



INTERNATIONAL JOURNAL OF PURE AND APPLIED RESEARCH IN ENGINEERING AND TECHNOLOGY

A PATH FOR HORIZING YOUR INNOVATIVE WORK

DIGITAL WATERMARKING DESIGNED FOR ASTUTENESS MULTIMEDIA

Dr. B. V. SUBBA RAO¹, Mr. T. RAM KUMAR², Dr. J. RAJENDRA PRASAD³

1. Professor, Department of IT, P.V.P Siddhartha Institute of Technology, Vijayawada-, INDIA.
2. Associate Professor, Department of CSE, Vignan's Institute of Technology and Aeronautical Engineering, Deshmukhi, Nalgonda (Dt), A.P.,INDIA,
3. Professor, Department of IT, P.V.P Siddhartha Institute of Technology, Vijayawada, INDIA.

Abstract

Accepted Date:

16/10/2012

Publish Date:

01/11/2012

Keywords

Watermarking,
Multimedia,
Cryptography,
Intelligence

Corresponding Author

Dr. B. V. SUBBA RAO¹

Professor, Department of
IT, P.V.P Siddhartha
Institute of Technology.
Vijayawada-, INDIA.

Digital watermarking is a promising technology to embed information as unperceivable signals in digital contents. Various watermarking techniques have been proposed to protect copyrights of multimedia digital contents over Internet trading so that ownership of the contents can be determined in subsequent copyrights disputes. In this paper, we propose a watermark-based document distribution protocol, which complements conventional cryptography-based access control schemes; to address the problem of tracing unauthorized distribution of sensitive intelligence documents. The reinforcement of document distribution policies requires a concrete support of non-repudiation in the distribution process. The distribution protocol is adapted from our previous work on the watermarking infrastructure for enterprise document management. It makes use of intelligence user certificates to embed the identity of the users into the intelligence documents to who are distributed. In particular, keeping the identity secrecy between document providers and users (but yet traceable upon disputes) is a key contribution of this protocol in order to support for intelligence applications. We also outline an implementation of the distribution protocol and watermarking scheme employed.

INTRODUCTION

The enforcement of distribution policies for sensitive intelligence documents is important but difficult. Sensitive documents may be found left behind in conference rooms, common areas, printing rooms, or public folders. Access control based on cryptography alone cannot address this problem. Once after obtaining access to a sensitive document may a person make unnecessary copies or handle it without care. A major challenge in the reinforcement of distribution policies for sensitive documents is the support of *non-repudiation* in the underlying process so that unauthorized copies of intelligence documents can be identified and traced back to their users. The reinforcement should also be applicable to both hard copies and soft copies of the documents. Conventional cryptographic schemes that cover only soft copies are inadequate to handle this requirement.

Digital watermarking is a promising technology employed by various digital rights management (DRM) systems to achieve rights management. It supports copyright information (such as the owner's

identity, transaction dates, and serial numbers) to be embedded as unperceivable signals into digital contents. The signals embedded can be perceivable or unperceivable to humans. In this paper, we focus on the application of invisible watermarking techniques for documents that are based on the imperfection of the human vision system. While visible watermarks should be perceptible enough to discourage theft but not perceptible enough to decrease the utility or appreciation of the document, invisible watermarks should be imperceptible. Furthermore, robust watermarking techniques have been designed to resist tampering and support later extraction and detection of these watermark signals. These signals recover the rights information originally embedded in the document.

In this paper, we apply digital watermarking techniques for the distribution intelligence multimedia documents such as images and audios. In particular, we present a novel distribution protocol for such documents. The protocol is adapted from two pieces of previous work (Cheung & Chiu; Memon & Wong, which describes an enterprise

