A WEB MULTI-TIER PLATFORM FOR ADAPTIVELY PROTECTING AND SECURELY DELIVERING MULTIMEDIA CONTENTS ON THE WEB

M. Ceccarelli[†], *M. Di Santo*[‡], *S. D'Onofrio*[‡], *F. Frattolillo*[‡]

[†]Department of Biological and Environmental Studies - [‡]Department of Engineering
Research Centre on Software Technology, University of Sannio, Benevento, Italy

ABSTRACT

The advances in multimedia technologies have created opportunities for Internet pirates, who can copy digital documents and illegally distribute them, thus violating the legal rights of document owners or web content providers. This paper describes a web, multi-tier platform for adaptively protecting and securely delivering multimedia contents in the Internet. The platform exploits web services technologies in order to comply with a distributed, component-based approach. In particular, it provides two main services: an adaptive, "on buyer" watermarking service and a SIP-based session management service. Thus, a content provider that exploits the proposed platform can offer advanced multimedia services in a security context.

1. INTRODUCTION AND MOTIVATIONS

Digital watermarking [6] is one of the security techniques that can be used for discouraging the unauthorized trading of multimedia documents in the Internet. It allows the perceptually invisible digital signature of the copyright holder to be embedded in the documents to be protected, and this enables the copyright holder to prove the ownership of the watermarked documents by extracting the signature.

The digital signature inserted in a protected document made available on payment allows web content providers (CPs) to potentially establish if a user is illegally in possession of the document. In addition, if the watermarking procedure is also able to generate signatures tied to both the copyright holder and document buyer, whenever an illegally owned copy of a document is found, the copyright holder can always potentially prosecute both the illegal owner and who has started the unauthorized sharing of the document, that is, the original buyer.

Since a robust watermarking can be computationally intensive or increase the size of the protected contents, it is important to adapt it to the characteristics of both the terminal used to open the required content and the transaction carried out between the user and CP. For example, a PDA or a mobile phone with no storing capacity or limited visualization capacities could receive "lightly watermarked" multimedia contents during transactions taking place on low-bandwidth networks. Therefore, an advanced watermarking procedure should exhibit an "adaptive" behavior, that is, it should take into account the identity of buyers and the characteristics of terminals and of the network transactions carried out to transfer protected contents. However, CPs often have neither the competence nor the economical advantage to directly apply effective watermarking procedures to all the distributed multimedia documents. On the other hand, the core business of web service providers (SPs) is just to supply specialized back-end software services to CPs. However, this is possible only if the integration of back-end and front-end services accessible from a variety of terminal types can be assured without requiring a tight coupling among different web entities, such as CPs end SPs. To this end, web services technologies can be used to simplify the process of dynamically integrating elemental system and business services into more complex customer services.

In this paper a web, multi-tier platform for adaptively protecting and securely delivering multimedia contents in the Internet is described. The main goals in devising the platform have been: (1) to make the platform modular and extensible; (2) to create a security context for all the transactions that take place among end users and the service entities composing the platform; (3) to make an adaptive, "on buyer" watermarking procedure available to CPs as a security service outsourced to SPs; (4) to show that CPs, adhering to the service model proposed for the platform, can supply advanced services without having to change their original role and main applications.

The literature is rich of proposals, but very few watermarking procedures exhibit both an adaptive and on buyer behavior. Moreover, the proposed interpretation of the procedure described in [3, 4] has been integrated in a web platform that, differently from others, such as MOSES [5], allows distinct web entities to dynamically interoperate in a security context without having to make the provided services compliant to complex frameworks.

The paper is organized as follows. Section 2 describes the platform architecture. Section 3 describes the interaction model adopted by the platform. Section 4 describes the watermarking procedure implemented as a web service. Section 5 reports first experimental results essentially focused on the watermarking procedure. Finally, section 6 reports conclusion remarks.

