Traitor Tracing Technology of Streaming Contents Delivery using Traffic Pattern in Wired/Wireless Environments

Masaru Dobashi, Hidehisa Nakayama,
Nei Kato and Yoshiaki Nemoto
Graduate School of Information Sciences
Tohoku University
Sendai-shi, Miyagi, 980–8579, Japan

Abbas Jamalipour
School of Electrical and Information Engineering
The University of Sydney
Sydney NSW 2006, Australia

***Abstract*— Today, with the rapid advance in broadband technology, digital contents delivery applications have been used widely and the streaming technology has made the contents delivery more popular. Nowadays, there is high expectation on Digital Rights Management (DRM). Traitor Tracing is one of the DRM technologies and enables us to observe user's contents streaming and detect illegal contents streaming. However, malicious users can interrupt tracing with illegal processes at user-side computers. To prevent all illegal processes at the user-side, routers should analyze information embedded into packets, which is unrealistic. In this article, we propose a system to detect illegal contents streaming by using only traffic patterns which are constructed from the amount of traffic traversing routers. We also investigate a method to cope with random errors and burst errors which occur frequently in wireless environment and show the satisfactory result which we have obtained in a practical testing environment.**

I. Introduction

In recent years, with the rapid advance in broadband technology, digital contents delivery applications have been used widely. The streaming technology has made the contents delivery more popular. Thanks to the widespread of wireless access, contents delivery in wireless environment has been actively studied [1], [2]. Nowadays, there is high expectation on the Digital Rights Management (DRM), which has been studied at many institutes [3], [4]. Traitor Tracing is one of them. It enables us to observe users' contents streaming and to detect illegal contents streaming [5], [6].

In general, Traitor Tracing systems embed copyright notices and server-side information into contents, using the digital watermark technology [7]–[9]. Recently, the method which use cryptographic keys to detect illegal transfers of contents was also proposed [10]. However, because those methods depend on information embedded into contents, malicious users can interrupt tracing with illegal processes at user-side computers [11]. To prevent all illegal processes at user-side computers, routers should analyze information embedded into packets, which is unrealistic because it needs very high computation. Furthermore, observing information in packets may cause problems from the perspective of the protection of personal information. Therefore, we should examine the Traitor Tracing method which does not use embedded information.

Using Variable Bit Rate(VBR), the movie's bit rate changes according to the change of the scene. When these movies are delivered and played as contents by streaming, the changes of the amount of traffic will appear as a unique waveform on the contents. This server information-based waveform can be related to user information which is independent of the movie's contents. Therefore, by matching the waveform at the server-side and the waveform at the user-side, we can detect the reception of the contents. If there is a user who receives the contents without permission, we judge that this is an illegal reception.

Accordingly, in this paper, we propose a method to detect an illegal streaming using only the amount of traffic which is observed in a short period (approximately $20[s]$) from the routers located just before the users. This system can be operated with less processing time than packet analysis. Furthermore, it can prevent illegal processes of malicious users efficiently, because no process is needed at user-side computers. In this paper, we also investigate a dynamic determination method and a matching method for impaired waveforms, taking into account for wireless environment where random errors and burst errors occur.

II. Detection System for Streaming of Contents

The outline of the proposed scheme is shown in Figure 1. In this figure, the Contents Server distributes contents, each user receives contents and each Router observes the amount of traffic. This information is sent to the Management Server which has an authorized users list, a list of users allowed to receive contents. The Management Server constructs server-side and user-side traffic patterns from information about the amount of traffic and matches patterns with each other. With matching results at the Management Server, contents streamings of users are detected. Here, we assume that both U1 and U2 are receiving contents but only U1 was an authorized user who exists in the list. In this case, U2 would be considered to be an illegal user receiving contents without legitimate

1-4244-0357-X/06/$20.00 ©2006 IEEE
This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE GLOBECOM 2006 proceedings.