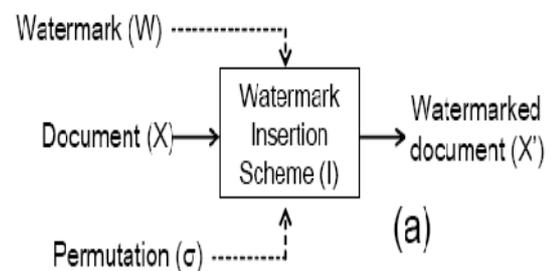
document management system and a watermarking protocol for purchasing digital contents over the Internet, respectively. It introduces the concepts of *intelligence user certificates* and trusted authorities responsible to issue these certificates. Document users may use the *intelligence user certificates* obtained from a trusted authority to identify themselves in acquisitions of intelligence documents. The same *intelligence user certificate* may be used in multiple acquisitions. These watermarks, once inserted, are difficult to be removed from their watermarked documents without knowing the exact *insertion* parameters. Watermarks can be preserved across media. For instance, a watermark embedded in a text document in its digital form can be detected in the hard copies of the digital document. If multiple watermarks are applied to individual digital copies, watermarking may also be used to indicate the identity of the legitimate document user of each copy. This allows unauthorized copies to be traced back to the document user from which they originated and thereby deterring unauthorized distribution of sensitive

documents. As such, the document distribution protocol should be able to distinguish the copies made by the document users from those made by *intelligence* document providers. Further, the use of intelligence user certificates together with intermediaries in our protocol enforces the identity secrecy between document suppliers and users (but yet traceable upon disputes). This is a key contribution in order to support for intelligence applications.

BACKGROUND AND RELATED WORK

In this section, we present the basic principles of watermarking schemes and the advantages of our watermarking protocols, by comparing related work.

Principle of Watermarking Schemes



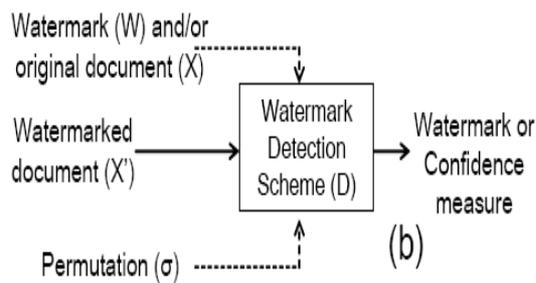


Figure 1 The Processes for (a) Watermark Insertion and (b) Watermark Detection

Watermarking schemes refer to the use of signal processing techniques to process watermarking signals in a piece of digital document. Existing watermarking schemes generally involve two stages: watermark insertion and watermark detection as shown in Figure 1. Suppose we have a digital document X , a watermark W , and a permutation function σ . A watermark insertion scheme I inserts a watermark W to the document X , where:

$$X' = I(X, W, \sigma)$$

For illustration, let us explain the principle of the insertion scheme based on a popular secure spread-spectrum watermarking technique proposed by Cox *et al.* [7]. The spread-spectrum technique assumes (i) the document is a vector of “features”, i.e., $X = \{x_1, x_2, \dots, x_n\}$ and (ii) the watermark signal

is a vector of “watermark elements”, i.e., $W = \{w_1, w_2, \dots, w_m\}$ with $n \geq m$. Note that the number of features in a document must be much greater than the number of components in a watermark signal so that the signal is unperceivable in the watermarked document X' . The permutation function σ is a bijection that shuffles the watermark elements before inserting them to the document X . As such, the shuffled watermark is a vector of $\sigma(W) = \{w_1', w_2', \dots, w_m'\}$, where $w_i' = \sigma(w_j)$ with $i, j \leq m$.

Corresponding to the watermark insertion scheme I , there is a watermark detection scheme D , which returns a confidence measure of the existence of a watermark W exists in a piece of document X' . A watermarking technique is referred to as *non-blind watermarking* when its detection scheme D requires the knowledge of the original document X , i.e., $D(X, X', W, \sigma)$ *false if W does not exist in X'* $D(X, X', W, \sigma)$ *true if W exists in X'*

Intelligence Document Distribution Infrastructure

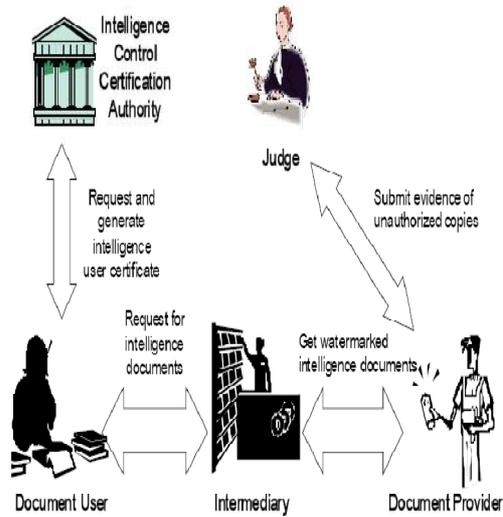


Figure 2 Overview of an intelligence document distribution infrastructure for intelligence document protection

