# An Efficient and Anonymous Buyer-Seller Watermarking Protocol

Chin-Laung Lei, Pei-Ling Yu, Pan-Lung Tsai, and Ming-Hwa Chan

*Abstract*—For the purpose of deterring unauthorized duplication and distribution of multimedia contents, a seller may insert a unique digital watermark into each copy of the multimedia contents to be sold. When an illegal replica is found in the market sometime later, the seller can determine the responsible distributor by examining the watermark embedded. However, the accusation against the charged distributor, who was the buyer in some earlier transaction, is objectionable because the seller also has access to the watermarked copies and, hence, is able to release such a replica on her own. In this paper, a watermarking protocol is proposed to avoid such difficulties, known as the *customer's right problem*, in the phase of arbitration. The proposed watermarking protocol also provides a fix to Memon and Wong's scheme by solving the *unbinding problem*. In addition, the buyer is no longer required to contact the watermark certification authority during transactions, and the anonymity of the buyer can be retained through a trusted third party. The result is an efficient and anonymous buyer-seller watermarking protocol.

*Index Terms*—Copyright protection, digital watermark, privacy, security.

## I. INTRODUCTION

THE ADOPTION of digital media has been growing rapidly over the past decade. The digitized form of numerous multimedia contents not only facilitates various operations like editing, transformation, and duplication, but also allows fast distribution from one physical location to another via the Internet. Unfortunately, as any advancement in technology can always be used in good and bad ways, this one is no exception. Thanks to the proliferation of software tools for manipulating digital data, unauthorized duplication and distribution of multimedia contents has never been easier. Such effortless infringement of intellectual property and copyright, in turn, creates urgent demands for effective protection mechanisms. Among others, digital watermarking techniques are considered one of the most promising solutions.

Since the very first introduction of its concept, digital watermarking has been employed to develop a wide variety of schemes [1]–[5] aiming at protecting the intellectual property

C.-L. Lei, P.-L. Yu, and P.-L. Tsai are with the Department of Electrical Engineering, National Taiwan University, Taipei 106, Taiwan, R.O.C. (e-mail: lei@cc.ee.ntu.edu.tw; bey@fractal.ee.ntu.edu.tw; charles@fractal.ee.ntu.edu.tw).

M.-H. Chan was with the Department of Electrical Engineering, National Taiwan University, Taipei 106, Taiwan, R.O.C. He is now with the Department of Atmospheric Sciences, National Taiwan University, Taipei 106, Taiwan, R.O.C. (e-mail: jorchan@fractal.ee.ntu.edu.tw).

Digital Object Identifier 10.1109/TIP.2004.837553

and copyright of still images, sound clips, video streams, and software programs. A digital watermark can be characterized as additional signals added to certain digital contents and can be extracted sometime later to assert the lawful ownership of the original digital contents. In general, watermarks are classified into two categories, visible watermarks and invisible ones. Visible watermarks typically consist of text messages (e.g., copyright declarations) or company logos and can be easily perceived. In contrast, the only way to find out whether a digital content contains an invisible watermark is to examine the digital content with proper watermark detection and extraction algorithms.

In addition, watermarking schemes can be divided into blind, semiblind, and private watermarking schemes [6]. For blind watermarking schemes, the watermark detection and extraction algorithms do not require the presence of the original digital contents during the process of examination, and everyone is able to run the detection and extraction algorithms on each copy of the digital contents whenever necessary. On the other hand, the detection and extraction algorithms of nonblind watermarking schemes require the original digital contents as an input and, therefore, can only be performed by the legitimate owners of the digital contents. Nonblind watermarking schemes are typically considered more robust than blind ones. For example, when used to protect the copyright of still images, nonblind watermarking schemes usually exhibit better tolerance against various image-processing techniques, while watermarks embedded in the images using blind watermarking schemes can be distorted easily.

Because of the inherent robustness, we decide to focus on the application of invisible and private watermarking schemes for identifying the responsible distributor from whom an illegal replica of certain digital content is originated. This is accomplished by requiring the seller of the digital content to insert a unique watermark into each copy she sells. As a result, when a pirated copy is found, the embedded watermark will serve as both a traceable fingerprint and a hard evidence for the involvement of the particular buyer in some earlier transaction.