2. THE PLATFORM ARCHITECTURE

The proposed platform, whose architecture is sketched in Figure 1, consists of two main parts: the former includes the web servers of CPs and represents the "front-end" tier of the platform; the latter, which represents the "back-end" tier of the platform, is composed of the web services implemented by SPs. In particular, in the proposed architecture an SP does not directly expose the web services that it supplies, but hides them behind a "dispatcher", which acts as

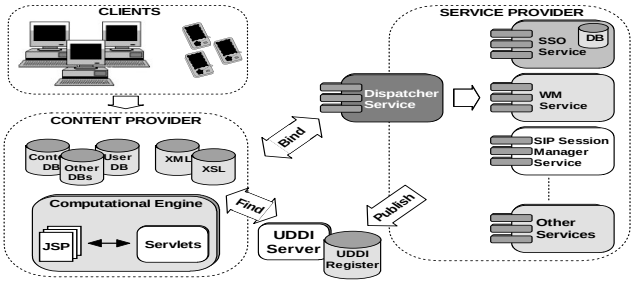


Fig. 1. The service platform.

a unique interface towards CPs for all web services implemented. Such an interface, designed itself as a web service, takes charge of receiving the service requests from CPs and dispatches them to the actual web services.

The choice of implementing an SP as a dispatcher and a set of web services hidden from users and CPs is motivated by the following considerations: (1) the dispatcher can act as a proxy for the web services; (2) a "single sign-on" (SSO) service can be exploited to control the interactions to and among the web services. In fact, the presence of an SSO service, which can itself be structured as a web service within the back-end tier, enables the dispatcher to sign-on only once in order to gain access to all the web services implemented by the SP. Therefore, even though the adoption of a dispatcher might represent a bottleneck for the back-end tier, the presence of a unique point of authentication/authorization enhances the security and increases the performances of the back-end tier, because the security credentials of the dispatcher do not need to be communicated to web services each time it wants to access them.

3. THE INTERACTION MODEL

Figure 2 shows the interaction model assumed by the platform, in which the CP allows users to access its web servers via terminals with a varying degree of functionalities. It exposes a "registration" service finalized both to acquire user information and to implement access control, user tracking and billing. The registration phase and all the subsequent communications involving the exchange of private information, such as payment, take place over SSL/TLS channels.

When a registered user chooses a content, a servlet running on the CP server decides if to protect it. If the content is, for example, an MPEG video, it may decide to watermark it. Then, the CP can directly contact an SP that offers a watermarking procedure for that video type or search UDDI registries to discover the demanded service.

To watermark the selected video according to the procedure proposed and described in section 4, the following steps are to be taken: (1) the CP sends the SP the video and some other information, such as the user profile, his terminal type, and the characteristics of user network connection; (2) the CP generates some data that it communicates both to the SP and the user, and that will be then exploited by the user to issue a challenge to authenticate the SP from which it will download the protected video; (3) the SP dispatches the video and all the information received from the

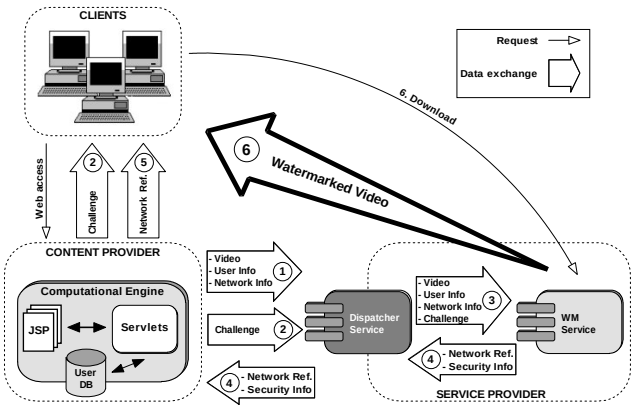


Fig. 2. The interaction scheme.

