Traffic Pattern based Content Leakage Detection for Trusted Content Delivery Networks

Amulya A[1],Anusha K A[2], Jeevitha T S[3], Komala B E
Dept. of Computer Science & Engineering
K S Institute of Technology
Bengaluru ,India

Mr Kushal Kumar B N
Assistant Professor
Dept. of CSE, KSIT

***Abstract*—Due to the increasing popularity of multimedia streaming applications and services in recent years, the issue of trusted video delivery to prevent undesirable content-leakage has, indeed, become critical. While preserving user privacy, conventional systems have addressed this issue by proposing methods based on the observation of streamed traffic throughout the network. These conventional systems maintain a high detection accuracy while coping with some of the traffic variation in the network (e.g., network delay and packet loss), however, their detection performance substantially degrades owing to the significant variation of video lengths. In this paper, we focus on overcoming this issue by proposing a novel content-leakage detection scheme that is robust to the variation of the video length. By comparing videos of different lengths, we determine a relation between the length of videos to be compared and the similarity between the compared videos. Therefore, we enhance the detection performance of the proposed scheme even in an environment subjected to variation in length of video. Through a testbed experiment, the effectiveness of our proposed scheme is evaluated in terms of variation of video length, delay variation, and packet loss.**

Index Terms — Streaming content, leakage detection, traffic pattern, degree of similarity

I.INTRODUCTION

In recent years, with the rapid development of broadband technologies and the advancement of high-speed wired/wireless networks, the popularity of real-time video streaming applications and services over the Internet has increased by leaps and bounds. YouTube and Microsoft network video are notable examples of such applications. They serve a huge population of users from all around the world with diverse contents, ranging from daily news feeds to entertainment feeds including music, videos, sports, and so forth, by using streaming transmission technologies. In addition, real-time video streaming communications such as web conference [1], [2], in intracompany networks or via Internet with virtual private networks (VPNs) are being widely deployed in a large number of corporations as a powerful means of efficiently promoting business activities without additional costs .A crucial concern in video streaming services is the protection of the bit stream from unauthorized use duplication and distribution. One of the most popular approaches to prevent undesirable contents distribution to unauthorized users and/or to protect authors' copyrights is the digital rights management (DRM) technology. Most DRM techniques employ cryptographic or digital watermark techniques.[3], [4]. However, this kind of approaches have significant effect on redistribution of contents, decrypted restored at the user-side by authorized yet malicious Moreover, redistribution is technically no longer difficult using peer-to-peer (P2P) streaming software [5]. Hence, streaming traffic may be leaked to P2P networks.

On the other hand, packet filtering by firewall-equipped egress nodes is an easy solution to avoid leakage streaming contents to external networks. In this solution, the packet header information (e.g., destination and source Internet protocol addresses, protocol type, and port number of outgoing traffic) of every streamed packet is inspected. case the inspected packets do not verify the predefined filtering policy, they are blocked and dropped However, it is difficult to entirely prevent streaming content leakage by means of packet filtering alone because packet header information of malicious users is unspecified beforehand and can be easily spoofed.

In this paper, we focus on the illegal redistribution of streaming content by an authorized user to external networks. The existing proposals in [6] and [7] monitor information obtained at different nodes in the middle of the streaming path. The retrieved information are used to generate traffic patterns which appear as unique waveform per content [8], just like a fingerprint. The generation of traffic pattern does not require any information on the packet header, and therefore preserves the user's privacy. Leakage detection is then performed by comparing the generated traffic patterns. However, the existence of videos of different length in the network environment causes a considerable degradation in the leakage detection performance. Thus, developing an innovative leakage detection method robust to the variation of video lengths is, indeed required. In this paper, by comparing different length videos, we determine a relationship between the length of videos to be compared and their similarity. Based on this relationship, we determine decision threshold

enabling accurate leakage detection even in an environment with different length videos.