A number of watermarking protocols have been proposed to track down the distributors of illegal replicas [7], [8]. However, most of them ignore the fairness to the customers at all, and the others address the issue ineffectively, considering the current practice of law enforcement. Another common shortcoming of these protocols is the lack of appropriate mechanisms to protect customer privacy during transactions. Privacy is an important human right in real life, and is even more important in the electronic world. Shopping anonymously has almost become one of the requirements in the basic design of every protocol regarding
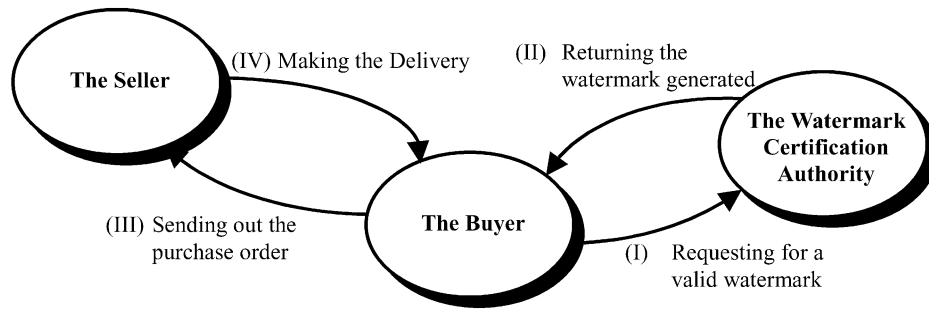
Fig. 1.   Interactions among the buyer, the seller, and the watermark certification authority in Memon and Wong's protocol.

electronic commerce. Anonymity helps customers to get rid of unsolicited advertisements and prevents possible leakage of critical personal information.

In this paper, a buyer-seller watermarking protocol that is more secure, more flexible, and more convenient for the general public than previous solutions is proposed. The proposed protocol provides a fix to Memon and Wong's scheme [7] by solving both the *customer's right problem* and the *unbinding problem*, which are explained later in Section II. The proposed protocol also enables innocent buyers to keep anonymous during transactions, while unambiguously identifying the responsible distributors (the guilty buyers) of illegal replicas. As a result, the proposed buyer-seller watermarking protocol is both effective and practical in terms of intellectual property and copyright protection.

The rest of this paper is organized as follows. In Section II, some important watermarking protocols are reviewed, and Section III describes the proposed watermarking protocol in detail. Section IV examines how the proposed protocol achieves the design goals and discusses the issues related to practicality. Section V then summarizes our achievements, including the improvements over other watermarking protocols.

## II. RELATED WORK

The intuitive idea of watermark-based fingerprinting has been implemented by a number of schemes using cryptographic techniques before the *customer's right problem* was first identified in [8]. Memon and Wong proposed a buyer-seller watermarking protocol in [7] to deal with the *customer's right problem*, but also introduced a new issue, the *unbinding problem*, in their solution. In this section, these two problems are described at length along with related protocols.

### A. Customer's Right Problem and Owner-Customer Watermarking Protocol

All of the traditional watermark-based fingerprinting protocols were designed to protect the seller's ownership of digital contents rather than the buyer's right. In these protocols, the seller is entitled to the responsibility of generating and inserting digital watermarks. As a result, the seller is granted access to each watermarked copy, exactly what the buyer finally gets when the transaction completes. Doing so is equivalent to assuming that the seller is always trustworthy, which is not true. Owing to this implicit assumption, a malicious seller can easily frame the buyer involved in a particular transaction

by releasing the corresponding watermarked copy afterwards. Even with a decent seller, it is still possible for a pirated copy that is actually leaked out because of a security breach in the selling system to mistakenly point to an innocent buyer in the phase of arbitration.

Without appropriate mechanisms to unambiguously determine the actual distributor of a pirated copy, these protocols suffer from the *customer's right problem*, as stated above. In an attempt to tackle the *customer's right problem*, Qiao and Nahrstedt proposed an owner-customer watermarking protocol in [8]. In this protocol, the buyer (customer) first encrypts a predetermined sequence of bits with a secret key only known to him, and sends the encrypted data to the seller (owner). The seller then generates a unique watermark based on the encrypted data received and inserts it into a copy of the digital content, and sends the watermarked copy back to the buyer. Since only the buyer knows the secret key, he can prove to anyone his legitimate possession of the watermarked copy. However, when a pirated copy is found, the charge against the buyer pointed by the embedded watermark is objectionable because the seller in Qiao and Nahrstedt's protocol still has access to the watermarked copy in its final form. Consequently, the buyer's repudiation to the distribution of the pirated copy cannot be overruled in the trial.