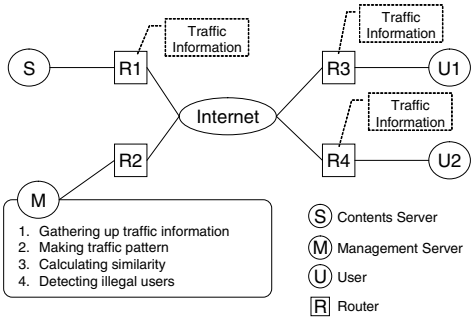

Fig. 1. Proposed Detection System

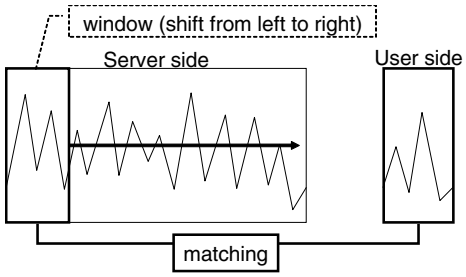

Fig. 2. Mechanism of Traffic Pattern Matching

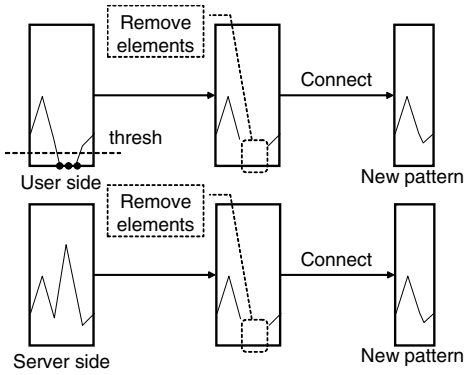

Fig. 3. Traffic Pattern in Wireless Environment and Measure against Burst Error Losses

authentication. These are the outline of the proposed system, which detects illegal receptions in streaming contents delivery.

In this section, we propose and explain the method to detect streaming of contents by observing the amount of packets at routers.

A. Traffic Pattern

In animated contents (MPEG), the bit rate is automatically adjusted to the changes of the scene. Each content is considered to have its own characteristic feature just like fingerprint, therefore, the unique information of these contents appear in waveforms. In this paper, we focus on VBR traffic, which is a typical type used in contents delivery, and call these waveforms "*traffic pattern*". Here, each content is distributed independently of each streaming server.

Traffic pattern is defined as the amount of traffic for one time-slot, a certain period of time $\Delta t[s]$ and expressed for N dimension in the following expression.

$$\boldsymbol{X} = \left(x_1, x_2, \cdots, x_N\right)^t, \quad T = N\Delta t$$

Here, $T[s]$ is the whole length of traffic pattern.

We pay attention to the similarity between a certain user-side pattern Y_U (U-dimension) and a part of server-side pattern X_U (U-dimension), and use cross-correlation coefficient as a criterion to judge the similarity of the two patterns. More specifically, we find the mean $\bar{x}, \bar{y}$ and the standard deviation s_x, s_y of each vector, and calculate the cross-correlation coefficient R_{XY} using the following equation.

$$R_{XY} = \frac{\boldsymbol{X_U^{'}}^t \boldsymbol{Y_U^{'}}}{\sqrt{||\boldsymbol{X_U^{'}}||^2 \cdot ||\boldsymbol{Y_U^{'}}||^2}} \tag{1}$$

where $-1 < R_{XY} < 1$ and $\boldsymbol{X_U^{'}}, \boldsymbol{Y_U^{'}}$ are the normalized traffic patterns when the $mean = 0$, and $variance = 1$.

$$\boldsymbol{X_U^{'}} = \begin{pmatrix} (x_1 - \bar{x})/s_x \\ (x_2 - \bar{x})/s_x \\ \vdots \\ (x_U - \bar{x})/s_x \end{pmatrix}, \boldsymbol{Y_U^{'}} = \begin{pmatrix} (y_1 - \bar{y})/s_y \\ (y_2 - \bar{y})/s_y \\ \vdots \\ (y_U - \bar{y})/s_y \end{pmatrix}$$

R_{XY}'s value would be near to 1, if two vectors were similar to each other.

B. Comparison of Traffic Pattern

As shown in Figure 2, the detection of illegal reception of contents is conducted by pattern matching between user-side patterns and server-side patterns. If a user-side pattern is a partial pattern of the server-side's one, this user is considered to be receiving the contents. To find such a matched pattern, we first set the window (the rectangular box within the Server-side pattern in Figure 2), which contains the same number of time-slots as the User-side pattern. Next, the Transform Process in Figure 3 is conducted to prevent influences of burst errors in wireless environment.