CP to the web service implementing the watermarking; (4) the web service runs the procedure and, as soon as it has watermarked a sufficient part of the video, returns its network reference and information about the watermarked inserted into the video to the dispatcher, which forwards them to the CP; (5) the CP communicates the network reference of the web service to the user and allows him to download the protected video; (6) the user authenticates the web service and starts the download. It is worth noting that: (1) the communications taking place during these steps exploit SSL/TLS channels; (2) the information returned by the watermarking service to the CP, together with other information about the outstanding transaction, has to be tied to the user and saved into a user database; (3) to speed up the service, the download takes place directly between the user and the watermarking server, and the watermarking procedure exploits Java multithreading. In fact, the web service can start watermarking before receiving the whole video, and can also begin to return the video to the user before having completely watermarked it.

The steps described above refer to the delivering of a protected document to a user. In practice, users could require advanced services based on mobility or on interactive communications. To manage these services, a session management web service implementing the Session Initiation Protocol (SIP) is provided by the platform. In fact, SIP is considered as one of the most promising ways to manage sessions by which to distribute multimedia contents according to the characteristics of user terminals and needs and to deploy innovative multimedia services over the Internet. It has also the ability to track down and connect to a mobile user via "register/lookup" methods.

4. THE WATERMARKING PROCEDURE

The watermarking provided in the developed platform is inspired by the method proposed in [3, 4] and is intended for MPEG-2 compressed video streams. The procedure has been modified in order to make it both exploitable in a web environment as an "on buyer" service and adaptive with respect to the characteristics of user terminal and network connection. It has been chosen for three main reasons: (1) it depends on few parameters that can be exploited to im-

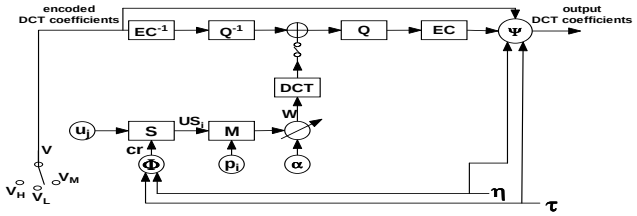


Fig. 3. The watermarking scheme.

plement the adaptive behavior of watermarking; (2) watermarking extraction does not require the original video, and this is a fundamental advantage for security in a web environment [7]; (3) both watermarking and watermarking extraction appear to be fast an reliable, since they directly operate in the compressed domain [2], and this can improve the quality of the service globally provided by CPs .

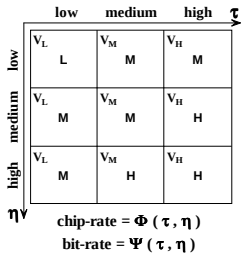
4.1. On buyer watermarking

The adopted procedure adds a noise-like signal to the encoded video signal processed block by block. To accomplish this, the approach from "direct-sequence spread spectrum" communications is followed [1].

To generate the watermark signal (see Figure 3) and implement the "on buyer" behavior, a sequence of bits identifying a user $u_j \in \{-1, 1\}$ is spread (S) by a large factor cr, called *chip-rate*, thus obtaining the spread sequence $us_i = u_j$, $j \cdot cr \leq i < (j+1) \cdot cr$. The sequence is then modulated (M) with a binary pseudo-noise sequence $p_i \in \{-1, 1\}$, yielding a noise-like signal which may be amplified with an amplitude factor α that can be varied according to local properties of the video. In particular, in the proposed solution, the sequence p_i has to be the same for each video made available by a CP. This way, once a protected video has been selected, it is possible to identify the pseudo-noise sequence p_i associated to it in the watermarking procedure. Then, to retrieve the watermarking, the video signal can be correlated with this sequence over a cr wide correlation window. The extracted watermark can be then analyzed to obtain the sequence of bits identifying the user who bought the video.

To implement the behavior described above, the sequences u_j and p_i have to be returned from the SP to a CP as "security info" whenever the watermarking is performed for the first time on a given pair (user, video) (see Figure 2). In practice, once a CP has stored the sequences u_j and p_i in its databases, it can send them to the SP as "user info" in the subsequent transactions involving videos that have been already protected or users who have already bought a video in previous transactions.