We identify five distinguished roles in an intelligence document distribution infrastructure, namely, *Document user*, *document provider*, *Intelligence control certification authority*, *Judge*, and *Intermediary*, as shown in Figure 2. Document users are the ones who want to acquire some intelligence documents. Each copy of the intelligence document can be individually watermarked to identify its authorized user. Document providers are producers of these intelligence documents. Document providers employ their own, possibly proprietary, watermarking

techniques to encode watermarks into intelligence documents before distributing them to document users. *Intelligence control certification authorities* are trusted parties that generate *intelligence user certificate* identifying document users. *Judges* are trusted parties to resolve allegations filed by document providers against document users. Based on the evidence submitted by the document provider, a judge will decide whether the allegation is justified. *Intermediaries* are third party agents between document users and providers; they know both the document providers and the users. Intermediaries are only necessary in the cases where users and providers must remain anonymous to each other. Otherwise, if the user and provider know each other, the intermediaries can be bypassed. Intermediaries do not produce intelligence documents themselves and they do not need to be trusted.

In this paper, we are primarily interested in those intelligence documents with content values that can be preserved only by duplicating the source documents. Examples of these documents are often

multimedia in nature such as films, maps, photographs, and so on.

Although various application of watermarking schemes for the trading of digital contents have been studied in the literature (such as [4], [10], [12], [14], [15], [17]), a comprehensive treatment of these five roles altogether in the proposed intelligence document distribution infrastructure has not been studied, in particular, regarding the issuance of *intelligence user certificates* and their utilization in the control of intelligence documents.

Watermarking Protocols

In general, watermarking protocols govern the process of exchanging watermarks and watermarked digital contents between a user and a provider (traditionally a buyer and a merchant in contents trading over the Internet). Such watermarking protocols have been mainly deployed for complementing digital rights management. With the advancement of watermarking research and increasing adoptions, various problems of attacks to watermarking protocols are being discovered. A

watermarking protocol generally comprises three major processes:

Watermark generation, watermark insertion & distribution, and dispute resolution. The watermark generation process concerns the creation of a legitimate watermark that can identify a buyer. The watermark insertion process concerns the insertion of watermarks to digital contents by a merchant and the distribution of watermarked contents reliably to buyers (some work separates watermark insertion and content distribution into two processes). The dispute resolution concerns the resolution of copyrights upon the detection of suspected copies. In connection to these three major processes, latest researches on watermarking protocol generally address the six issues as tabulated in Table

1. Regarding the issues (b) to (f), different work makes different assumptions on the degree of trusts.

Table 1 Research Issues of Watermarking Protocols

Process	Issues
Watermark Generation	(a) Protection of watermark secrecy
Watermark Insertion & Distribution	(b) Buyers cannot be trusted
	(c) Merchants cannot be trusted
Dispute Resolution	(d) Buyers cannot be trusted
	(e) Merchants cannot be trusted
	(f) Judges cannot be trusted

The technical research issues in the protection of watermark secrecy in the process of watermark generation are similar to those occur in the public key infrastructure. As such, most existing works on watermarking protocols do not explicitly address that issue. Memon & Wong [17] and Cheung & Curreem [4] address the issue by requiring the buyers to present a valid public key on requesting a trusted *certification authority* for a legitimate watermark. Issue (b) is addressed by most existing watermarking protocols in the way that buyers are not trusted to provide a legitimate watermark. To resolve this issue, most protocols require intermediaries to be responsible for the watermark generation, while our protocol does not require this. Several studies attempt the problem that content merchants may not be trusted in the process of watermark insertion, i.e., the issue (c) in Table 1. Qiao and Nahrstedt [19]

suggest two ways to tackle the problem. One is to introduce a *trusted third party* (TTP). The merchant first sends the original content to the TTP, the content is encrypted with a symmetric key system. Then the watermark is generated at the TTP and inserted to the original content. Finally, the watermarked content is delivered to the buyer through a secured channel between the TTP and the buyer. Another alternative is to use cryptographic protocols between merchants and buyers.