In the Transform Process, first, vector's elements whose values are equal or less than a certain threshold T_P are removed from the User-side pattern Y_U (U-dimension) and new User-side pattern $Y_U^{'}$ ($U^{'}$-dimension) are constructed. For example, three elements are removed in Figure 3. Next, the same part of the Server-side pattern's elements as the User-side pattern is also removed and new Server-side pattern $X_U^{'}$ ($U^{'}$-dimension) is constructed. In Figure 3, three corresponding elements are removed. After the Transform Process, cross-correlation coefficient R_{XY} is calculated with Equation (1).

After these, sliding the window from left to right is done by one slot and the whole server-side pattern is scanned. We repeat the extraction of pattern X_U (U-dimension) from server-side pattern X_S (S-dimension), the Transform Process and also calculate the cross-correlation coefficient. If whole sever-side pattern X_S had S-dimension and user-side pattern Y_U had U-dimension, the number of the calculation would be

1-4244-0357-X/06/$20.00 ©2006 IEEE
This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE GLOBECOM 2006 proceedings.

$S - U + 1$ times.

If a large value exists in cross-correlation coefficient graph, it means that a certain user-side pattern is similar to the part of the server-side pattern and such a pattern is called a "matched pattern". In this case, the user is considered to be receiving contents. In following section, we explain about a threshold to decide whether a matched pattern exists or not.

C. The Method of Dynamic Determination of Discrimination Threshold

Compared to wired environment, in wireless environment, packet losses occur more frequently due to random errors and burst errors. This causes user-side traffic patterns to be impaired and the overall value of the cross-correlation coefficient to be decreased. Thus, the detection would be almost impossible, if we set the discrimination threshold value to be fixed, for example $T_R = 0.9$ (the value that is generally considered to be high correlation). For this reason, we introduce a method to dynamically determine the discrimination threshold value T_R from the statistical tendency of the cross-correlation coefficient. In this way, we can flexibly cope with the network state.

Most parts of $S - U + 1$ cross-correlation coefficients have small values, whose distribution is considered to be the normal distribution, since the distribution of cross-correlation coefficients of two different waveforms is approximated to the normal distribution [12]. However, the cross-correlation coefficient of two patterns similar to each other has a large value compared to most other values and such a large value is called as an "outlier". The part which should be detected is one part among $S-U+1$ parts and such a part is statistically considered as an outlier. The mathematical way of detecting outlier is based on the mean and variance is as follows. Let μ_R be the mean and σ_R be the variance of cross-correlation coefficient, the discrimination threshold value T_R is defined as

$$T_R = \min(\mu_R + 4\sigma_R, \quad 1.0) \tag{2}$$

According to the Chebyshev's inequality, the probability that the data are equal or greater than the threshold value T_R is about $6[\%]$. As a result, the coefficient greater than T_R calculated from equation (2) is regarded as one outlier.

Figure 4 shows the proposed determination algorithm which can be applied in both wired and wireless environments. Therefore, without taking into account which environment users belong to, we can detect an illegal streaming by using this algorithm.

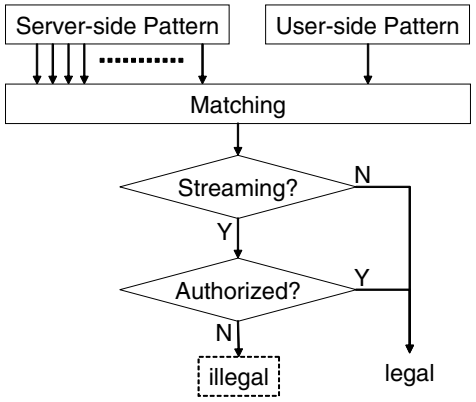

Fig. 4. Algorithm of Discrimination

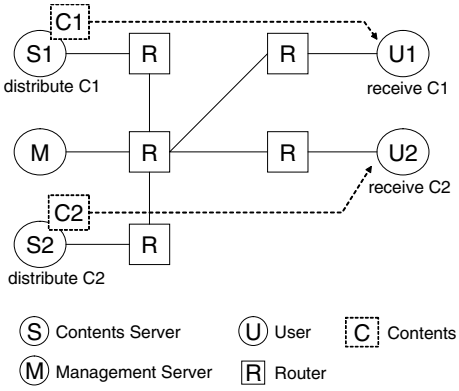

Fig. 5. Topology

TABLE I
COMPUTER CONFIGURATION

Device	Specification
CPU	Intel Pentium4 2.4GHz
Memory	256Mbytes
OS/Kernel	Fedora Core 3/Linux2.6.11
NIC	Intel82801DB PRO/100 VE(CNR) Ethernet Controller
NIC driver	e100 3.3.6-k2-NAPI