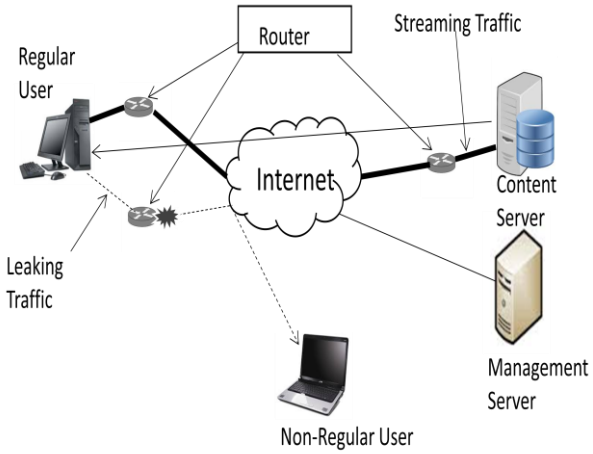

Figure 1: Overview of a leakage scenario and leakage detection scenario

II. CONTENT LEAKAGE DETECTION

In this section, we first take a look at a typical video leakage scenario, and we present an overview of existing traffic pattern based leakage detection technologies.

A. Typical video leakage scenario

Due to the popularity of streaming delivery of movies, development of P2P streaming software has attracted much attention. These technologies enhance the distribution of any type of information over the Internet [13]. A typical content leakage scenario can be described by the following steps. First, a regular user in a secure network receives streaming content from a content server. Then, with the use of a P2P streaming software, the regular yet malicious user redistributes the streaming content to a non-regular user outside its network. Such content-leakage is hardly detected blocked by watermarking and DRM based techniques.

B. Leakage detection procedures

Throughout the video streaming process, the changes of the amount of traffic appear as a unique waveform specific to the content. Thus by monitoring these information, content-leakage can be detected. An overview of the network topology of the proposed leakage detection system is shown in Fig. 1. This topology consists of two main components, namely the traffic pattern generation engine embedded in each router, and the content leakage detection engine. Therefore each router can observe its traffic volume and generate traffic pattern. meanwhile, the leakage detection engine computes the threshold, and based on specific criterion, detects contents leakage. The result is then notified to the target edge router in order to block leaked traffic.

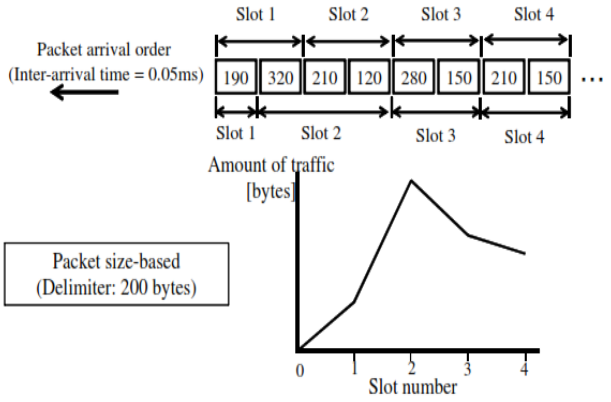

Figure 2: Traffic Pattern Process Generation

C. Pattern generation algorithm

Here, we describe the traffic pattern generation process performed in conventional methods. Traffic pattern generation process is based on a packet size-based algorithm. The traffic pattern generated is expressed as an N-dimension vector as follows,

$$X_N=(x_1,x_2,\ldots.x_N)^T \qquad (1)$$

where x_i indicates the volume of the i^{th} chunk, and N is the total number of chunks.

Packet size-based algorithm defines a slot as the summation of amount of arrival traffic until the observation of a certain packet size. This algorithm only make use of the packet arrival order and packet size, therefore is robust to change in environment such as delay and jitter. However, packet size based algorithm shows no robustness to packet loss. Fig. 2. describes an example of packet size-based generation process. Here, in the packet size-based process, slots are generated by summing the amount of arrival traffic until observing a packet of size less than 200 bytes and based on this observation the graph is generated.

D. Leakage detection criterion

The cross correlation matching algorithm is performed on both the traffic patterns generated through packet size based algorithm [6]. Therefore, [6] uses a dynamic decision threshold based on the Chebyshev's inequality, and given by the following equation:

$$\Theta = \min(\mu_R + 4\sigma_R, 1.0) \qquad (2)$$

where μ_R and σ_R represent the mean and variance respectively.

$$R_{XuYu} = \frac{(X't\ Y't)}{\sqrt{\|X'_U\|^2 \|Y'_U\|}} \qquad (3)$$

Here, whether or not compared patterns are similar is decided by comparing the maximum value of R_{XuYu} with Θ from Eq.2. Meanwhile, during the matching process of packet size based generated traffic patterns, the similarity resulting from the comparison of different videos

is considerably small, while the similarity resulting from the comparison of similar videos is considerably large. A suitable fixed value is therefore used as the decision threshold [6]. To determine whether or not the compared traffic patterns are similar, the maximum value of R_{XuYu} is retrieved and compared to the decision
threshold, i.e., $\max(R_{XuYu}) >$ threshold, which indicates that the compared traffic patterns are similar.