### B. Buyer-Seller Watermarking Protocol

In [7], Memon and Wong proposed a buyer-seller watermarking protocol to deal with the *customer's right problem* by preventing the seller from direct access to the final watermarked copy. In their protocol, still images are taken as the example digital contents, and three roles, the seller, the buyer, and the watermark certification authority (denoted by **WCA**), are involved as shown in Fig. 1. To complete a transaction, the following steps are carried out in sequence.

1) The buyer sends his identity and public key $B$ to **WCA** and requests a valid watermark. Upon receiving the request, **WCA** randomly generates a watermark $W$ with its vector representation being $\{w_1, w_2, w_3, \ldots, w_n\}$, and encrypts it as $E_B(W) = \{E_B(w_1), E_B(w_2), E_B(w_3), \ldots, E_B(w_n)\}$, where $E_K$ denotes the *privacy-homomorphic* encryption function with respect to the watermark insertion operation of the underlying watermarking scheme using public key $K$. Then, **WCA** signs $E_B(W)$ and sends it back to the buyer.

2) The buyer forward $E_B(W)$ and **WCA**'s signature to the seller. After verifying that $E_B(W)$ is perfectly signed and, hence, $W$ is, indeed, a valid watermark generated by **WCA**, the seller chooses a permutation function $\sigma$ to shuffle the elements in $E_B(W)$ to get $E_B(W') = \sigma(E_B(W)) = E_B(\sigma(W)) = \{E_B(w'_1), E_B(w'_2), E_B(w'_3), \ldots, E_B(w'_n)\}$, and inserts the permuted watermark into $X$, the image to be sold, in the encrypted domain by computing $E_B(X') = E_B(X \oplus \sigma(W)) = E_B(X) \oplus E_B(\sigma(W)) = E_B(X) \oplus E_B(W')$, where $\oplus$ is the watermark insertion operation of the underlying *linear* watermarking scheme, and $X'$, or $X \oplus \sigma(W)$, is the watermarked copy of $X$ in its final form (not known to the seller). After the seller finishes the computation, she sends $E_B(X')$ to the buyer.

3) When the buyer receives $E_B(X')$, he decrypts it with his private key $B^{-1}$ by computing $X' = D_{B^{-1}}(E_B(X'))$, where $D_{K^{-1}}$ denotes the decryption function using private key $K^{-1}$, and gets the correctly watermarked copy $X'$.

In step 1), $E_K$ is required to be *privacy homomorphic* with respect to $\oplus$. $E_K$ is said to be *privacy homomorphic* with respect to $\oplus$ if $E_K(a \oplus b) = E_K(a) \oplus E_K(b)$ holds for every $a$ and $b$ in the message space. In step 2), the underlying watermarking scheme is required to be *linear*. A watermarking scheme is said to be *linear* if its watermark insertion operation can be represented as the computation of $Y' = Y \oplus Z = \{y_1 \oplus z_1, y_2 \oplus z_2, y_3 \oplus z_3, \ldots, y_q \oplus z_q, y_{q+1}, y_{p+2}, y_{q+3}, \ldots, y_p\}$, where $Y = \{y_1, y_2, y_3, \ldots, y_p\}$ is the vector representation of the original digital content, $Z = \{z_1, z_2, z_3, \ldots, z_q\}$ is the vector representation of the watermark, $p \geq q$, and $Y'$ is the watermarked copy of $Y$.

When a pirated copy is found sometime later, and the watermark embedded in it leads to certain buyer who denies the unauthorized distribution, the protocol goes into the phase of dispute resolution. At this stage, the seller needs to send $E_B(W)$, **WCA**'s signature associated with $E_B(W)$, and $\sigma$ to the arbiter, and the arbiter will ask the buyer to decrypt $E_B(W)$ with his private key $B^{-1}$ to get $W$. After knowing $W$, the arbiter executes the watermark detection and extraction algorithm to check whether $\sigma(W)$ is indeed present in the pirated copy. If $\sigma(W)$ is found, the arbiter will adjudicate the buyer to be guilty; otherwise, the buyer is innocent. In Memon and Wong's protocol, if the buyer refuses to or cannot decrypt $E_B(W)$, he will also be considered guilty as charged.

