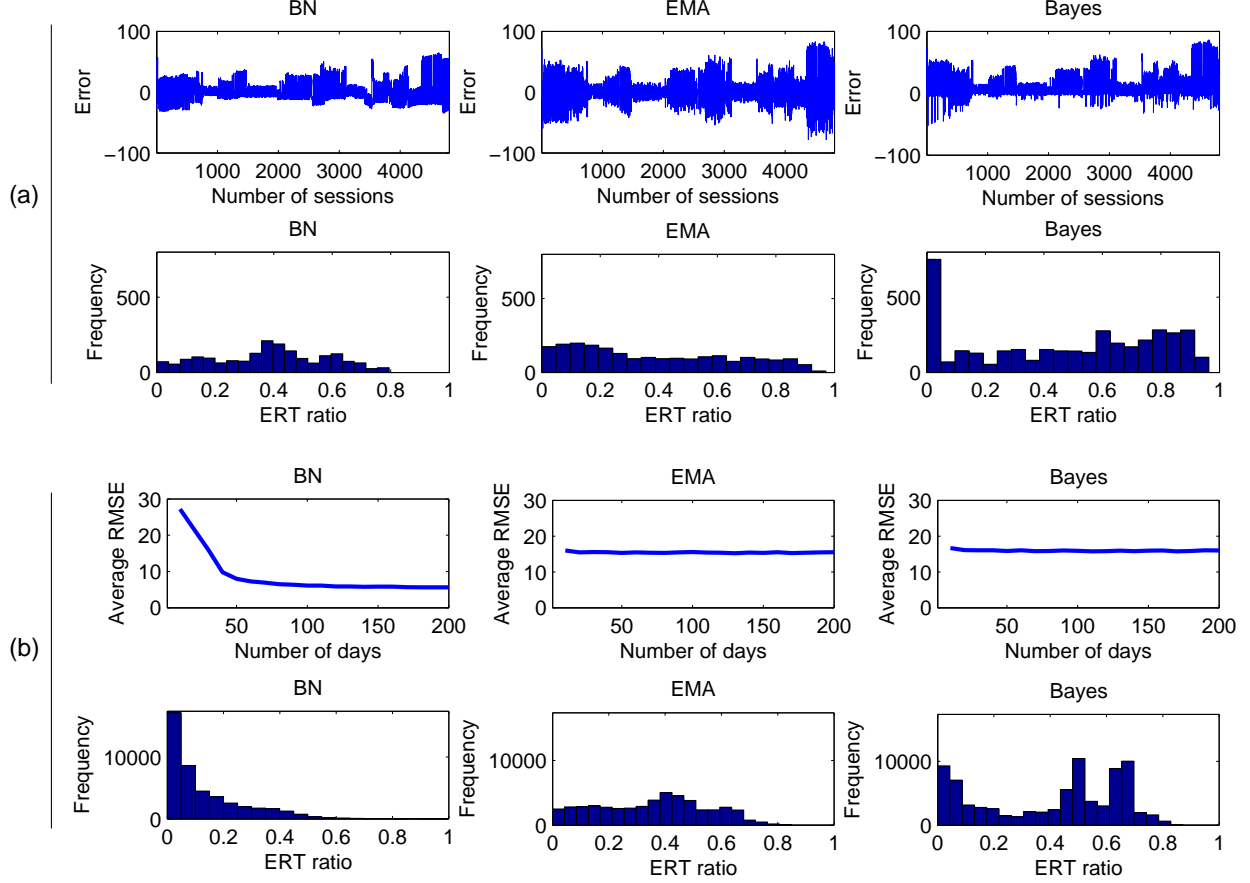


Fig. 6. Comparison: (a) Global case; (b) Local case

an under-estimating tendency. As for EMA, it presents a fairly distributed error. For successful estimations, ERT is far better for BR since most of estimated sessions present a ratio under 0.05, exhibiting a very good accuracy. EMA and BN present similar distribution with a slight advantage for BN whose ERT does not exceed 0.8.