III. EXPERIMENT OF DETECTION USING TRAFFIC PATTERN

A. Experimental Setup

We performed an experiment using two animated contents encoded in MPEG4's VBR formula to verify that streaming contents are detected with the proposed method. The topology of the experiment and the specification of PCs is shown in Figure 5 and Table I. In Figure 5, Contents Server S1 and S2 distributed Contents C1 and C2 independently of each other. User U1 and U2 received C1 and C2 respectively. Here, we set one time-slot as $1[slot] = 0.2[sec]$ and the threshold in section II-B as $T_P = 2000[bytes]$. The lengths of C1 and C2 are both $900[slot]$ $(= 180[sec])$ and the lengths of user-side patterns are both $100[slot]$ $(= 20[sec])$. Management Server M constructed traffic patterns from information about the amount of traffic and calculated cross-correlation coefficients. In this experiment, no huge background traffic existed except for packets sent to control networks. In following paragraph, we show the scenario of this experiment.

First, the amount of traffic at each router nearest to S1, S2 and U1 were observed and sent to M. Next, traffic patterns were constructed and cross-correlation coefficients between traffic patterns of S1 and U1, and between traffic patterns of

1-4244-0357-X/06/$20.00 ©2006 IEEE
This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE GLOBECOM 2006 proceedings.

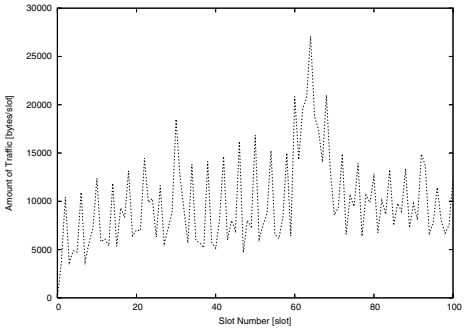

Fig. 6. Traffic Pattern of U1 (Wired Environment)

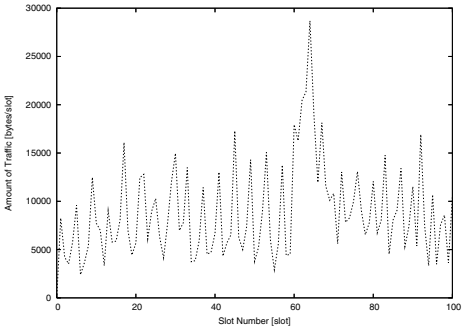

Fig. 9. Traffic Pattern of U1 (Random Error)

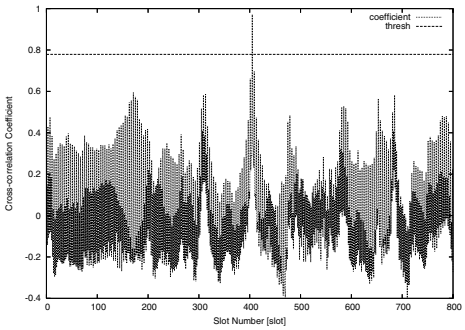

Fig. 7. Correlation between S1 and U1 (Wired Environment)

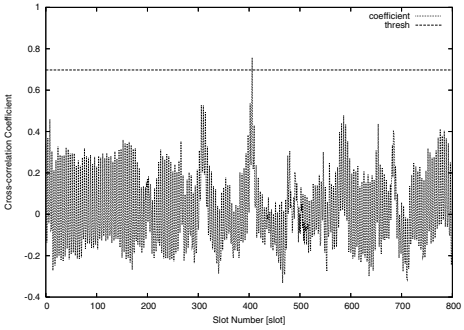

Fig. 10. Correlation between S1 and U1 (Random Error)