Memon and Wong's protocol successfully solves the *customer's right problem* since the watermark insertion operation is performed in the encrypted domain and thus the seller has no access to the watermarked copy of the digital content in its final form. To prevent the seller from accessing the watermarked copy afterwards does not imply that the buyer must make the copy untouchable by locking it in a safety box, because showing a picture to a friend or playing a song in the living room is not the same as giving away the watermarked copy in its digital form. If the seller wants to verify the buyer's legitimate possession of the watermarked copy, she can always ask the arbiter to do so, or performs the verification via zero-knowledge watermarking schemes as in [9], [10]. To sum up, the solution to the *customer's*

*right problem* by preventing the seller from accessing the watermarked copy to be sold neither puts extra burdens on the buyer nor overly restricts the use of the legally purchased watermarked copy. On the other hand, the seller's ownership is well protected because the buyer has no idea about the permutation function and, therefore, is unable to remove the embedded watermark.

In spite of solving the *customer's right problem*, Memon and Wong's protocol has several issues. First of all, the seller is able to intentionally transplant a watermark initially embedded in a copy of certain digital content into another copy of a completely different digital content, provided both copies are sold to the same buyer. We refer to this as the *unbinding problem*, which will be explained further in Section II-C. Another deficiency is concerning the arbitration. In the phase of dispute resolution, Memon and Wong's protocol requires the suspected buyer to decrypt the previously encrypted watermark. Whether the buyer is innocent or guilty, in actuality, he is unlikely to cooperate. Furthermore, it is not reasonable for a protocol to depend on such cooperation because doing so will then allow a malicious seller to easily harass an innocent buyer by repeatedly requiring cooperation. Besides, the disposition of presuming the guilt of an uncooperative buyer is also questionable because in the general practice of law, it is the responsibility of the accuser to proof the guilt of the defendant, not in reverse. In addition, the protocol restricts itself to the use of *linear* watermarking schemes and, hence, provides limited flexibility in practice. Finally, the buyer in Memon and Wong's protocol needs to communicate with both **WCA** and the seller to complete a single transaction. Contacting more than one party is inconvenient, and sometimes more expensive or even prohibitive, especially when implementing electronic commerce in the paradigm of mobile computing.

### C. Unbinding Problem

The *unbinding problem* arises because most of the previously proposed watermarking protocols, including Memon and Wong's, fail to provide proper mechanisms on binding a chosen watermark to a specific digital content or a specific transaction. Therefore, once the seller discovers a pirated copy, it is possible for her to transplant the watermark embedded in the pirated copy into another copy of a higher-priced digital content to produce made-up piracy so that she can get compensated more.

The details of producing such made-up piracy are explained as follows. Assume that the buyer has bought $U'$ and $V'$ from the seller, where $U' = U \oplus \sigma_1(W_1)$ and $V' = V \oplus \sigma_2(W_2)$ are the watermarked images of two different still images $U$ and $V$, respectively, and $U$ is much more expensive than $V$, $\sigma_1$ and $\sigma_2$ are two different permutation functions, and $W_1$ and $W_2$ are different watermarks. Suppose sometime later the seller gets a pirated copy of $V'$, which the buyer distributes without authorization. Since the seller has the corresponding original image $V$, it is trivial for her to derive $\sigma_2(W_2)$ from $V'$. Then, she is able to insert $\sigma_2(W_2)$ into $U$ and encrypts the resulting watermarked image with the buyer's public key $B$ to obtain $E_B(U \oplus \sigma_2(W_2))$. With $E_B(W_2)$, **WCA**'s signature associated with $E_B(W_2)$, and $\sigma_2$ in hand, now the seller has all the evidences needed and is ready to claim that the buyer illegally

distributes $U'' = U \oplus \sigma_2(W_2)$, which is actually never sold to the buyer.

