IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 6, No 2, November 2011
ISSN (Online): 1694-0814
www.IJCSI.org

Study of Reputation Based Defense Mechanism in Peer To Peer Live Streaming

Prof. Dr. M.A. Mottalib[1], Md. Ali-Al Mamun[2], Reazul Hoque[3], Md. Ehsanuzzaman[4] and Jonayed Kaysar[5]

[1,2,3,4,5] Department of Computer Science and Information Technology (CIT)
Islamic University of Technology (IUT)
Board Bazar, Gazipur-1704, Bangladesh

Abstract

Peer-to-peer live video streaming systems are having more popularity as the information technology is advancing fast. Peer-to-peer live video streaming systems are like peer-to-peer file sharing system and they are also vulnerable to content pollution attack as file sharing. In this type of attack, the attacker mixes polluted or unnecessary data into the streaming data and forwards the polluted data to normal peers and hence the perceived video quality is decreased. In this paper, a comparative study among three pollution defense mechanisms in peer-to-peer live streaming systems: Blacklisting, Simple Decentralized Reputation System and Bayesian Approach to Reputation System. Finally performance analysis and the derived result has been presented.

Keywords: *coolstreaming, pollution index, pollution attack, peer-to-peer.*

1. Introduction

In peer-to-peer live video streaming system, at first the video to be streamed is divided into segments and distributed to its partners, known as peers or nodes. These peers forward the segments to other peers, that means video segments can be downloaded from other peers thus eliminates the need for powerful servers.

Popular systems, such as CoolStreaming/DONet [2] use data driven and mesh-pull based overlay network. The content of the video is segmented into chunks by only one source. When peers join into the channel, it makes neighbor ship with a subset of peers. They exchanges buffer map with their neighbors to request for and download video chunks from the neighborhood. When an attacker joins into the network, it downloads chunks and inserts corrupted chunks making the data polluted. When a peer naively request for chunks, attacker may send the forged data, thus the naïve peer gets polluted and to keep efficiency constant, have to keep and integrate the forged data into video and the perceived video quality is decreased. These polluted peers send polluted data to other peers and at a point of time, the whole network becomes polluted.

In this paper, we present through simulation, an evaluation of efficiency of some proposed defense mechanisms. In section 2, we discuss the procedure of the defense mechanisms: Blacklisting, Simple Decentralized Reputation System and Bayesian Approach to Reputation System. In section 3, we discuss our experiment model and in section 4, we show the performance evaluation of the defense mechanisms.

2. Discussion On Defense Mechanisms

2.1 Blacklisting

This is a reputation based defense mechanism [8]. In this system, each peer actively monitors other peers' behavior and according to that, reputes a score for that peer; this reputation information is reported to a centralized server, which calculates the total reputation of each peers. The main goal is to identify and isolate the attacker. Each peer *i* after downloading chunk from peer *j* , checks the ratio of total number of polluted chunks (*n*) downloaded from *j*

and total number of requested chunks (r) to peer j. peer i reputes peer j according to (1)[9].

$$I_{(i,j)} = \begin{cases} \max\left(0, I_{(i,j)}^{t-1} - \alpha_p * (1 + n_i/r)^{|r|}\right) & \text{if } n_i/r > thr \\ \min\left(1, I_{(i,j)}^{t-1} + \alpha_g * n_i/r\right) & \text{otherwise} \end{cases} \quad (1)$$

Here, α_p and α_g are reward and penalty value for the classification as clean and polluted interaction respectively and $\alpha_p > \alpha_g$. $I_{(i,j)}^{t}$ denotes the reputation score of peer j, given by peer i. Each peer i periodically sends the reputation information to a centralized server which calculates the reputation of j by weighting the reported score by the reputation score of i, R_i as in (2)[9].

$$R_j = \begin{cases} \frac{\sum_{i \in N} I_{(i,j)} * R_i}{\sum_{i \in N} R_i} \end{cases} \quad (2)$$

Here, N is the set of peers that have reputed peer j. After calculating the score, it is forwarded to the peer i, and if the reputation score is less than a threshold value, then peer i will stop it's partnership with peer j.