4) *Local case*: Success rate provided in Table III provides mitigated results in the local case. Although BR seems to be better than other estimators in terms of success rate, its accuracy given by RMSE is far worse than that of BN. Concerning EMA, it cumulates the weaknesses of both BR and BN producing a weak success and high RMSE. Figure 6 confirms these results by showing that once the learning period

is over, BN presents the best RMSE with an excellent accuracy illustrated by its ERT ratio distribution. As a conclusion, in a local case, BN is the most precise estimator. Its average success rate is explained by the fact that the network is not biased towards under- or over-estimation of sessions. In a peer's departure anticipation application it should integrate safety interval by lowering estimated sessions by a given factor. We already tested a 20% factor which makes in this case BN the best estimator according to both success-rate and accuracy.

V. CONCLUSION

In this paper, we presented a Bayesian network modeling a user behavior in a P2P live streaming system. This model gathers in a single formalism the phenomena and causal relationships that have been established through measurement campaigns. More specifically, the information we get consists of components of users behavior, its external impacting factors from the environment and network, and the relationships among these elements. Such a Bayesian model can be used in different applications. The one we focus on is the estimation of peers' session time which aims at anticipating their departure. Simulation results we have shown let us conclude in the following way.

In a system composed of a variety of users having different behaviors, the use of our Bayesian network in a central entity based on-and-providing system-scale estimation is not the best solution since the global behavior of the system is averaged and its specific aspects cannot be handled. A simplest Bayesian estimator based on the sole previous session times outperforms the Bayesian network in terms of success rate while providing a comparable accuracy. By contrast, using a Bayesian network directly into peers which present a consistent and constant behavior is relevant since it is the sole estimator able to estimate the peer behavior very accurately. This result has been confirmed whatever the type of user we considered. Improvements in the success rate for the peer's departure application should be easily obtained through the use of a small safety interval in session time estimations.

Perspectives opened by this work are numerous. Our current work consists in consolidating simulations through extended scenarios (e.g. mix of personas in peers to address the case of multi-viewers). Short term future work will consist in designing, implementing and evaluating a protocol that can be integrated into peers in order to ensure the communication part of our Bayesian network. We are especially looking at existing implementations to get experimental results. As long term future work, we plan to use our model in other application contexts, such as the estimation of the streaming quality or the anticipation of population changes and especially flash crowds that can be useful to provide a constant quality of service to users whatever the environmental context is.