Here is another possible exploit of the unbinding vulnerability. Following Memon and Wong's protocol, the seller keeps records of $E_B(W)$, **WCA**'s signature associated with $E_B(W)$, and $\sigma$ after a successful transaction. The next time the same buyer comes to purchase another image, the ill-behaved seller simply ignores the newly generated and encrypted watermark forwarded by the buyer and inserts the same permuted watermark $\sigma(W)$ into the copy of the requested image in the encrypted domain. Under such circumstances, the same $E_B(W)$, **WCA**'s signature associated with $E_B(W)$, and $\sigma$ can be used to testify against the particular buyer regardless which still image actually being pirated, provided that the buyer has committed piracy. As a result, when a pirated copy is found, the seller can always claim that the buyer also illegally distributes copies of other images purchased by him.

## III. PROPOSED SCHEME

Derived from Memon and Wong's buyer-seller watermarking protocol [7], the proposed protocol addresses many issues found in its predecessor. In addition to dealing with the *customer's right problem*, the proposed protocol provides a fix to the *unbinding problem* mentioned above, and also employs Public-Key Infrastructure (PKI) to attain several important achievements that are missing in the watermarking protocols proposed previously. In this section, we first define the roles and notations to be used throughout the rest of this paper and explain the goals as well as the assumptions. Then, we continue to elaborate the three subprotocols that comprise the proposed protocol: the *registration protocol*, the *watermarking protocol*, and the *identification and arbitration protocol*.

In the model of the proposed protocol, five different roles involved are as follows.

1) **S**: The seller, who wants to make a profit on the sales of certain digital content. The seller may be the rightful owner of the original digital content, or an authorized reselling agent.
2) **B**: The buyer, who wants to purchase a copy of the digital content from **S**.
3) **CA**: A trusted certification authority, which is responsible for issuing anonymous certificates. Before a transaction is carried out, **B** applies to **CA** for an anonymous certificate. In this case, **B**'s public key serves as his pseudonym, and the anonymous certificate is used to certify that the pseudonym is correctly registered to **CA**, and **CA** actually knows about the real identity behind the pseudonym [11]. The existence of **CA** provides the possibility of **B**'s anonymity as well as the assurance to **S**.
4) **WCA**: A trusted watermark certification authority, which is responsible for the generation of random and valid watermarks. The existence of WCA asserts the validity of watermarks and ensures that the watermarks generated are not revealed to **S**.
5) **ARB**: An arbiter, who adjudicates lawsuits against the infringement of copyright and intellectual property.

The notations are defined as follows.

| | |
|---|---|
| $(pk_{\mathbf{I}}, sk_{\mathbf{I}})$ | A public-private key pair, where **I** is the identity of its owner. That is, $pk_{\mathbf{I}}$ is **I**'s public key, while $sk_{\mathbf{I}}$ is **I**'s private key. |
| $\mathrm{Sign}_{\mathbf{I}}(M)$ | The signature of message $M$ signed by **I** with his private key. |
| $E_{pk_{\mathbf{I}}}(M)$ | The ciphertext of message $M$ encrypted with **I**'s public key. The encryption can be performed by anyone. |
| $D_{sk_{\mathbf{I}}}(C)$ | The original message of ciphertext $C$ decrypted by **I** with $sk_{\mathbf{I}}$. |
| $\mathrm{Cert}_{\mathbf{J}}(\mathbf{I})$ | The digital certificate issued to subject **I** by certification authority **J**. Throughout the rest of this paper, we purposely use X.509-compliant digital certificates [12] so that anyone is able to verify the validity of any certificate, and the public key associated with a particular subject can be easily obtained from his certificate. |
| $X \oplus W$ | The watermarked copy of digital content $X$. The binary operator $\oplus$ denotes the operation of watermark insertion, and $W$ is the watermark to be inserted. |

The goals to be achieved by the proposed protocol are as follows.

1) The proposed watermarking protocol is secure and fair to both **B** and **S**. In particular, the proposed protocol must guarantee that it is impossible for **S** to frame **B** by any means, and that there is no chance for **B** to successfully remove the watermark embedded in any copy he purchased. Besides, the proposed protocol is also supposed to solve the *customer's right problem* and the *unbinding problem* mentioned previously. To the best of our knowledge, no existing watermarking protocol in the literature satisfies this requirement.
2) The operations of watermark insertion are not performed by **WCA**. Watermark insertion is time consuming and may require numerous computational resources, especially for large-volume digital contents, such as high-resolution still images and video streams. We do not want to overload **WCA**, and we do not want **WCA** to hold all of the original digital contents, either. Qiao and Nahrstedt were also aware of this concern in [8].
3) **B** is not required to contact anyone but **S** in each transaction. As discussed before, contacting more than one party in each transaction is inconvenient, more expensive, or even prohibitive to **B**, and it also significantly limits the applicability of the protocol, as in the case of Memon and Wong's one.
4) **B**'s privacy is well protected. The identity of **B** is not supposed to be exposed unless he is proven to have committed piracy. Surprisingly, the watermarking protocols proposed previously seldom consider the issue of privacy protection of **B**.