To insert the watermark in the compressed video V (see Figure 3), the encoded 8×8 blocks of the video are extracted and processed together with the corresponding blocks of the watermarking signal W. In particular, the MPEG-2 bitstream is split into its main components, and only the DCT encoded signal blocks are modified. Each encoded block is represented by a sequence of Huffman codes, each representing one (run-level)-pair and, thus, one quantized non-zero DCT coefficient of the current signal block. Therefore, to insert the watermark, each Huffman code is decoded



	low	medium	high
τ	$\leq (320\times240)$	(320×240) – (640×480)	$\geq (640\times480)$
η	modem links	DSL, LAN	T1, T3 lines

	L	M	H
Φ	10,000 – 100,000	100,000 – 500,000	500,000 – 1,000,000
Ψ	15 – 30 %	30 – 40 %	40 – 50 %

Fig. 4. The output of the functions Φ and Ψ.

(EC^{-1}) and inversely quantized (Q^{-1}), that is, the mapping from the quantizer index to the quantizer representative is performed. After this processing, a quantized DCT is added to the corresponding DCT coefficient from the transformed W signal, yielding a watermarked DCT coefficient. This is then quantized (Q) and Huffman encoded (EC).

Finally, it is worth noting that the scheme for drift compensation reported in [4] is not shown in Figure 3 for the sake of brevity.

4.2. The adaptive behavior

The adaptive behavior of the procedure is due to two main functions: Φ and Ψ (see Figure 3). These functions determine respectively the chip-rate cr and the video output bit-rate, both depending on user information, such as the characteristics of his/her terminal and network connection. To this end, the user terminal type is qualified by the variable τ, which essentially captures its visualization capacities (i.e. video resolution), while the characteristics of user network connection are synthesized by the variable η, which captures network bandwidth and latency.

τ is derived from what declared by users during the video selection web phase, while η can be directly estimated by a CP during the transaction web phase with users. Moreover, the procedure assumes that CPs send SPs videos whose quality already depends on the visualization capacities of the user terminal and on the applied billing. This is taken into account by assuming that the original video quality only depends on τ.

In Figure 4 three tables are shown. The first table summarizes the behavior of Φ and Ψ depending on τ and η. The characters L, M and H stand for "low", "medium" and "high" respectively, and identify the values of Φ and Ψ as reported in the third table. In particular, Φ and Ψ behave similarly (thus, their behaviors are shown in the same table), even though the values they assume in each interval are different for each function. The second table specifies when τ and η assume low, medium and high values, respectively. Moreover, in the first table, the quality of the video sent by a CP and depending on τ is also indicated: V_L, V_M and V_H stand for a low, medium and high quality video,

respectively.

The behaviors of Φ and Ψ have been determined taking into account the original philosophy of the procedure. In particular, a high value for cr increases the robustness of the watermarking, but at the same time decreases the data rate for watermark. On the other hand, controlling the bit-rate means to determine the fraction of the watermark signal that can be successfully embedded in the protected video: increasing the bit-rate means to increase this fraction and to improve the robustness of the watermarking.

To this end, in [3, 4] the watermarking is assumed not to increase the output bit-rate. On the contrary, in the proposed procedure, Ψ may increase the bit-rate, since the video size is assumed to change according to both the protection level required and the actual conditions of the service: the former is essentially identified by τ, while the latter are captured by η. Therefore, once the increment for a video size has been determined, Ψ sets a counter to the increment value. Then, Ψ updates the counter by subtracting from it the difference between the number of bits needed to represent a codeword for the watermarked signal sent to output and the number of bits used to represent the same codeword for the original video signal: positive differences are considered "debts", while negative differences are considered "credits". Thus, when the increment reaches 0, further codewords for the watermarked signal are sent to output only if further credits occur that balance debts. This way, the increment of the video size remains constant.