2.2 Simple Decentralized Reputation System

This mechanism was proposed by Alex Borges Vieira, Sergio Campos and Jussara Almeida [3]. This approach is decentralized and simpler than other proposed mechanisms. Here each node takes into consideration only its individual experience. In this approach, node p_i computes reputation of node p_j periodically based on the ratio of polluted chunks received and total chunks received. If p_j's response to p_i's request includes n polluted chunks and if the fraction of polluted chunks is below the limit (limit is a value chosen by each peer), then p_i considers its interaction with p_j as good and increases its reputation score. In all other cases p_i decreases p_j's reputation score. If p_j's score becomes lower than a threshold value than node p_i stops its partnership with node p_j.

$$R_i[p_j] < R_{min} \quad (3)$$

Where *Rmin* is the threshold value. This system allows rehabilitation of peers. For that two system states are used, calm and storm state. When more than one polluter or malicious peer is attacking the system, it is in storm state. Otherwise the system is in calm state.

When the system is in storm state the threshold value, $R_{min,}$ is increased and when the system is in calm state the threshold value is relaxed.

$$Rmin_t = \begin{cases} \min\left(Rt_{maxt}, Rmin_t + \gamma_{pt}\right) & Storm \\ \min\left(Rt_{mint}, Rmin_t - \gamma_{gt}\right) & Clam \end{cases} \quad (4)$$

Equation(4) shows the calculation of dynamic threshold value. The threshold value may vary from Rt_{maxi} (worst storm state) to Rt_{mini} (best calm state). Here $\gamma_{pi} > \gamma_{gi}$, so R_{mini} increases faster [3].

2.3 Bayesian Approach to Reputation System

This mechanism was proposed by Sonja Buchegger and Jean-Yves Le Boudec [7]. In this approach every node i maintains two ratings about its partner. The two ratings are reputation rating and trust rating. Reputation rating is the opinion of node i about its partner node j and the trust rating is node i's opinion about how honest node j is as a participator in the reputation system. Data structure $R_{(i,j)}$ is used for reputation and $T_{(I,j)}$ for trust rating. Node i also maintains data structure $F_{(i,j)}$ for record of first hand information about node j. This approach uses two rating as it takes into account other's experience with its own. Here first, when node i makes a firsthand observation of node j's behavior, $R_{(i,j)}$ and $F_{(i,j)}$ are updated. Second, nodes periodically publish their first hand information to its neighbors. For example node i receives first hand information about node j from node k and if node k is classified as "trustworthy" by i or if $F_{(k,j)}$ is close to $R_{(i,j)}$ then $F_{(k,j)}$ is accepted and $R_{(i,j)}$ is updated slightly. Else, $R_{(i,j)}$ is not updated. In all cases trust rating $T_{(i,k)}$ is updated. If $F_{(k,j)}$ is close to $R_{(i,j)}$ then $T_{(i,k)}$ slightly improves, else it slightly worsens. In this method only first hand information is shared with others.

Every node uses its ratings to classify other nodes from time to time, according to two criteria: 1- *normal/misbehaving* 2- *trustworthy/untrustworthy*. Both classifications are computed using the Bayesian approach, based on reputation ratings for the first and trust ratings for the latter.

This decentralized reputation-based defense mechanism is based on the Bayesian trust model that uses the Beta distribution to monitor the behavior of peers in the network, *Beta(α, β)* , where α denotes misbehavior cases and β denotes normal behavior cases. Initially, the prior is *Beta(1, 1)* and there is a parameter θ such that a peer is misbehaving with probability θ. If θ is constant, after n observations $\alpha \sim n\theta$ and $\beta \sim n(1-\theta)$.

The first hand information ($F_{i,j}$) has the form *(α, β)*. α and β of peer j are updated by peer i using following equations: $\alpha := u\alpha + s$ and $\beta := u\beta + (1-s)$. Where s is 1 for each misbehavior case and 0 for each normal

behavior case and u is the discount factor for past experiences which works as fading mechanism

In this approach, the values of α and β are periodically decayed. When the inactivity time expires, values of α and β are set as follows: $\alpha = u\alpha$ and $\beta = u\beta$ .This is done for redemption in absence of observations.

