Efficient Traducer Tracing System Using Traffic Volume Information

K.V. Ramana Ph.D.,
Department of Computer Science
Jawaharlal Nehru Technological University
Kakinada, 533003, India

N. Praveen Kumar
Department of Computer Science
Jawaharlal Nehru Technological University
Kakinada, 533003, India

Abstract--- Many leading Broadband access technologies and their accessing abilities have the capability to meet the future requirements of the Broadband consumer. With the enormous growth in the broadband technologies, there is a need of applying simmering technology in many applications like video conference systems and content transmission systems. Streaming content enables users to get access the files quickly and not have to wait until the file is done downloading. Security remains one of the main challenges for Content Streaming. Digital Rights Management (DRM) system must be implemented to avoid content spreading and un-intentional content usage. Watermarking technology can also be used to implement the DRM system but it has its own limitations and attacks. A control method for the steaming content delivery is required to prevent abuse of the content. For this reason, authors have proposed a contended methodology that uses traffic volume information obtained from routers. Traducer tracing is one of the essential technologies that designs DRM systems, and empowers content distributors to notice and control content acquisition. This technology utilizes the main concept of traffic contours that helps to determine who is watching the streaming content and whether or not a secondary content delivery exists i.e., mainly used to determine whether the network is being traced out by the intruder or not.

Keywords- Content Streaming; Traffic Contours; Traducer Tracing; Digital Rights Management.

I. INTRODUCTION

There are many new technologies that use Local Area Networks and Internet to replace previous systems which use leased lines [1], [2]. Streaming technology can also be implemented to content delivery systems [3]. For example, Internet TV came into existence for telecasting content streaming [4] using broadband technology plays a major role in the evolution of streaming network infrastructure. The most important feature of this kind of system is that a person can obtain the information very efficiently and not to have to wait for the complete download of the required information. However, these systems need to be securely implemented. For this reason, an efficient DRM (Digital Rights Management) system has to be implemented [5], [6]. DRM incorporated system can efficiently manage the problem of content spreading and un-intended content usage. DRM technology is necessitated to avoid some threatened problems like possibility of existence of authenticity issues i.e., using content without the knowledge of Content holder and Content provider. DRM strategy incorporated system can efficiently manage the content transmission and users operations on the content in particular phenomena [7]. Watermarking [8] [9] and Traffic patterns are the two efficient techniques that can effectively drive the DRM system. However, watermarking has its own limitations.

In regards of protection towards the content, an effective DRM technology has to be implemented. For this, the content is encrypted and decryption keys are securely transmitted to users and they can use these keys to avoid data interception [10]. This type of protection [8] is unable to control secondary distributions of decrypted data [11]. An efficient management technology is implemented by the content providers to manage and monitor the use of content [12], [13]. This technology is called Traducer-tracing technology which provides flexibility for the content providers to manage the data.

Traducer tracing system stands for incorporating unique identity to the propagating contents by embedding watermarks into them [14], [15]. Use of watermarking concept for unique content identification involves much complexity in terms of computations. Major drawbacks in using water marking concept are:

1. System that incorporates water-marking embedded techniques need surplus amount of computations to encode complex contents. This result in difficulty for evaluating the computational cost related to real-time streaming.
2. The major lapse of watermarking, content conversion, and the known removal attacks [10], [16] has become a serious issue.

Thence, an efficient system must be introduced so that it enhances the content distribution to explore the process of finding the narrow list of users who may be traducers. In order to avoid the user’s intervention in the normal functioning of the system, an effective monitoring process should be implemented based on the information obtained at the routers in the middle of the streaming path. Process includes gathering information about traffic amounts and using this information to track user’s contents reception. This content is helpful for the content providers to lessen the list of traducers and by using this concept, content provider can efficiently trace out whether the desired content is received by the user or not i.e., he can be able to know whether the user is watching the content or not and also can be able to determine whether the secondary content distribution exists or not [17], [18]. Since, the proposed concept is not using the packet information, privacy issues of the application is least bothered [19].

The Content is streamed with the Variable Bit Rate. Bit rate of this streaming content varies in accordance with the variations in the content. During the process of content delivery to the users from the servers, routers present in the middle of the streaming path are configured in such a way that they are capable to observe the streaming content and can generate unique contours associated with the traffic amount. It is possible to determine whether the targeted content is being watched out by the user or not, during the process of content distribution, only by comparing the user-side contours to the server side contours.