Thus, by combining the behaviors of Φ and Ψ, a user requiring a video from a terminal with limited visualization capacities or connected to a CP via a low bandwidth or high latency network receives a lightly watermarked, low quality video. On the contrary, a user provided with a high resolution terminal and a high performance network connection receives a strongly watermarked, high quality video.

In the proposed procedure, cr may vary in order to achieve the adaptive behavior. As a consequence, to extract the watermark from a video, it is necessary to have the associated sequence p_i as well as the value of cr used to watermark the video. To this end, watermarking is actually performed in two phases. In the former, a cr_v value constantly associated to the video is used to embed the first n values of the sequence u_i. These values are used to identify the chip-rate cr calculated by Φ and that has to be used to watermark, in the latter phase, the remaining part of the video. Thus, given the video, the sequence p_i and the value cr_v can be identified and then applied to retrieve the first n values of the sequence u_i, which identify the cr value to be used to extract the watermark from the remaining part of the video. Anyway, every time a new value of cr is used, the SP has to communicate it as "security info" to the CP, which can store it in its databases.

5. EXPERIMENTAL RESULTS

We have implemented a first release of the proposed procedure in Java. It implements a simplified version of the functions Φ and Ψ. In particular, Φ allows cr to assume only a few values for each interval reported in Figure 4, while Ψ allows 15–50% of the DCT coefficients to be altered.

We are working at an improved release of the procedure, in which a more defined model for Φ and Ψ is adopted. Some of the main goals are: (1) to better tune the behavior of Φ and Ψ according to the variables τ and η; (2) to define a model to correctly determine the values of τ and η according to the information provided by a CP; (3) to specify what information a CP has to assume from a user to enable an SP to correctly estimate τ and η.

Finally, it is worth noting that the watermarking scheme adopted can be considered secure against many common manipulations, and the addition of the adaptive behavior enhances the robustness of the procedure whenever the user is provided with a high quality terminal.

6. CONCLUSIONS

In this paper a web, multi-tier platform for adaptively protecting and securely delivering multimedia contents in the Internet has been presented. The platform exploits the web services technology and this allows SPs to make available complex services to CPs without requiring a tight coupling among different web entities or heavy changes in the core business of CPs. In addition, the platform has been designed to implement a high level of security in distributing multimedia contents. Finally, watermarking implements an adaptive, "on buyer" service, able also to protect a video according to some user information, such as the visualization capacities of user terminal and the bandwidth and latency of user network connection. This way, the watermark embedded in a video is tied to the user who buys it, and a trade-off can be achieved between protection needs and the computational power required to satisfying them.

7. REFERENCES

[1] I. Cox, J. Kilian, T. Shamoon, "Secure spread spectrum watermarking for images, audio and video", in *Proc. of IEEE Int. Conf. on Image Processing*, pp. 243–246, 1996, IEEE Press.

[2] T. Y. Chung, M. S. Hong et al., "Digital watermarking for copyright protection of MPEG-2 compressed video", *IEEE Transactions on Consumer Electronics*, vol. 44, n. 3, pp. 895–901, 1998.

[3] F. Hartung, B. Girod, "Digital Watermarking of Raw and Compressed Video", *SPIE Procs Series*, vol. 2952, pp. 205–213, Oct. 1996.

[4] F. Hartung, B. Girod, "Digital Watermarking of MPEG-2 Coded Video in the Bitstream Domain", *Proc. of Int. Conf. on Acoustics, Speech, and Signal Processing*, vol. 4, pp. 2621–2624, Munich, April 1997.

[5] Web site of MOSES EC IST project: http://www.crl.co.uk/projects/moses

[6] A. Piva, F. Bartolini, M. Barni, "Managing Copyrights in Open Networks", *IEEE Internet Computing*, vol. 6, pp. 18–26, May-June 2002.

[7] W. Zeng, B. Liu, "A Statistical watermark detection technique without using original image for resolving rightful ownerships of digital images", *IEEE Trans. on Image Processing*, vol. 8, n. 11, pp. 1534–1548, 1999.