In this system, reputation rating $(R_{i,j})$ has the *form* (α', β'). Initially they are set to $(1,1)$. There are two cases when the reputation rating $(R_{i,j})$ is updated:

1. In the first case the update is same as first hand information update.
2. In the second case the update is done when a reputation rating published by other peer is accepted and copied. For example let, peer i receives $F_{k,j}$ from peer k. If $T_{i,k}$ is such that k is trustworthy to peer i, according to the (5) :

$$\begin{cases} trustworthy & if\ E(Beta(\gamma', \delta')) < t \\ untrustworthy & if E\left(Beta\left(\gamma', \delta'\right)\right) \geq t \end{cases} \quad (5)$$

Here, threshold t is tolerance; then peer i modifies $R_{i,j}$ as follows :

$$R_{i,j} := R_{i,j} + wF_{k,j} \quad (6)$$

w is a small positive constant. Otherwise, i considers k as untrustworthy and does a deviation test :

$$|E(Beta(\alpha_F, \beta_F)) - E(Beta(\alpha, \beta))| \geq d \quad (7)$$

Where $F_{k,j} = (\alpha_F, \beta_F)$ and $R_{i,j} = (\alpha, \beta)$ and $E(Beta(\alpha, \beta))$ is the expectation of distribution $Beta(\alpha, \beta)$ and d is a positive constant. If the test is positive, that means peer k is praising j much and is not used. Otherwise, $R_{i,j}$ is updated according to (6).

The trust rating is always updated and is updated in the same way as reputation rating.
Finally node i classify j's behavior and trustworthiness as

$$\begin{cases} normal & if\ E(Beta(\alpha', \beta')) < r \\ misbehaving & if E\left(Beta\left(\alpha', \beta'\right)\right) \geq r \end{cases} \quad (8)$$

Here, r is tolerance.

$$\begin{cases} trustworthy & if\ E(Beta(\gamma', \delta')) < t \\ untrustworthy & if E\left(Beta\left(\gamma', \delta'\right)\right) \geq t \end{cases} \quad (9)$$

Here, t is tolerance and in this way the system works.

3. Experiment Model

We have created a simulator using the programming language Java to simulate DONet [2] system which stands for the Data Driven Overlay Network; this is commercially known as Cool-Streaming. We have chosen this specific system because it is a mesh-based network and it has gained commercial success among the mass people after being deployed commercially.

3.1 System Model

In our simulated peer-to-peer mesh based system, each node has a unique identifier (ID) and a membership cache (mCache) which contains list of identifiers of neighbor nodes. When a node joins the network, it first contacts the origin node which chooses a deputy for the node randomly from its mCache and redirects the newly joined node to the deputy it has selected. Then the node can get a list of partner candidates from the deputy and establishes partnership with these candidates. Each node also contains buffer map which represents the data availability of that node, and this is continuously exchanged with the partners and schedules the exchanging of data segments based on the buffer map it has received. Each node also has a cache where the downloaded segments are stored.

For the validation of our simulator we have generated a graph from our result depicting the continuity index of peers with number of partners equaling to 4. Continuity index is the number of segments that arrive on or before playback deadline over the total number of segments.

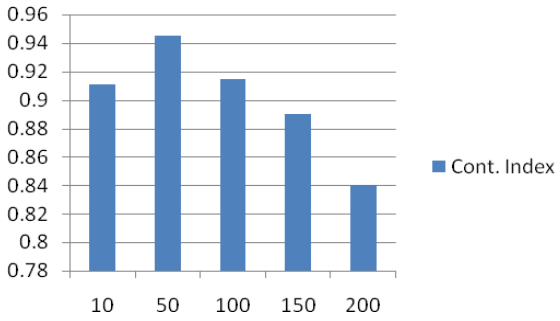

Fig.1 Continuity Index (x-axis denotes number of peers and y-axis denotes Continuity Index)

Our result matches very closely with the result given in [2].

3.2 Experiment Parameters

In this simulation, we have assumed that all peers have 4 neighbors as in [2] it is showed as the optimum one. Number of total nodes are 100. Each node joins the network at an interval of 2 seconds. We have used heterogeneous bandwidth and for that we have assigned random bandwidth between100 to 150 kBps. The length of the streaming video is 500 seconds. The whole streaming video is divided into 500 segments each containing 1 sec video. Each node maintains a sliding window of 12 segments. The playback starts 25 second after receiving the 1st downloaded segment.

4. Experiment Result

We have used three data sets to compare the reputation based defense mechanisms. The data sets are Network Pollution index, Percentage of Polluted peer and peer Pollution index. We considered the percentage of polluted peer up to 40% as within this range we got a clear picture of which mechanism was most effective.