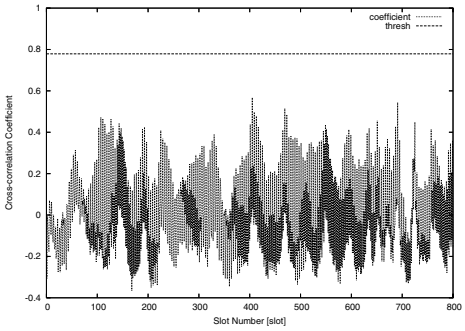

Fig. 8. Correlation between S2 and U1 (Wired Environment)

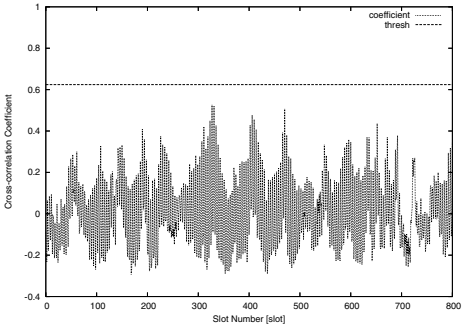

Fig. 11. Correlation between S2 and U1 (Random Error)

S2 and U1 were calculated.

We performed an experiment in pseudo-wireless environment, where random errors or burst errors are generated with software as well as in wired environment. The error rates of random errors and burst errors were set to $10[\%]$ and $30[\%]$ respectively.

B. The Experiment in Wired Environment

Figure 6 shows the traffic pattern of U1 in wired environment. Figure 7 and Figure 8 show the cross-correlation coefficients between traffic patterns of S1 and U1 and between traffic patterns of S2 and U1 in wired environment. In Figure 7, we can see a part greater than threshold T_R while there is no part greater than threshold T_R in Figure 8. As a result, we can say that the proposed system correctly detected whether a user was receiving certain contents in wired environment.

C. The Experiment in Wireless Environment

First, we performed an experiment, generating random errors. Figure 9 shows the traffic pattern of U1 in random error environment. Figure 10 and Figure 11 show the cross-correlation coefficients between traffic patterns of S1 and U1 and between traffic patterns of S2 and U1 in random error environment.

Next, we performed an experiment, generating burst errors. Figure 12 shows the traffic pattern of U1 in burst error environment. Figure 13 and Figure 14 show the cross-correlation coefficients between traffic patterns of S1 and U1 and between traffic patterns of S2 and U1 in burst error environment.

In both environments, the proposed system correctly detected whether a user is receiving certain contents, with the Transform Process in section II-B and the dynamic threshold T_R in section II-C.

1-4244-0357-X/06/$20.00 ©2006 IEEE
This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE GLOBECOM 2006 proceedings.

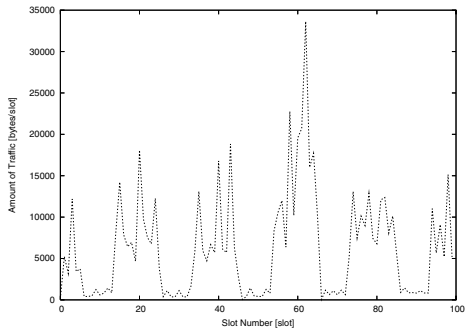

Fig. 12. Traffic Pattern of U1 (Burst Error)

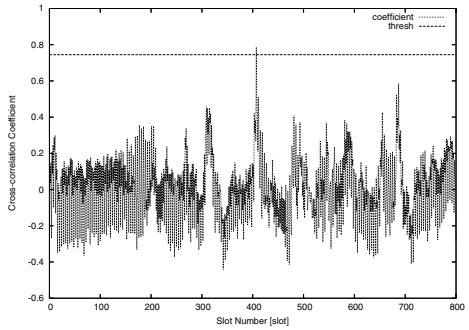

Fig. 13. Correlation between S1 and U1 (Burst Error)