REFERENCES

- [1] S. Spoto, R. Gaeta, M. Grangetto, and M. Sereno, "Analysis of PPLive through active and passive measurements," in *IEEE International Symposium on Parallel&Distributed Processing*, 2009, pp. 1–7.
- [2] L. Vu, I. Gupta, J. Liang, and K. Nahrstedt, "Measurement and modeling of a large-scale overlay for multimedia streaming," in *International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness*, 2007, pp. 1–7.
- [3] F. Wang, J. Liu, and Y. Xiong, "Stable peers: Existence, importance, and application in Peer-to-Peer live video streaming," in *IEEE INFOCOM*, 2008, pp. 1364–1372.
- [4] X. Hei, C. Liang, J. Liang, Y. Liu, and K. W. Ross, "A measurement study of a large-scale P2P IPTV system," *IEEE Transactions on Multimedia*, vol. 9, no. 8, pp. 1672–1687, 2007.
- [5] T. Silverston, O. Fourmaux, A. Botta, A. Dainotti, A. Pescapè, G. Ventre, and K. Salamatin, "Traffic analysis of Peer-to-Peer IPTV communities," *Computer Networks*, vol. 53, no. 4, pp. 470–484, 2009.
- [6] J. Jia, C. Li, and C. Chen, "Characterizing PPStream across Internet," in *IFIP International Conference on Network and Parallel Computing Workshops*, 2007, pp. 413–418.
- [7] K. Shami, D. Magoni, H. Chang, W. Wang, and S. Jamin, "Impacts of peer characteristics on P2PTV networks scalability," in *IEEE INFOCOM*, June 2009, pp. 2736–2740.
- [8] X. Zhang, J. Liu, and B. Li, "On large scale Peer-to-Peer live video distribution: Coolstreaming and its preliminary experimental results," in *IEEE International Workshop on Multimedia Signal Processing*, 2005.
- [9] S. Xie, G. Y. Keung, and B. Li, "A measurement of a large-scale Peer-to-Peer live video streaming system," in *International Conference on Parallel Processing Workshops*, 2007, pp. 57–62.
- [10] B. Li, S. Xie, G. Y. Keung, J. Liu, I. Stoica, H. Zhang, and X. Zhang, "An empirical study of the Coolstreaming+ system," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 9, pp. 1627–1639, 2007.
- [11] S. Agarwal, "A case study of large scale P2P video multicast," in *International Conference on IP Multimedia Subsystem Architecture and Applications*, 2007, pp. 1–5.
- [12] Y. Tang, L. Sun, K. Zhang, S.-Q. Yang, and Y. Zhong, "Longer, better: On extending user online duration to improve quality of streaming service in P2P networks," in *IEEE International Conference on Multimedia and Expo*, 2007, pp. 2158–2161.
- [13] Z. Liu, C. Wu, B. Li, and S. Zhao, "Distilling superior peers in large-scale P2P streaming systems," in *IEEE INFOCOM*, April 2009, pp. 82–90.
- [14] Y. Tang, L. Sun, J.-G. Luo, and Y. Zhong, "Characterizing user behavior to improve quality of streaming service over P2P networks," in *Advances in Multimedia Information Processing*, 2006, pp. 175–184.
- [15] K. Sripanidkulchai, B. Maggs, and H. Zhang, "An analysis of live streaming workloads on the Internet," in *IMC'04*, 2004, pp. 41–54.
- [16] E. Veloso, V. Almeida, W. J. Meira, A. Bestavros, and S. Jin, "A hierarchical characterization of a live streaming media workload," *IEEE/ACM Transactions on Networking*, vol. 14, no. 1, pp. 133–146, 2006.
- [17] T. Qiu, Z. Ge, S. Lee, J. Wang, J. Xu, and Q. Zhao, "Modeling user activities in a large IPTV system," in *9th ACM SIGCOMM Conference on Internet Measurement*, 2009, pp. 430–441.
- [18] T. Qiu, Z. Ge, S. Lee, J. Wang, Q. Zhao, and J. Xu, "Modeling channel popularity dynamics in a large IPTV system," in *11th international joint conference on Measurement and Modeling of Computer Systems*, 2009, pp. 275–286.
- [19] M. Cha, P. Rodriguez, J. Crowcroft, S. Moon, and X. Amatriain, "Watching television over an IP network," in *8th ACM SIGCOMM conference on Internet measurement*, 2008, pp. 71–84.
- [20] S. Horovitz and D. Dolev, "Collabrium: Active traffic pattern prediction for boosting P2P collaboration," in *IEEE International Workshops on Enabling Technologies*, 2009, pp. 116–121.
- [21] I. Ullah, G. Bonnet, G. Doyen, and D. Gaïti, "Improving performance of ALM systems with Bayesian estimation of peers dynamics," in *MMNS*, 2009, pp. 157–169.
- [22] —, "Modeling user behavior in P2P live video streaming systems through a Bayesian network," in *AIMS*, 2010, pp. 2–13.
- [23] L. Guo, S. Chen, Z. Xiao, E. Tan, X. Ding, and X. Zhang, "Measurements, analysis, and modeling of BitTorrent-like systems," in *IMC '05*, 2005, p. 3548.
- [24] Z. Liu, C. Wu, B. Li, and S. Zhao, "Why are peers less stable in unpopular P2P streaming channels?" in *Networking*, 2009, pp. 274–286.
- [25] A. Rudstrom and M. Sjolinder, "Capturing TV user behaviour in fictional character descriptions," Swedish Institute of Computer Science, Tech. Rep., October 2008.
- [26] H. Yu, D. Zheng, B. Y. Zhao, and W. Zheng, "Understanding user behavior in large-scale video-on-demand systems," *ACM SIGOPS Operating Systems Review*, vol. 40, no. 4, pp. 333–344, 2006.
- [27] A. Brampton, A. MacQuire, I. Rai, N. J. P. Race, L. Mathy, and M. Fry, "Characterising user interactivity for sports video-on-demand," in *International workshop on Network and Operating Systems Support for Digital Audio & Video*, June 2007.