A DISTRIBUTION PROTOCOL FOR INTELLIGENCE DOCUMENTS

Our distribution protocol consists of three processes: (i) generation of watermarks and intelligence user certificates, (ii) acquisition of watermarked intelligence documents, and (iii) resolution of policy violation. The processes and the data relations involved will be diagrammatically specified in the Unified Modeling Language (UML) [16], which is a well defined modeling language widely used for specifying, constructing, and documenting software systems. To support flexible enterprise document management policies, our distribution

protocol is designed to address the following two issues.

- **Maintenance of watermark secrecy:** The secrecy of document users' watermarks must be maintained because these watermarks identify document users. This issue is particularly important in the processes of document distribution where a party can be at the same time a document provider and a document user. Watermarks must not be released to document providers. In our protocol, a document user does *not* need to release his/her watermark to any parties after acquisition of the legitimate watermark.

- **Prevention of Trojan horse attacks:** A document user cannot use the intelligence user certificate of another user to obtain a watermarked document.

The document distribution protocol comprises three major processes: intelligence user certificate generation, intelligence document acquisition, policy violation resolution. The intelligence user certificate generation process concerns the creation of a registration certification, which embeds an encrypted version of a

legitimate watermark that identifies a document user. The watermarked document creation process governs the creation of watermarked documents and their reliable distributions to document users. The policy violation resolution process focuses on the collection of evidence and justification of a policy violation allegation against a document user.

Generation of an Intelligence User Certificate

Figure 3 and Figure 4 present the process of acquiring an intelligence user certificate and the associated data relations, respectively. Before applying for an intelligence user certificate, a document user should have obtained a Valid Public Key Infrastructure (PKI) Certificate, which contains a public key to be used in the purchase of digital contents. A legitimate certificate must be issued by a trusted PKI Certification Authority.

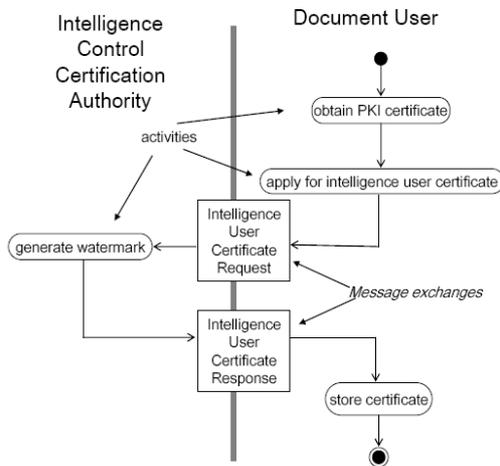


Figure 3 An UML Activity Diagram for the Acquisition of an Intelligence user certificate

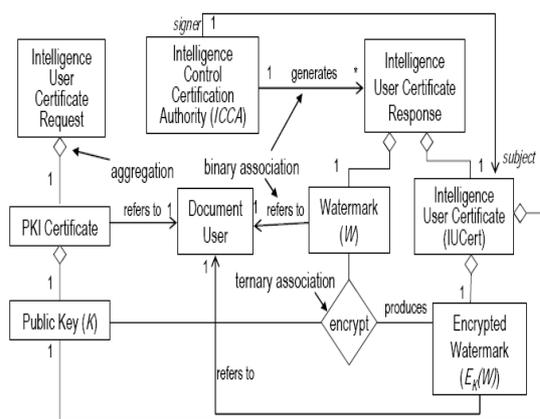


Figure 4 An UML Class Diagram of Data Relations for the Acquisition of Intelligence user certificates

When the document user wants to acquire a watermark for accessing a document, he/she attaches his/her PKI certificate in an intelligence user certificate request and submits it to a trusted intelligence control

certification authority. In response to the intelligence user certificate request, the intelligence control certification authority generates a legitimate watermark (W) and prepares an intelligence user certificate containing an encrypted copy ($EK(W)$) of W based on the document user's public key K . By $EK(W)$, we mean:

$$EK(W) = EK(\{w_1, w_2, \dots, w_m\}) = \{EK(w_1), EK(w_2), \dots, EK(w_m)\}.$$

The document user can verify the encrypted watermark, if necessary, using his/her private key and the received watermark. The watermark (W) uniquely identifies the document user. Like PKI private key, the watermark (W) is to be kept confidentially. Only the encrypted copy ($EK(W)$) is used in the subsequent acquisition of intelligence documents in order to protect the secrecy of the user's watermark. In addition, this allows the document provider to verify the consistency between $EK(W)$ and K .