On the other hand, the DP matching algorithm is perfomed on traffic patterns generated through packet size-based algorithm. therefore, a fixed predefined value is used as the decision threshold [7]. Whether or not patterns are similar is decided by comparing the distance computed through DP matching with the decision threshold, i.e., the distance less than the threshold indicates that the compared traffic patterns are similar.

III. PERFORMANCE FOR DIFFERENT LENGTHS OF VIDEOS

In this experiment, we use a set of different videos having the same length and can be perfectly distinguished using the conventional methods, P-TRAT and DP-TRAT. From this set, we generate portions of video of different lengths varying from 30 to 300 seconds. From the generated portions of videos, we randomly choose and send 10, 20 and 30 videos from the server to the user. We then observe the amount of traffic, generate the traffic pattern and perform the matching process.

In other word, the performance degradation observed in this experiment can be considered to be caused by the existence of videos with different lengths. P-TRAT and DPTRAT are used for comparison.

Fig. 3 demonstrates that with the DP-TRAT, the increase in the number of videos decreases the accuracy. the absence of an adequate method to set the decision threshold handling videos of different length causes the occurrence of erroneous decision in the detection performance of the DP-TRAT.

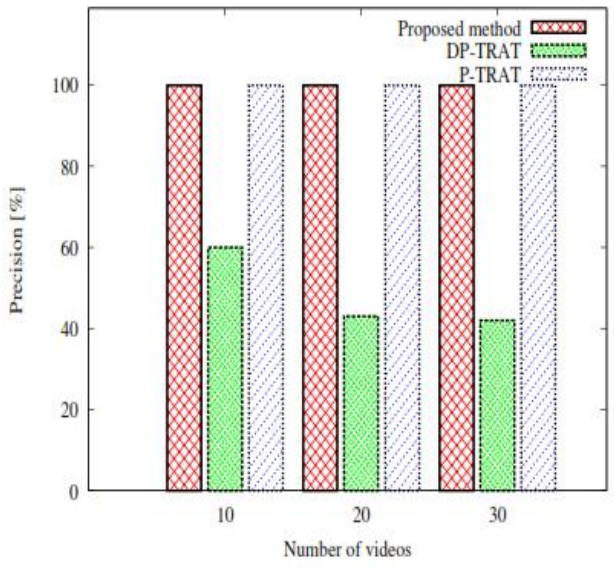

Figure 3: Accuracy

Fig. 4 shows that with the conventional methods (P-TRAT, DPTRAT), the recall ratio is slightly affected by the variation of video length. Fig 5 shows that compare to the conventional methods, the proposed scheme is not affected by the variation of video length.