The following is a concise procedure to determine whether the user is watching the targeted content or not. Server side traffic contours are generated with the help of traffic volume information, traffic amount, obtained at the router near the content server. User-side contours are generated using traffic amount obtained at the routers near the user. These two obtained contours are compared by the management server, to decide whether the user is watching the content or not. This process does not require any operations on the user’s computer. On a theoretical basis, this mechanism makes impossible for a user to tamper with the tracing process. The major benefit of the proposed mechanism is that, it can be implemented with the lower computational cost since it uses only traffic volume information but not using mechanisms, like water marking which consumes higher computational costs.

A Traducer tracing technique is said to be Network Based Traducer Tracing Technique, when it is applied in the network comprising of Servers and Users. In this paper, the details of the proposed method, and evaluating it by creating an effective environment, using an efficient Simulator called Network Simulator (NS2) is shown.

The rest of this paper is organized as follows. In Section II, Overview of DRM Technology, and the research works on these technologies are discussed. In Section III, details on technologies that construct the proposed Traducer tracing system is presented. In Section IV, the results of simulations with Network Simulator 2 (NS2) are shown. Finally, Section V concludes the paper.

II. RELATED WORK

There have been many tremendous works done in exploring DRM technology [4] and techniques that come under DRM technology. Traditional systems make use of classical technologies like Encryption [20], [21] and Access-conditioning [22] for driving DRM incorporated systems. There are many crucial technologies that construct DRM system. Traducer tracing system is one among them. With this technique, content provider is able to monitor content usage and also able to know whether the network is being traced out by the intruder or not i.e., whether the user is appropriately using the content or not.

A Simple Traducer tracing system is depicted in fig.1

Figure 1: Example of Simple Traducer tracing.

The working procedure of the mentioned system is described as below:

- Unique information is embedded into the Content using Digital watermarking, by the content provider and generates copies of the content.
- The generated copies are propagated to different users.
- Application running at user analyses the content data and re-assembles the embedded information.
- According to the analyzed and extracted information, user’s application notifies that he or she is watching the content.

The following figure shows a scenario of a network comprises of a Content Server, Users and the two contents, which are made unique by using the concept of Digital Watermarking, and depicts a mechanism to find out whether secondary distribution of content exists or not and to trace out the originator or run-off source of the secondary content.
The following is the mechanism to find out whether the secondary content distribution exists or not.

- Content server $S$ delivers a unique content $C_1$ to $U_1$ and the user application at $U_1$ analyzes the Watermarked information, $W_1$, and notifies it to $S$.
- In case of existence of unauthorized content distribution between user 1 and user 2, $W_1$ also found from $U_2$’s content and application running at $U_2$ notifies $W_1$ to $S$.
- Finally, $S$ finds duplicated $W_1$ and concludes that there is an existence of secondary content distribution says $C_1$ in detail $W_1$, in between $U_1$ and $U_2$.

There are two main issues that are associated with the above mechanism. They are

1. It includes lot of computational costs in encoding the content and also in embedding the watermarks, since an individual content has to be propagated to each user. Let ‘$C$’ and ‘$W$’ be the costs involved in encoding the content and embedding watermarks into it. Then, there will be at least $N(C+W)$ total cost involved in propagating the content, which has the traceability of traducers. ‘$N$’ is the number of content users. This does not suit better for the real-time streaming contents. That is why an efficient method is needed to track the traducers with minimum cost.

2. Second issue is the use of Watermarking concepts for tracing traducers and for embedding information into the host signal has its own limitations. There are known attacks against Watermarks such as Copy attack, Collusion attack, Removal attack and Sensitivity attack [23], [24].

Watermarking concept does not go forward or does not suit in the field of complex networks. Since, it shows problems because of unclear network environment and users.

Hence, Traducer tracing systems with Watermarking need additional mechanisms to tighten the scope of application by shortening down the possible list of traducers. A convenient and the most efficient method to track the content stream without using the Watermarking technology is by the use of Traffic Volume Information obtained at the routers present in the middle of the streaming path. This method avoids the need of decrypting and decoding the streaming content.