4.1 Network Pollution Index

At first we have calculated the network pollution index .Network pollution index is the total number of polluted segments divided by the total number of segments in the network. Below four graph is given, each graph shows the Network Pollution index with time for the three chosen reputation systems –Blacklisting, Simple decentralized reputation system and Bayesian approach to reputation system and the streaming model without any defense mechanism- DONet. The first graph shows for 10% malicious peers and the next increases by 10 up to 40%.

Fig.2 For 10% malicious peers(x-axis denotes time and y axis denotes Network Pollution Index)

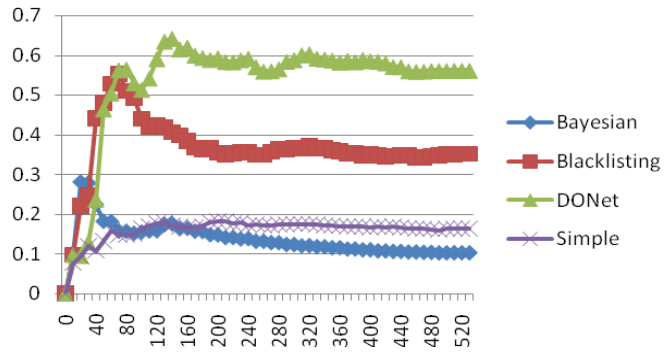

Fig.3 For 20% malicious peers(x-axis denotes time and y axis denotes Network Pollution Index)

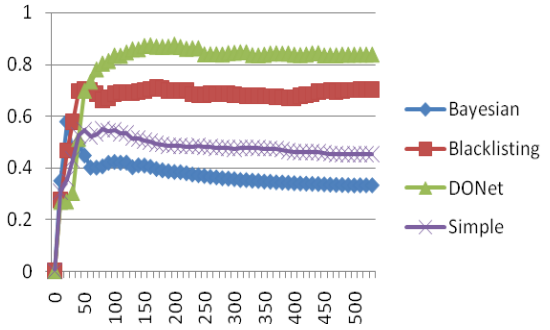

Fig.4 For 30% malicious peers(x-axis denotes time and y axis denotes Network Pollution Index)

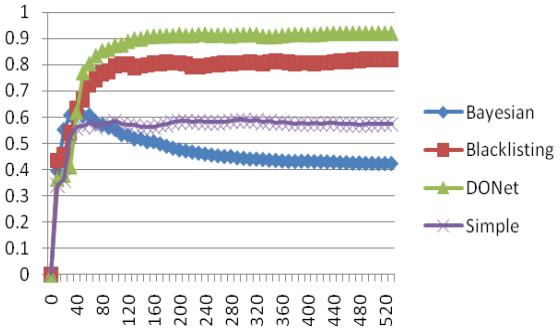

Fig.5 For 40% malicious peers(x-axis denotes time and y axis denotes Network Pollution Index)

4.2 Percentage of Polluted Peer

Next we have calculated the Percentage of Polluted Peer in the network. Percentage of polluted peer is the number of polluted peers divided by the total number of peers in the network. A peer is considered polluted when 10% of its downloaded segments are polluted. Below four graph is given, each graph shows the percentage of polluted peer in the network with time for the three chosen reputation systems- Blacklisting, Simple decentralized reputation

system and Bayesian approach to reputation system and the streaming model without any defense mechanism- DONet. The first graph shows for 10% malicious peer and the next increases by 10 up to 40%.

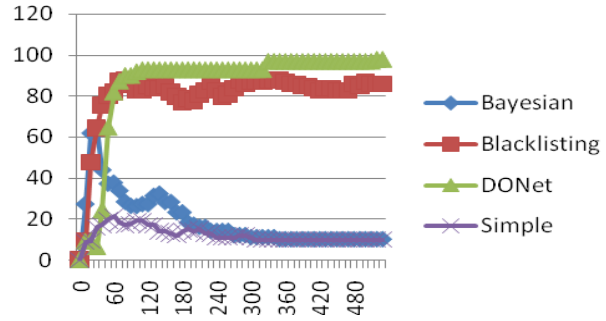

Fig.6 For 10% malicious peers(x-axis denotes time and y axis denotes Percentage of Polluted Peer)