An important assumption of the proposed watermarking protocol is that PKI has been well deployed so that each party has its own public-private key pair as well as a digital certificate issued by **CA**. Once the assumption holds, everyone is able to authenticate anyone else and the message exchange between any two
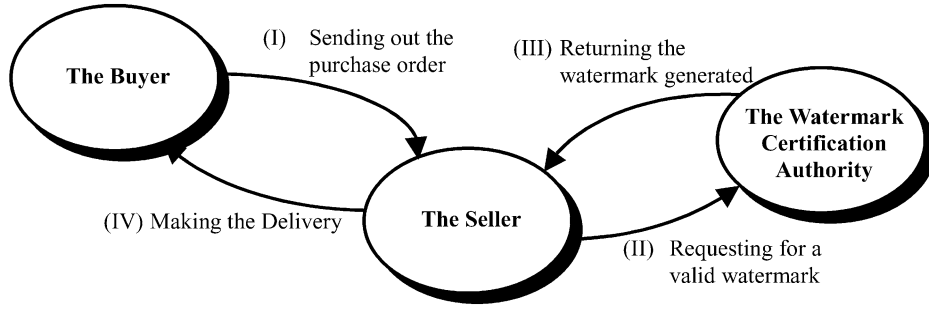
Fig. 2.   Interactions among the buyer, the seller, and the watermark certification authority in the proposed watermarking protocol.
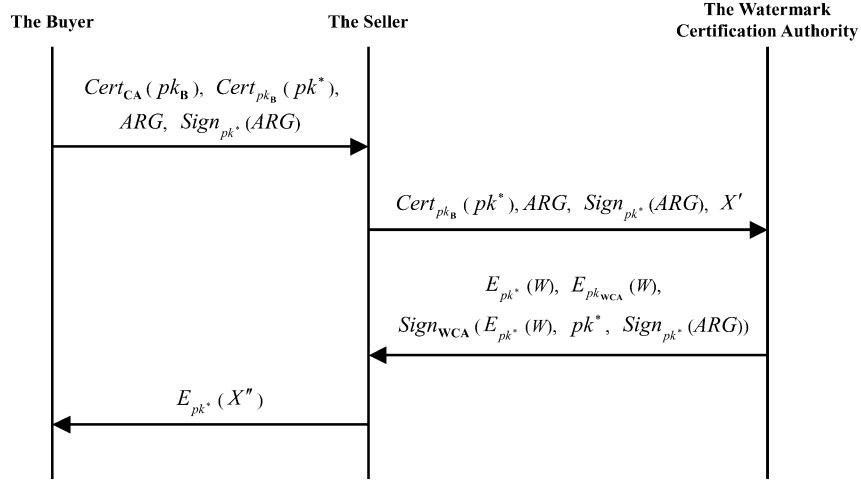


Fig. 3.   Details of a transaction in the proposed watermarking protocol.

parties can be made secure with no problem. We also assume that the encryption function used in PKI is *privacy homomorphic* with respect to the watermark insertion operation. Recall that $E_{pk_I}$ is said to be *privacy homomorphic* with respect to $\oplus$ if $E_{pk_I}(a \oplus b) = E_{pk_I}(a) \oplus E_{pk_I}(b)$ holds for every $a$ and $b$ in the message space. For example, the well-known RSA cryptosystem [13] is a *privacy homomorphism* with respect to multiplication [14]. Another public-key encryption function that is *privacy homomorphic* with respect to addition can be found in [15].

One of the major differences between the proposed watermarking protocol and its predecessor is that the underlying watermarking scheme is not required to be *linear* in the proposed protocol. That is, we do not require the computation of the watermark insertion to be performed on an element-by-element basis. As long as *privacy homomorphism* is preserved, any kind of watermarking schemes, including those that do not tolerate the permutation of watermarks, can be adopted.