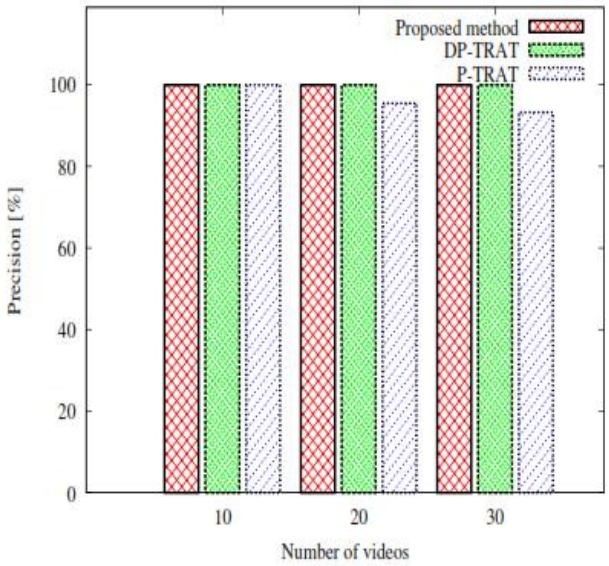

Figure 4: Recall ratio

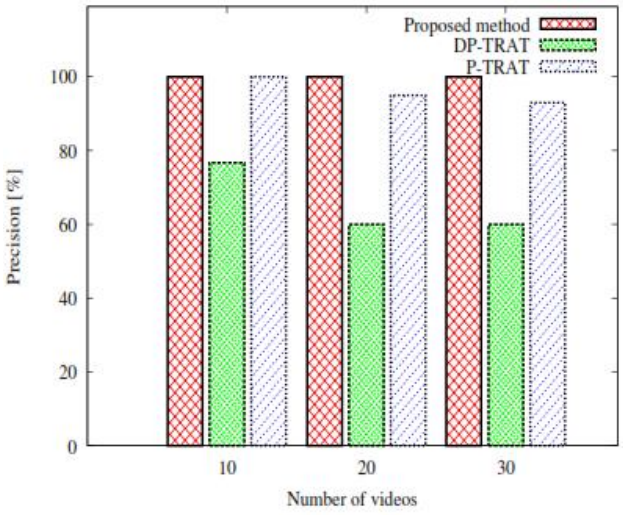

Figure 5: F-measure

IV. CONCLUSION

The content leakage detection system based on the act that each streaming content has a unique traffic pattern is an innovative solution to prevent illegal re-distribution of contents by a regular, yet malicious user. Though three typical conventional methods, namely T-TRAT, P-TRAT, DP-TRAT, show robustness to delay, jitter or packet loss, the detection performance decreases with considerable variation of video lengths. This paper attempts to solve these issues by introducing a dynamic leakage detection scheme. Moreover, in this paper, we investigate the performance of the proposed method under a real network environment with videos of different lengths. The proposed

method allows flexible and accurate streaming content leakage detection independent of the length of the streaming content, which enhances secured and trusted content delivery.

V. REFERENCES

[1] Y. Chu, S. G. Rao, S. Seshan and H. Zhang, "Enabling conferencing applications on the Internet using an overlay multicast architecture," in Proc. ACM SIGCOM, pp.55-67, California, USA, Aug. 2001.

[2] Z. Yang, H. Ma, and J. Zhang, "A dynamic scalable service model for SIP-based video conference," in Proc. 9th nternational Conference on Computer Supported Cooperative Work in DE.

[3]E.I. Lin, A.M. Eskicioglu, R.L. Lagendijk, and E.J. Delp, "Advances in digital video content protection," Proc. IEEE, vol.93, no.1, pp.171-183,Jan. 2005

[4] E. Diehl and T. Furon, "Watermark: Closing the analog hole," in Proc. IEEE Int. Conf. Consumer Electronics, pp.52-53, 2003.

[5] Y. Liu, Y. Guo, and C. Liang, "A survey on peer-to-peer video streaming systems," Peer-to-Peer Networking and applications, Vol.1, No.1, pp.1828, Mar. 2008.

[6] M. Dobashi, H. Nakayama, N. Kato, Y. Nemoto, and A. Jamalipour, "Traitor Tracing Technology of Streaming Contents Delivery using Traffic Pattern in Wired/Wireless Environments," in Proc. IEEE Global Telecommunications Conference, pp.1-5, San Francisco, USA, Nov./Dec. 2006.

[7] K. Matsuda, H. Nakayama, and N. Kato, "A Study on Streaming Video Detection using Dynamic Traffic Pattern," IEICE Transactions on Communications (Japanese Edition), vol.J19-B, no.02, 2010.

[8] S. Amarasing and M. Lertwatechakul, "The Study of Streaming Traffic behavior," KKU Engineering Journal, vol.33, no.5, pp.541-553, Sept.Oct.2006.

[9] Y. Zhang, P. Ma, and X. Su, "Pattern Recognition Using Interval-valued intuitionistic Fuzzy set and Its similarity Degree," IEEE International Conference on Intelligent Computing and Intelligent Systems, 2009.

[10] A. Golaup, and H. Aghvami, "A multimedia traffic modeling framework for simulation-based performance evaluation studies," Computer Network, vol. 50, no. 12, pp. 2071-2087, 2006.