Acquisition of Intelligence Documents

Figure 5 and Figure 6 present the process of acquiring an intelligence document and the associated data relations, respectively. In this process, a document user places a

request containing his/her intelligence user certificate (*IUCert*) to an intermediary that knows where to find a provider of the requested document. The intermediary then forwards the *IUCert* to the corresponding document provider. The provider retrieves the encrypted watermark ($E_K(W)$) and the user's public key (K) from the *IUCert* and verifies their consistency based on the digital signature $Sign_{ICCA}(IUCert)$ by the intelligence control certification authority. If the verification succeeds, the document provider generates a unique identifier (V) and prepares a hashed value $H(\sigma)$ of a selected permutation function σ using an one way hash function, such as MD5 (RSA **¡Error! No se encuentra el origen de la referencia.**). The permutation is to increase the watermark robustness so that the watermarked intelligence documents can better resist tampering.

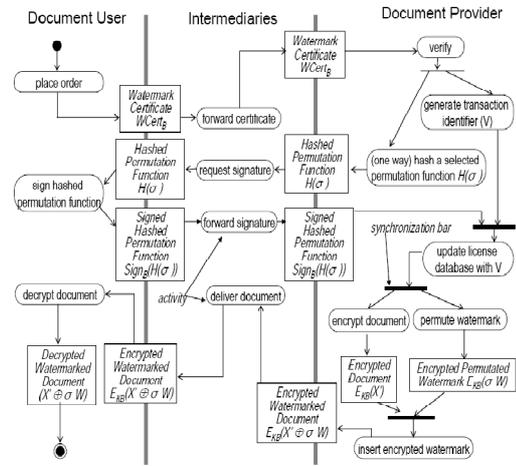


Figure 5 An UML Activity Diagram for the Acquisition of an Intelligence Document

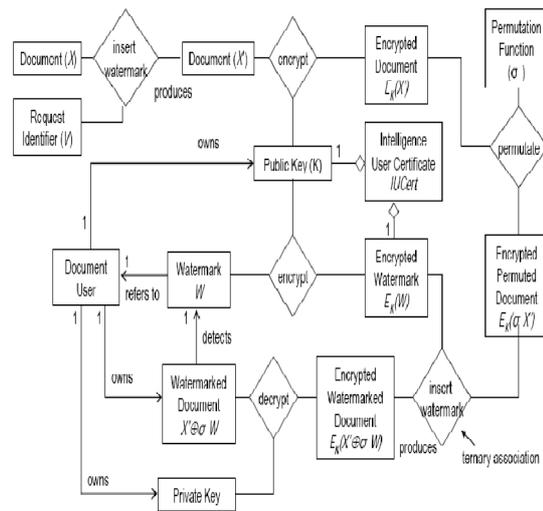


Figure 6 An UML Class Diagram of the Data Relations for the Acquisition of an Intelligence Watermarked Document

The hashed value is then signed with the private key of the user to produce

$Sign(H(\sigma))$. The private key used must match the public key (K) in the intelligence user certificate. It can be readily checked by using the user's public key after receiving $Sign(H(\sigma))$. This procedure allows the user to acknowledge the permutation function to be used in the subsequent watermark insertion process. After receiving the signed hashed value $Sign(H(\sigma))$, the document provider validates the signature using K . If the validation succeeds, the request details and the signed hashed value are recorded to a database; otherwise the request is aborted. To facilitate the detection of access right violation, the document X' is watermarked with the unique identifier V . The document X' is then encrypted to $E_K(X')$ using the public key K . The provider also permutes the encrypted watermark $E_K(W)$ with the function σ , resulting in $\sigma(E_K(W))$. Since $E_K(W)$ is a vector in the form of $\{E_K(w1), E_K(w2), \dots, E_K(wm)\}$, the resultant value gives the encrypted permuted watermark $E_K(\sigma W)$. For example, the well known RSA public key cryptosystem is one of those that exhibit

privacy homomorphism with respect to an addition operator.