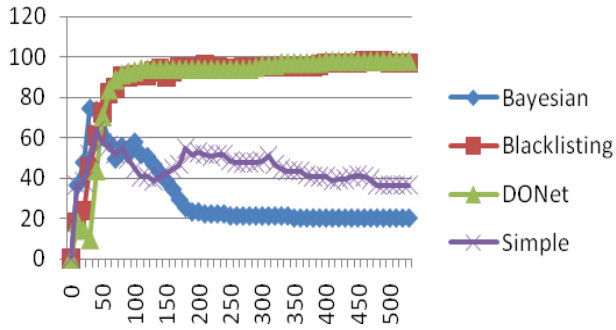

Fig.7 For 20% malicious peers(x-axis denotes time and y axis denotes Percentage of Polluted Peer)

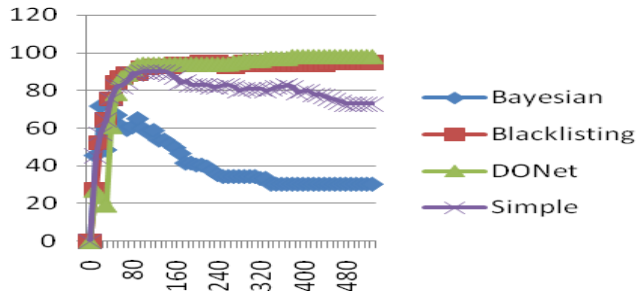

Fig.8 For 30% malicious peers(x-axis denotes time and y axis denotes Percentage of Polluted Peer)

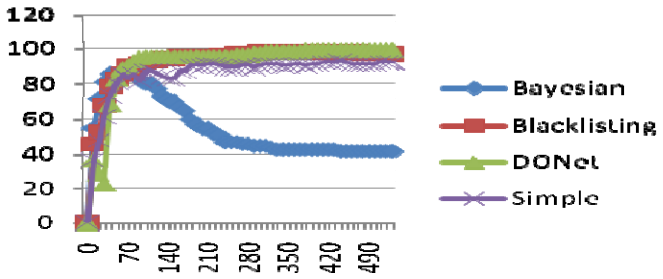

Fig.9 For 40% malicious peers(x-axis denotes time and y axis denotes Percentage of Polluted Peer)

4.3 Peer Pollution Index

Lastly we have calculated Peer Pollution index. Peer Pollution index is the average number of polluted segments divided by the average number of segments in each peer. Below four graph is given , each graph shows the peer pollution index with time for the three chosen reputation systems- Blacklisting, Simple decentralized reputation system and Bayesian approach to reputation system and the streaming model without any defense mechanism- DONet. The first graph shows for 10% malicious peer and the next increases by 10 up to 40%.

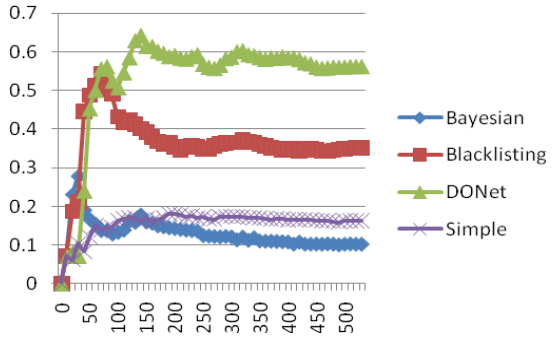

Fig.10 For 10% malicious peers(x-axis denotes time and y axis denotes Peer Pollution Index)

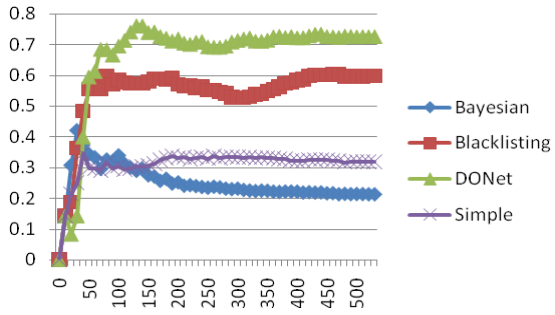

Fig.11 For 20% malicious peers(x-axis denotes time and y axis denotes Peer Pollution Index)

