TRAFFIC PATTERN BASED DATA LEAKAGE DETECTION FOR SECURED DATA TRANSFER NETWORKS

Ganesh Udmale, Somesh Gore, Nikhil Hiran, Swapnil Kadam, Prof.Jagtap V.V

[1] *Student, Computer Engineering, Vishwabharti Academy's College Of Engineering, Maharashtra, India*
[2] *Student, Computer Engineering, Vishwabharti Academy's College Of Engineering, Maharashtra, India*
[3] *Student, Computer Engineering, Vishwabharti Academy's College Of Engineering, Maharashtra, India*
[4] *Student, Computer Engineering, Vishwabharti Academy's College Of Engineering, Maharashtra, India*
[5] *Lecturer, Computer Engineering, Vishwabharti Academy's College Of Engineering, Maharashtra, India*

ABSTRACT

As we know that the increasing popularity of Data streaming and services in recent years, the issue of trusted data 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 some e.g. network delay and packet loss, however, their detection performance substantially degrades owing to the significant variation of Data lengths. In this paper, we focus on overcoming this issue by proposing a any type of data content-leakage detection scheme that is robust to the variation of the data length. By comparing videos of different lengths, we determine a relation between the length of data to be compared and the similarity between the compared data. Therefore, we enhance the detection performance of, The effectiveness of our proposed scheme is evaluated in terms of variation of data length, delay variation, and packet loss. It helps to solve dada leakage or content leakage. The advantage lies mainly in advanced content availability protecting the secured data.

Keyword: - *Traffic pattern, Data Leakages, Data transfer network ,Data Detection.*

1. INTRODUCTION:

Internet has increased with the rapid development of broadband Technologies and the advancement of high-speed wired/wireless networks, the popularity of real-time Data transfer applications and services over the Internet has increased. YouTube SAAVAN and Microsoft Network (MSN) video, Audios 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.

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. However, this kind of approaches have no significant effect on re-distribution of contents, decrypted or restored at the user-side by authorized yet malicious users. Moreover, redistribution is technically no longer difficult by using Peer to Peer (P2P) streaming software . Hence, streaming traffic may be leaked to P2P networks.packet filtering by firewall-equipped egress nodes is an easy solution to avoid leakage of streaming contents to external networks. In this solution, the packet header information (e.g., destination and source Internet Protocol (IP) addresses, protocol type, and port number of outgoing traffic) of every streamed packet is inspected.

In this work, we focus on the illegal re-distribution of streaming content by an authorized user to external networks. 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 required. In this paper, by comparing different data videos, we determine a relationship between the length of data contents 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 data.

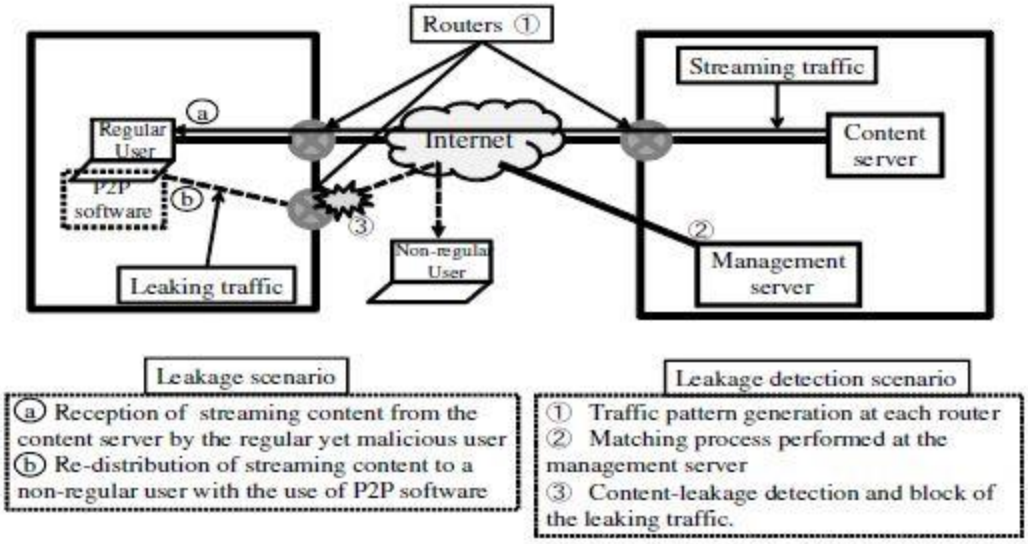

Chart -1: leakage scenario and leakage detection scenario Overview.

2. DATA LEAKAGE DETECTION

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

2.1 Data leakage scenario

Popularity of Transfering of data,development of P2P streaming software has attracted much attention. These technologies enhance the distribution of any type of information over the Internet. A typical data leakage scenario can be described by the following steps as depicted in Fig. 1. 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 or blocked by watermarking and DRM based techniques.

2.2 Leakage detection procedures

The data transferring process, the changes of the amount of traffic appear as a unique waveform specific to the content. Thus by monitoring these information retrieved at different nodes in the network, content-leakage can be detected. An overview of the network topology of the proposed leakage detection system is shown in Fig. 1. Therefore each router can observe its traffic volume and generate traffic pattern. the traffic pattern matching engine computes the similarity between traffic patterns through a matching process, and based on specific criterion, detects contents leakage. The result is then notified to the target edge router in order to block leaked traffic.

2.3 Pattern generation algorithm

The traffic pattern generation process performed in conventional methods. Traffic pattern generation process is based on a either time slot-based algorithm or a packet size-based algorithm. 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 makes 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.