### A. Registration Protocol

If **B** wants to stay anonymous during transactions, he may apply to **CA** for an anonymous certificate in advance, via the *registration protocol* presented here. An anonymous certificate is a normal digital certificate except that the content of its subject field is a pseudonym rather than the real identity of the applicant. By issuing the anonymous certificate to **B**, **CA** is responsible for binding this particular anonymous certificate to **B**. **CA** also

guarantees that the binding is not revealed to any other party unless requested by **ARB** when **B** is proven to have committed piracy.

To apply for an anonymous certificate, **B** first randomly selects a key pair $(pk_B, sk_B)$ and sends $pk_B$ to **CA**. When **CA** receives $pk_B$, it generates an anonymous certificate, $\text{Cert}_{CA}(pk_B)$, and sends it back to **B**. In the proposed protocol, we let $pk_B$ be the pseudonym associated with the anonymous certificate issued to **B**, as mentioned in the standard of Simple Public-Key Infrastructure (SPKI) [16], [17], which allows for the use of pseudonymous public keys. A single anonymous certificate can be used in multiple transactions, and it is up to **B** to decide when to apply for another anonymous certificate. Alternatively, **B** may skip the entire registration process and use his (normal) digital certificate if anonymity is not a concern.

### B. Watermarking Protocol

To carry out a transaction, **B**, **S** and **WCA** follow the *watermarking protocol* described in this subsection. Fig. 2 shows the interaction among **B**, **S**, and **WCA**, and Fig. 3 visualizes the details of the following steps.

1) To acquire a copy of digital content $X$, **B** first negotiates with **S** to set up a common agreement, ARG, which explicitly states the rights and obligations of both parties, and specifies the digital content of interest. ARG uniquely binds this particular transaction to $X$ and can be regarded as a purchase order. Note that **B** may use his pseudonym

in the negotiation to keep his identity unexposed. Alternatively, since ARGs are different only in the part specifying the digital content of interest, it is possible for **S** to generate ARGs beforehand and put them in a public place (e.g., her Web site), along with the catalog of the digital contents to be sold, so that **B** may have anonymous access to ARGs.

2) After the initial negotiation, **B** randomly selects a one-time key pair $(pk^*, sk^*)$ and generates an anonymous certificate, $\mathrm{Cert}_{pk_\mathbf{B}}(pk^*)$, with $pk^*$ being the associated pseudonym, on the honor of $pk_\mathbf{B}$, **B**'s pseudonym associated with $\mathrm{Cert}_\mathbf{CA}(pk_\mathbf{B})$. In this case, **CA** assures the legality of $pk_\mathbf{B}$, and $pk_\mathbf{B}$, in turn, assures the legality of $pk^*$. Then, $B$ transmits $\mathrm{Cert}_\mathbf{CA}(pk_\mathbf{B})$, $\mathrm{Cert}_{pk_\mathbf{B}}(pk^*)$, ARG, and $\mathrm{Sign}_{pk^*}(\mathrm{ARG})$ to **S**.

3) Upon receiving $\mathrm{Cert}_\mathbf{CA}(pk_\mathbf{B})$, $\mathrm{Cert}_{pk_\mathbf{B}}(pk^*)$, ARG, and $\mathrm{Sign}_{pk^*}(\mathrm{ARG})$, **S** verifies the validity of the certificates and the signature, and aborts the transaction if any of them is invalid. Otherwise, she generates a unique watermark $V$ for this particular transaction and computes $X' = X \oplus V$, where $X'$ is the watermarked digital content. In this step, **S** may employ any watermarking scheme she likes, provided that it is able to resist various possible distortions resulted from malicious attacks and still can be successfully extracted sometime later. The existence of $V$ in $X'$ is to produce a key of search for **S** to quickly locate a specific entry in her sales records when a pirated copy is found. It is not meant for proving the involvement of a cheating buyer. Then, **S** sends $\mathrm{Cert}_{pk_\mathbf{B}}(pk^*)$, ARG, $\mathrm{Sign}_{pk^*}(\mathrm{ARG})$, and $X'$ to **WCA**.