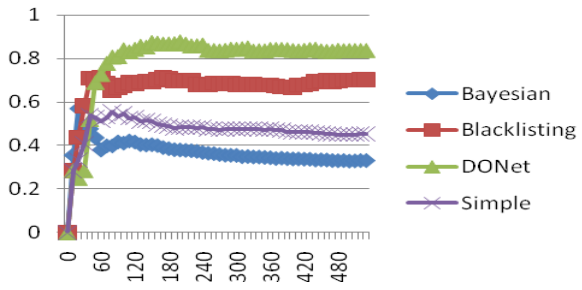

Fig.12 For 30% malicious peers(x-axis denotes time and y axis denotes Peer Pollution Index)

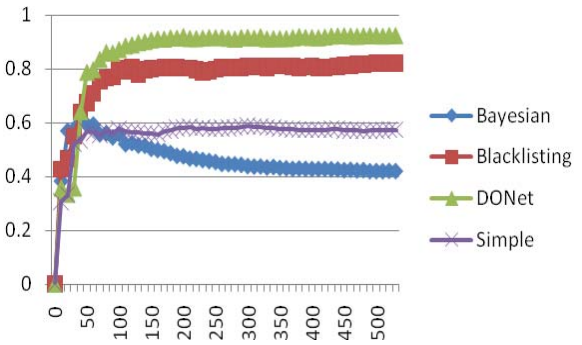

Fig.13 For 40% malicious peers(x-axis denotes time and y axis denotes Peer Pollution Index)

5. Conclusion

In this paper, we have studied peer to peer live streaming, peer to peer live streaming under pollution attack and peer to peer live streaming with defense mechanism under pollution attack. Peer to peer live streaming without any defense mechanism under pollution attack shows devastating result. We have simulated three defense mechanisms namely Blacklisting, Simple decentralized Reputation System and Bayesian Approach to Reputation System.

By analyzing the data obtained from the simulation, we can see that the Network Pollution index of Bayesian Approach to Reputation System is considerably lower for all 10-40% of malicious peers.

We can also see that, the Peer Pollution index and Percentage of Polluted Peers of Bayesian Approach to Reputation System is also considerably lower for all 10-40% of malicious peers.

Hence we can come to the conclusion that Bayesian Approach to Reputation System is a feasible and suitable defense mechanism to be used in peer-to-peer live video streaming applications.

References

[1] Rudiger Schollmeier, ”A Definition of Peer-to-Peer Networking for the Classifiction of Peer-to-Peer Architectures and Applications”. Proceedings of the First Interbational Conference on Peer-to-Peer Computing (P2P.01).2002.0-7695-1503-7.

[2] Xinyan Zhang, Jiangchuan Liu, Bo Li and Tak-Shing Peter Yum, “CoolStreaming/DoNet:A Data-Driven Overlay Network for Efficient Live Media Streaming ”.in Proc. Of the 24th Conference of the IEEE Communication Society (INFOCOM 2005), March 2005.

[3] Alex Borges Vieira, Sergio Campos and Jussara Almeida, “Fighting Attacks In P2P Live Streaming. Simpler is Better”. in IEEE International Conference on Multimedia & Expo,Hanover, Germany, 2008.

[4] Prithula Dhungel, Xiaojun Hei, Keith W. Ross, and Nitesh Saxena, “The Pollution Attack in P2P Live Video Streaming: Measurement Results and Defenses”. in Proc. of ACM Workshop on Peer-to-peer Streaming and IP-TV (P2P-TV 2007),New York, NY, USA, August 2007, pp. 323–328.

[5] Eric Lin, Daniel Medeiros Nunes de Castro, Mea Wang and John Aycock, “SPoIM: A Close Look at Pollution Attacks in P2P Live Streaming”. On quality of service (IWQoS), 2010 18th International Workshop.1548-615X.

[6] Samarth Shetty, Patricio Galdames, Wallapak Tavanapong and Ying Cai, “Detecting Malicious Peers in Overlay Multicast Streaming”.Proceedings on Local Computer Networks 2006 31st IEEE Conference. 0742-1303 .

[7] Sonja Buchegger and Jean-Yves Le Boudec, “A Robust Reputation System for P2P and Mobile Ad-hoc Networks”. Proceedings of P2PEcon 2004, Harvard University, Cambridge MA, U.S.A., June 2004.

[8] Alex Borges, Jussara Almeida, Sergio Campos, *“Fighting Pollution In P2p Live Streaming Systems”*. on IEEE International Conference Multimedia and Expo, 2008