The assumption for a watermark generation algorithm that supports an insertion function $X'=I(X,W,\sigma)$, a detection function $D(X',X,W,\sigma)$, and a privacy homomorphism are commonly supported by most watermark generation algorithms. The specific implementation of I and D does not affect the applicability of our protocol and therefore is not the focus of this paper. As such, the protocol can be used with most existing watermark generation algorithm.

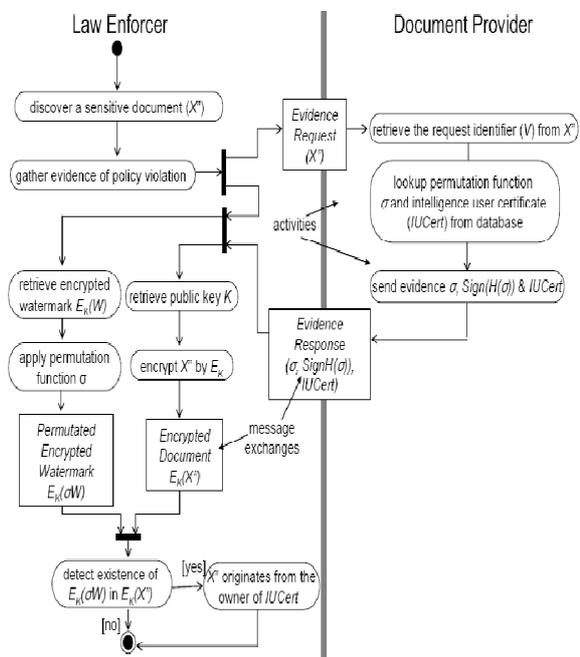


Figure 7 presents the protocol for the policy violation resolution process. When an unauthorized copy of document, say X'' , is found, the affected document provider can extract the unique request identifier V encoded in X'' .

IMPLEMENTATION FRAMEWORK

In this section, we present a case study based on our implementation framework to demonstrate the functionality and Practicability.

System Architecture

Figure 8 outlines the system architecture of an intelligence document distribution infrastructure centered on intermediaries. An intermediary needs to have a full-scale intelligence document management system, while document users or providers may rely on that of an intermediary. The main components of the management system are as follows.

1. The *front end application* is tightly coupled with an *access control layer* for authentication and control of document users and providers.
2. The *agencies / role manager* maintains the information of the document users and

providers in strict confidentiality. The roles captured the capability and authorization about which kind of documents they can use or provide.

3. The *document tracker* keeps track of all the document request and provision, validating the authorization.
4. The *watermark engine* processes watermark insertion and extraction. In case a document provider cannot process watermark, it may fall back to use the facility provided by the intermediary.
5. The *document repository and database* collectively serve as a backend to store all the above information, documents, and logs for non-repudiation purposes.

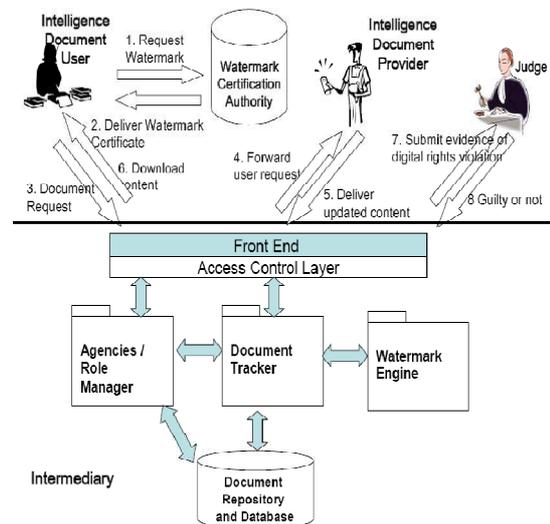


Figure 8 System Architecture Centered on Intermediaries

Furthermore, this architecture supports multiple tiers of intermediaries in an intelligence distribution network, which is an additional advantage. Instead of directly obtaining a document from a document provider, an intermediary can indirectly contact another capable intermediary for intelligence documents.