In Section III, an efficient system is proposed to trace a multimedia streaming content.

### III. METHODOLOGY

#### A. Modules

The proposed work has been segmented into three phases. The first phase deals with building network environment. The second phase establishes communication among the hosts and servers. This phase also deals with the re-configuration or re-construction of new paths in case of path failures. The third phase will explain about the mechanism to detect the flow of streaming content in variety of networks i.e., to find out whether the user is watching the streaming content or not. This is an effective mechanism find out whether the network is being traced out by the intruder or not.

Figure 3 represents the flow chart of the modules and their description was in below sections.

1) **Build network:**

In the first phase, a system is proposed by considering a network comprises of Servers, Clients (Users), Intermediate Routers [24]. Servers are configured to run two types of applications. One server is configured to implement effective content delivery. Content Server and the other, is configured to perform vital functionalities in detecting the traducer in the network, Management Server. There are intermediate routers at the side of each communicating party that have a capability to observe the traffic amount and propagate the obtained information to the Server that manages the built network.

2) **Establish Communication:**

In the second phase, a communication among the network entities has been established. The content has to be streamed from the Content Server to different Clients say Users, via intermediate routers. Content Servers make use of DRM technology in generating unique content and propagating it to different users.
3) Core Module of the Application:

In the third phase, a concise review of the proposed system is introduced. The framework of the proposed strategy is shown below.

![Flow chart to implement traducer tracing system using Traffic volume information.](image)

Figure 3: Flow chart to implement traducer tracing system using Traffic volume information.

![Framework of the proposed strategy.](image)

Figure 4: Framework of the proposed strategy.

Figure 4 represents a Network of Servers, Users and Routers which has the capability to drive the proposed mechanism.

The mechanism of the contended methodology is as follows:

Content Server spreads the content C and the users U1 and U2 receives the content. Routers in the middle of the streaming path say R1, R3 and R4 observe the traffic amount. For instance, Router R1 supervises the flow of content server S and mass up the sizes of the packets. Information obtained at the routers is sent to the Management Server using protocols such as ICMP and SNMP.

Management Server constructs the traffic contour from the traffic amount information obtained by the routers. Router R1 observes the content at the content server S during its propagation and transmits the observed data to the Management Server M, which generates the Server side traffic contour. Similarly, the routers at user side say R3 and R4 observes the traffic amount for few minutes and transmits the observed data to Management server that generates User side traffic contour. Transformations are performed on the obtained traffic contours to avoid breach errors. Management Server then performs comparison to find out the degree of similarity between the User-side and Server-side traffic contours. Based on the results obtained from the similarity matching, Content provider can able to know who is watching the streaming content. If the contours match, all the network behaviors are normal. If they don’t, the user is not watching the content i.e., some additional noise has been added to the streaming content which can be done by the interception of the participated communicating party. Thence, there is a possibility to say that there is a traducer in the network and by the mentioned mechanism, it is possible to decide that the network is being traced out by the intruder.

The suggested system has a capability to ascertain the flow of gushing content in different breeds of networks. The contended methodology is befitting for bountiful kinds of applications. Appliances include QuickTime servers administered by Apple. The recommended procedure is suitable for streaming content and is not applicable for the downloaded content since monitoring of the network can be done on the propagating content. The proposed system is brought to bear for a precise degree of confined networks. For instance, this totally suits for the video streaming systems using corporate LANS. With the contended mechanism, the executives of the streaming systems can effectively administer the propagating content and know who is watching them. This diminishes not only the aegis hazard on caper by the third party but also by the administrator.

B. Elucidation of Traffic Contours and Congruity

In this thesis, Variable bit rate simmering content is spotlighted and embrace of traffic contours to boast the streaming content continuance.

The traffic contour is elucidated as the heap of traffic for a certain period of time $\Delta t$ [sec] and expressed as an N-dimensional vector in the following expression.
\[ C = (c_1, c_2, \ldots, c_N)^T, \quad T = N \Delta t \tag{1} \]

Where ‘T’ in seconds is the length of the traffic contour and ‘N’ is the number of slots. \( c_1 \) is the traffic contour observed in first time slot and \( c_2 \) is the traffic contour observed in second time slot.