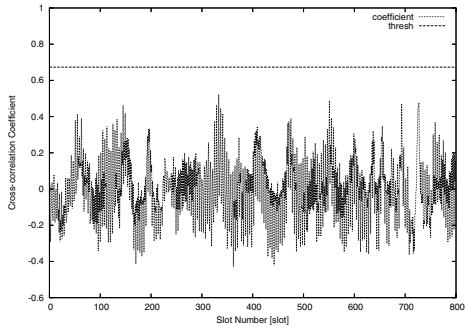

Fig. 14. Correlation between S2 and U1 (Burst Error)

IV. Consideration

A. Delay

The main problem in real-time delivery of movies is the delay due to the huge size of data. The traffic pattern in this case becomes the parallel displacement of the traffic pattern when there is no delay. However, even if there is a difference in detecting position, the characteristic of waveform would not be impaired. Therefore, delay should have small impact on our detection system.

B. Background Traffic

In general, there are not only contents traffic on the network, but also many other applications traffic and other users' traffic. However, to some extent, it is possible to extract the contents traffic from other users' one, using IP address information and packet type. Since the contents traffic has large amount of packets compared to other applications, we can apply our method similar to the method in section II-B and section II-C dealing with background traffics as random errors.

V. Conclusion

Nowadays, there is high expectation on the Digital Rights Management (DRM). The Traitor Tracing technology is one of the DRM technologies. Traitor Tracing enables us to observe user's contents streaming. However, malicious users can interrupt tracing with illegal processes at user-side computers. To prevent all illegal processes at the user-side, routers should analyze information embedded into packets but this is unrealistic.

In this paper, we proposed a system to detect illegal contents streaming by using traffic patterns which are constructed from the amount of traffic traversing routers and investigated a method to cope with random errors and burst errors. Finally, satisfactory results were shown.

We consider that simple delay in delivery of movies have small impact on our detection system. Moreover, even if there is a background traffic in the environment, we can apply our method effectively by extracting the contents traffic from background traffic with IP address information and packet types. To validate the proposed method, additional experiments using more contents and more complicated topology are planned.

References

[1] A. Toufik, M. Ahmed, and B. Raouf, "Interworking between sip and mpeg-4 dmif for heterogeneous ip video conferencing," Proc. of the IEEE ICC, vol.25, no.1, pp. 2469–2473, Apr. 2002.

[2] T. Liu, and C. Choudary, "Content-aware streaming of lecture videos over wireless networks," Proc. of the IEEE Multimedia Software Engineering, pp. 458–465, Dec. 2004.

[3] A. Seki, and W. Kameyama, "A proposal on open drm system coping with both benefits of rights-holders and users," Proc. of the IEEE Globcom, vol.22, no.1, pp. 4111–4115, Dec. 2003.

[4] F. Hartung, and F. Ramme, "Digital rights management and watermarking of multimedia content for m-commerce applications," IEEE Comm. Magazine, vol.38, no.11, pp. 78–84, Nov. 2000.

[5] A. Fiat, and T. Tassa, "Dynamic traitor tracing," Journal of CRYPTOLOGY, vol.14, no.3, pp. 211–223, 2001.

[6] R.S. Naini, and Y. Wang, "Sequential traitor tracing," IEEE Trans. on Information Theory, vol.49, no.5, pp. 1319–1326, 2003.

[7] K. Matsui, Basic knowledge of digital watermark, Morikita Shuppan, 1998.

[8] B. Turnbull, "Important legal developments regarding protection of copyrighted content against unauthorized copying," IEEE Comm. Magazine, vol.39, no.8, pp. 92–100, Aug. 2001.

[9] D. Kundur, and K. Karthik, "Video fingerprinting and encryption principles for digital rights management," Proc. of the IEEE, vol.92, no.6, pp. 918–932, June 2004.

[10] D. Boneh, and M. Franklin, "An efficient public key traitor tracing scheme," Advances in Cryptology - Crypto'99, pp. 338–353, 1999.

[11] S. Voloshynovskiy, S. Pereira, T. Pun, J. Eggers, and J. Su, "Attacks on digital watermarks: classification, estimation based attacks, and benchmarks," IEEE Comm. Magazine, vol.39, no.8, pp. 118–126, Aug. 2001.

[12] R. Duda, P. Hart, and D. Stork, Pattern Classification(2nd ed.), John Wiley & Sons, 2000.

1-4244-0357-X/06/$20.00 ©2006 IEEE