Structure of Intelligence User Certificate

Figure 9 depicts the structure of an intelligence user certificate. To support future extensions, each certificate carries a *version* number indicating its format. The *intelligence user certificate serial number*, assigned by the *intelligence control certification authority*, uniquely identifies each certificate. The *signature algorithm* identifier denotes the algorithm (say, md5RSA) used by the authority to sign this certificate. Fields are also contained in a certificate to indicate its issuer, owner, and effective period. The issuer of a certificate must be a trusted intelligence control certification authority. The *role* field specifies the role to be played by the owner of this certificate. Examples of role are individual, organization, and group. The role is used by document providers to define

various policies. Each certificate carries the public key of its owner, with which the certification authority encrypt the watermarks embedded in a certificate. Note that a document provider will permute the encrypted watermark before inserting it to a piece of digital content. The *watermarking scheme identifier* specifies the scheme to which the watermark is applicable. The *encrypted watermark* contains an encrypted value of each component of the watermark that the certification authority has issued to the owner of this certificate.



Figure 9 Format of an *intelligence user certificate* owned by a document user

Correlation between watermarks

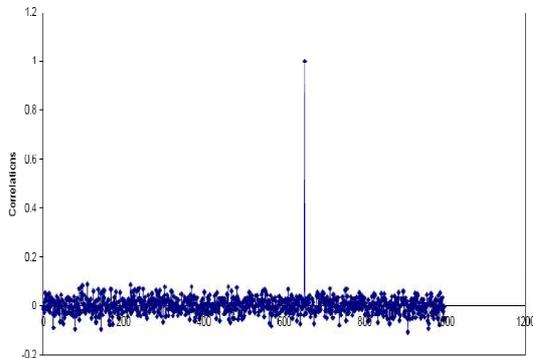


Figure 10 Correlation between watermarks

Next, we demonstrate our application of the RSA cryptosystem [51] for encryption in our protocol. In the encryption process, for a datum x and a public key a , the encrypted datum y is computed as:

$$y = E_a(x) = x^a \bmod n$$

where n is a product of two very large primes p and q . As for the decryption process, we use the private key b and compute:

$$x = D_b(y) = y^b \bmod n$$

Permutation function

During the watermark insertion phrase, the provider has to permute the document user's encrypted watermark. We implemented this function (σ) by randomly swapping the 1000 watermark coefficient.

The following code snippet illustrates how this can be done in the C programming language:

```
void permutefunc(VLONG wmark[], int size,
int seed) {int i, index1, index2; srand(seed);
for (i=0; i < rand() % 100 + 50) // min. 50
times, max 150 times
{VLONG tmp; index1=rand()%size; index2=ran
d()%size; // swap the two watermark
coefficient tmp
=wmark[index1]; wmark[index1]=wmark[ind
ex2]; wmark[index2]=tmp;}}
```

The VLONG structure can hold an integer of any size and seed represents the seed number used for random number generator. The provider will compile this code into an object code. Together with the random seed used, the provider will applied SHA-1 [22] to the object code to generate a message digest $H(\sigma)$. This message digest is put on the provider's site and must be downloaded by the document user beforehand. Then, the document user can sign this message digest $Sign(H(\sigma))$.

DISCUSSIONS

Watermark robustness is a key topic studied in the discipline of signal processing. Robustness refers to the ability to detect the watermark from a watermarked copy after common signal processing operations that do not destruct the contents. Various robust watermarking schemes (see section 2.3) have recently been proposed to survive different kinds of attacks, such as the insertion of malicious watermarks, spatial filtering, band-pass filtering, lossy compression, printing and scanning, re-sampling and noise addition, etc.

CONCLUSION

In this paper, a novel document distribution protocol has been proposed to address a problem in an intelligence distribution network so that document management policies can be properly reinforced. The protocol provides a concrete support for non-repudiation in the document distribution processes. It allows the document user, who has made each document copy, to be uniquely identified and accountable, and thus the route of document leakages can be identified. The

support of non-repudiation in fact reduced to the requirement of the absence of mutual trusts between document users and document providers. To realize the protocol, we have also outlined a possible implementation centered on intermediaries, which can isolate document users and providers.

References

1. H. Berghel, Watermarking Cyberspace, Communications of the ACM, 1997; 40 (11): 19-24
2. J. Bustos and K. Watson, Beginning .Net Web Services using C#, Birmingham, UK: Wrox Press Ltd., 2002.
3. S. C. Cheung and D. K. W. Chiu, A watermarking infrastructure for enterprise document management, in.
4. Proceedings of the 36th Hawaii International Conference on System Sciences (HICSS-36), CDROM, IEEE.
5. Press, Big Island, Hawaii, 2003, 10 pages.