In the proposed strategy, while the network is in running state, router present at the server observes the traffic amount throughout the content transmission and transmits this information to obtain the Server side traffic contour which is expressed as

\[ C_S = (c_1, c_2, \ldots, c_s)^T, \quad \text{according to (1)} \]

User side traffic contour is expressed as

\[ C_U = (c_1, c_2, \ldots, c_u)^T \]

In the above expressions, \( S \) and \( U \) are the number of time slots. The length of the Server side traffic contour or observation is greater than User side observation i.e., \( S > U \).

To find out the similarity of these contours, a partial pattern \( P_U \) of finite length say \( U \) (length of the User side contour) of the contour \( C_S \), is snipped off and is compared with the contour \( C_U \).

Before computing the congruity of two contours, \( P_U \) and \( C_U \) are normalized as,

\[ \begin{align*} 
\overline{P_U'} & = \frac{P_U - \overline{P_U}}{\sigma_P} \\
\overline{C_U'} & = \frac{C_U - \overline{C_U}}{\sigma_C} 
\end{align*} \]

In the above expression, \( \overline{\cdot} \) and \( \sigma \) are the means of each vector, \( \sigma_P \) and \( \sigma_C \) are the standard deviations. After normalizing, the means of \( P_U' \) and \( C_U' \) are zero and variances are 1. Finally, by obtaining the computed values, these values are used in the below equation to find out the similarity between the user side contour and the part of server side contour.

\[ R_{PC} = \frac{\overline{P_U'} \cdot \overline{C_U'}}{\sqrt{||\overline{P_U'}||^2 ||\overline{C_U'}||^2}} \tag{3} \]

When the two contours are in congruence, \( R_{PC} \) approximates to 1.

By using the above process, the similarity between the contours can be obtained only if there are of same length. As mentioned earlier, the length of the server side contour \( C_S \) is greater than the length of the user side contour \( C_U \). In this contended methodology, similarity between the contours of different length has to be computed. For this reason, a concept of “window” is used. Here, a part of window of length equals to the length of the user side contour is considered and is used to snip off the partial part \( P_U \), of the server side traffic contour. By using the partial contour \( P_U \), of the server side traffic contour, the contended strategy computes the similarity \( R_{PC} \) by moving the window from left to right on the server side traffic contour \( C_S \).

C. Confrontation of Traffic Contours using Similarity

The framework of comparison of traffic contours is explained in systematic way using the following three steps. In the first step, a window is considered that snips-off the partial contour \( P_U \) from the Server side contour \( C_S \). In the second step, similarity between the partial contour \( P_U \) and User side traffic contour \( C_U \). In this step, transformations are applied on the contour to avert from the consequences of burst errors. In the final step, window is advanced from left to right by one slot. After advanced to next slot, the mentioned three step process repeats and then applied to next slot, this process continues till the window arrives the rightmost component of the Server side contour. In total, \( S-U+1 \) values of similarity are obtained from the above computation. The following figure depicts the above mentioned process.

---

Figure 5: Framework for comparison of Traffic contours.

1) Allievation against Burst Errors

It becomes arduous to compute the similarity as burst errors vitiates the User side traffic contour. Transformations on the considered contours (partial contour and the user side contour) are performed to avert from the consequences of the burst errors. Due to the presence of the burst errors, the user...
side traffic contour moderately drops. The following is the
sequential procedure to transform the contours. A precise
value is computed which is referred to a threshold value and is
used as a reference to apply transformations. In the first step,
user side traffic contour is speculated and the values less than
or equal to the threshold are considered and are removed from
the contour. In the second step, partial contour is considered
and amputate its elements conferred to user side removal
routine. Characteristically, in the partial contour, the
components that are at the same position where the elements
of user side contour are removed, are taken out. In the third
step, the outlived components are consociated.

The above mentioned routine has to be implemented
before the similarity comparison mechanism invokes. The
traffic contours do not modify by this procedure, when there is
no burst errors. The accuracy and performance of the
contended mechanism depends on the appropriateness of the
Threshold value. To compute the threshold value, a
charismatic and dynamic procedure is used.