2.4 Pattern *matching algorithm*

In pattern recognition, the degree of similarity is defined to be the similarity measure between patterns. The server side Traffic patterns represents the original traffic pattern and is expressed as XS = (x1; x2; :::; xS) t. The user-side traffic pattern is expressed as YU = (y1; y2; :::; yU) t. Here, S and U are number of slots, and the length of the user-side observation is shorter than that of the server-side, i.e., S > U.

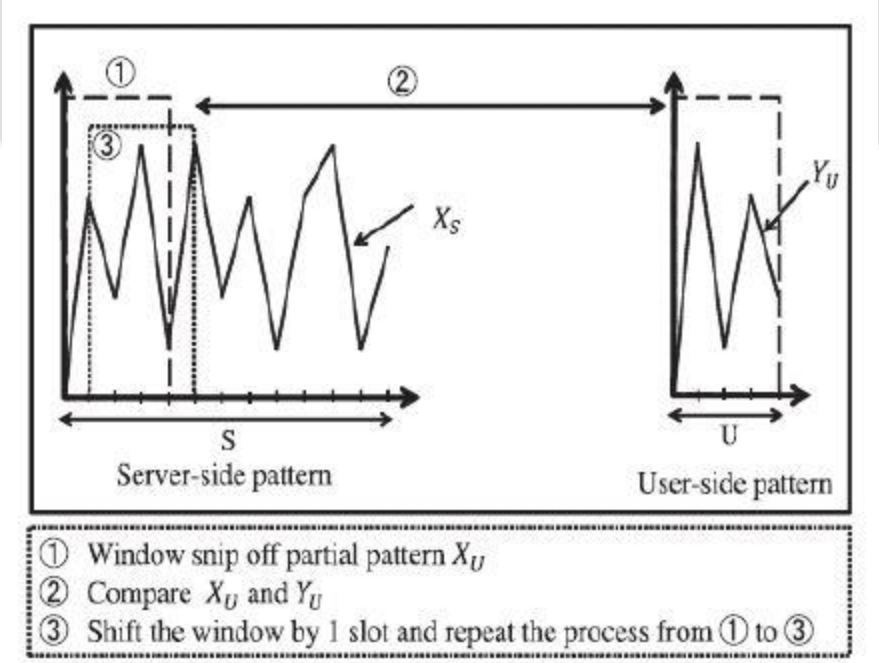

Chart -2: Traffic pattern matching.

3. DETECTION TECHNIQUE TO HANDLE DATA CONTENTS OF DIFFERENT LENGTHS

The conventional methods, DP-TRAT shows high robustness to packet delay, jitter, and packet loss. However, the existence of data of different lengths subjected to time variation in real content delivery environment causes DP-TRAT's accuracy to decrease. In this section, we take a look at the issue caused by the existence of different length data in network environments.

3.1 Issue due to different lengths of data

Traffic patterns of Contents represent the carrying their characteristics, and are unique per content. Therefore, the longer the traffic pattern is, the more information on the data it displays. In conventional methods, it is assumed that a certain length of content can always be obtained through the network for all contents. Therefore it is possible to utilize a fixed decision threshold in both PTRAT and DP-TRAT methods. However, there is no such guarantee in actual network environments.

3.2 Leakage detection

We compare the target traffic pattern to the original traffic pattern, and we adjust the obtained degree of similarity using the approximation curve, where x is the size of the target traffic pattern. Finally, we compare the adjusted degree of similarity to the decision threshold specific to the original data, and detect whether or not there is a leakage. The problem in comparison of short length data is then solved, and a flexible and accurate leakage detection method is possible.

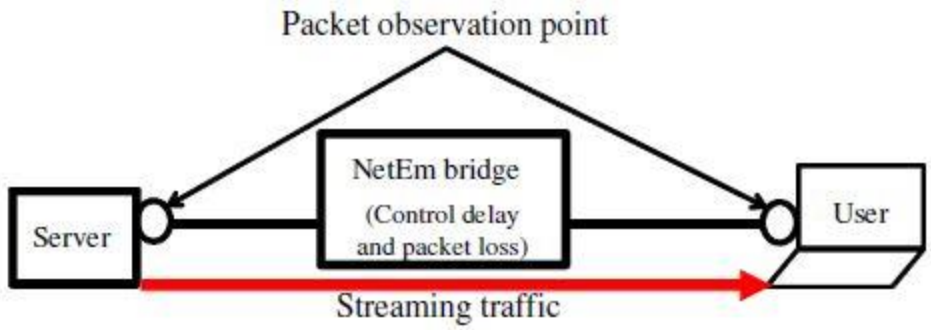

Chart -3: Topology of our conducted experiment.

4. CONCLUSIONS:

The content leakage detection system based on the fact 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.

5. ACKNOWLEDGEMENT:

We are greatly indebted to our Prof. S. G. Joshi, Head of the Department, Vishwabharti Academy's College of Engineering, Ahmednagar, for permitting us to do this paper. We would like to convey our heart-full thanks to our guide Prof. V. V. Jagatap and co-ordinator Prof. Natikar S. B., for their guidance and support in every step of this paper. We would like to convey our thanks to all the faculty and friends who have directly or indirectly helped us for successful completion of this paper.

6. REFERENCES

[1]. "The Study of Streaming Traffic behavior," KKU Engineering Journal, vol.33, no.5, pp.541-553, Sept.-Oct. 2006.

[2]. "Advances in digital video content protection," Proc. IEEE, vol.93, no.1, pp.171-183, Jan. 2005.

[3]. "A Study on Streaming Video Detection using Dynamic Traffic Pattern," IEICE Transactions on Communications Japanese Edition), vol.J19-B, no.02, 2010.

[4]. "A Survey on Peer-to-Peer Video Streaming Systems," Peer-to-Peer Networking and Applications, vol. 1, no. 1, pp. 18-28, Mar. 2008.4