2) Active resolution of Threshold value

It is trivial to discover the apex value in the similarity
graph using static threshold determination procedure. Packet
loss compels the contended methodology. If packets drop
during the process of content propagation, then the user side
traffic contour gets decayed i.e., the larger value of similarity
becomes smaller ones. In order to overcome this problem, a
dynamic strategy is introduced to find out the threshold value,
which can be easily adapted to the altering network
environment. The similarity data that is obtained by this
contended methodology is small and it is normally distributed
around zero since, the cross-correlation values of the two
distinguished random waveforms are approximated to be
distributed normally. More similar patterns yield greater value
of similarity. That value can be termed as an “Outlier”. The
following equation is used to determine the threshold value
that helps in finding out the outlier among some values.

\[ T_C = \min (\mu_C + 4\sigma_C, 1.0) \]

In the above equation, \( \mu_C \) and \( \sigma_C \) are the mean and variance of the obtained similarity values. They are estimated based on
the analysis. Each and every user side contour has its own \( \mu_C \),
\( \sigma_C \) values. The coefficient of variance can be adjusted to
balance the trade-off between the detection ratio and false-
positives. Greater value of variance coefficient results in low
false positive ratio at low detection ratio cost. Accordingly,
the similarity value which is greater than \( T_C \) is considered as
an outlier, which is numerically distant from the outlived
values on the profile.

The computational cost needed to deploy the contended
methodology is less when compared with the traditional
mechanisms which include water marking and encoding
techniques. In the proposed technique, Management Server is
configured to compute the similarity of traffic signifiers. The
total number of multiplications computed in the proposed
system can better be explored in terms of three parameters.
They are length of the user side contour (U), length of the
server side contour (S) and the number of multiplications
performed to compute the measure of similarity (M). The
total number of multiplications are \((S-U+1) M\).

IV. RESULTS AND DISCUSSIONS

The network components are configured in such a way
that, they can perform their respective functionalities with
minimum computation costs. This result in the less
computational cost incurred in implementation of contended
methodology.

Below figure depicts the network model of ten components
and their connections.
Figure 8 shows the scenario of labeling the network participants involved to implement the contended methodology.

Figure 9 depicts a scenario of content simmering from content server to intermediate router at user (Router4) through intermediate router at server (Router1).

Below figure shows the immediate instance of the previous scenario.

Figure 10: Content streaming from server to destined user.
Figure 10 represents streaming of contents from content server to destined user via routers in the middle of the streaming path.

The contended methodology employed in the simulation makes the network acquiring the property of self-healing. Self Healing means system can automatically recover to stable state without external interference. In this aspect, even there is a link failure in the network, the system has a capability to re-establish the path. In the meanwhile, the proposed strategy finds the alternate path to the destined user.

Below figure shows the scenario of link breakage between the routers present in the middle of the streaming path.

Below figure show the re-establishment of original path.

Figure 11 shows the instance describing about the link failure between the routers in the middle of the streaming path.

Below figure shows the adaptation of alternate path by the contended methodology to avoid data loss to the destined user.

Figure 12: Adaptation of alternate path.

Figure 13: Re-Construction of original path.

Figure 13 shows the scenario of re-establishment of original path in order to avoid the situation of content loss at the destined user.

Below figure shows the traffic contour generated by the Management server using the traffic volume information obtained from the router at the server.

Figure 14: Server-side traffic contour.

Below figure shows the User side traffic contour generated at an instance by the management server using the traffic volume information obtained from the router at the destined user.
enable betterment in the security of the content delivery, an efficient mechanism is introduced which is based on traffic contours. To assess the accomplishment of the contended methodology, simulations were administered.

The contended methodology can also be incorporated with the previous DRM technology however, the computational cost involved in implementing the proposed strategy is less when compared with the costs involved in implementing watermarking and encoding techniques used in general traducer tracing techniques. The network components are configured in such a way that, they can perform their respective functionalities with minimum computation costs. This result in the less computational cost incurred in implementation of contended methodology. The proposed concept is not using the packet information, privacy issues of the application is least bothered.

REFERENCES


V. CONCLUSION

To prevent exploitation of the content, a subjugate approach for simmering content delivery is required. For this purpose, traducer tracing technology is introduced. Conventional techniques to trace out the traducers involve high load and complex computations to produce multiple contents and watermarking has its own known limitations. To