4) When **WCA** receives $\mathrm{Cert}_{pk_\mathbf{B}}(pk^*)$, ARG, $\mathrm{Sign}_{pk^*}(\mathrm{ARG})$, and $X'$ from **S**, it verifies the validity of the certificate and the signature, and aborts the transaction if any of them is invalid. Otherwise, it generates a watermark $W$ specific to this transaction. Since $X'$ is also transmitted to **WCA**, it is possible for **WCA** to create a more robust watermark according to the characteristics of $X'$ so that the well-tailored watermark can be more difficult to disrupt and cause less perceivable defects when later inserted into $X'$. If **S** has concerns about sending $X'$ to **WCA**, she may offer a profile describing $X'$ instead. After $W$ is successfully generated, **WCA** computes $E_{pk^*}(W)$, $E_{pk_\mathbf{WCA}}(W)$, and $\mathrm{Sign}_\mathbf{WCA}\big(E_{pk^*}(W), pk^*, \mathrm{Sign}_{pk^*}(\mathrm{ARG})\big)$, and sends them back to **S**.

5) Upon receiving the response, **S** performs the second-round watermark insertion in the encrypted domain by computing $E_{pk^*}(X'') = E_{pk^*}(X' \oplus W) = E_{pk^*}(X') \oplus E_{pk^*}(W)$, without knowing the actual watermark, $W$. **S** also has no idea about the doubly watermarked copy of $X$ in its final form, $X''$, or $X' \oplus W$. Note that the computation of $E_{pk^*}(X'')$ is possible because $E_{pk^*}$ is *privacy homomorphic* with respect to $\oplus$. Afterwards, **S** delivers $E_{pk^*}(X'')$ to **B** and stores $V$, $\mathrm{Cert}_\mathbf{CA}(pk_\mathbf{B})$, $\mathrm{Cert}_{pk_\mathbf{B}}(pk^*)$, ARG, $\mathrm{Sign}_{pk^*}(\mathrm{ARG})$, $E_{pk^*}(W)$, $E_{pk_\mathbf{WCA}}(W)$, and $\mathrm{Sign}_\mathbf{WCA}\big(E_{pk^*}(W), pk^*, \mathrm{Sign}_{pk^*}(\mathrm{ARG})\big)$ in a new entry of $\mathrm{Table}_X$, her sales records with respect to digital content $X$.

6) After receiving $E_{pk^*}(X'')$, **B** decrypts it with $sk^*$ by computing $X'' = D_{sk^*}(E_{pk^*}(X''))$ and obtains the correctly watermarked copy $X''$. Note that **S** cannot intentionally alter the value of $E_{pk^*}(X'')$ via any kind of post processing after it is computed. Otherwise **B** may fail to restore $X''$ in the process of decryption and decide to abort the transaction.

Following the *watermarking protocol*, **B** eventually gets the watermarked copy of the digital content he wants to purchase. **S** cannot derive the same watermarked copy because she lacks for the knowledge of $X''$ and $W$. On the other hand, it is also impossible for **B** to remove the watermark since he knows nothing about $V$ and $W$.

## C. Identification and Arbitration Protocol

When a pirated copy $Y$ of certain digital content $X$ owned by **S** is found in the market, the *identification and arbitration protocol* depicted in this subsection can be conducted to determine the identity of the responsible distributor, who was the buyer in some earlier transaction, with undeniable evidences.

The first thing for **S** to do is to run the corresponding watermark detection and extraction algorithm on $Y$ to extract the watermark inserted in step 3) described in the previous subsection. Let $V'$ denote the watermark extracted. **S** then uses $V'$ as a key to search $\mathrm{Table}_X$, her sales records with respect to $X$, for a match, as shown in Fig. 4. The exact mechanism for finding a match completely depends on the watermarking scheme adopted. Generally, the searching is accomplished via correlating $V'$ with each $V_i$ in $\mathrm{Table}_X$ and selecting the one with the highest value of correlation over the threshold of a predetermined confidence level, where $V_i$ is the watermark stored in entry $i$ of $\mathrm{Table}_X$. When a match is found, **S** collects the associated information, $X'$ (computed by $X \oplus V_k$, where $V_k$ is the key of the matched entry), $\mathrm{Cert}_\mathbf{CA}(pk_\mathbf{B})$, $\mathrm{Cert}_{pk_\mathbf{B}}(pk^*)$, ARG, $\mathrm{Sign}_{pk^*}(\mathrm{ARG})$, $E_{pk^*}(W)$, $E_{pk_\mathbf{WCA}}(W)$, and $\mathrm{Sign}_\mathbf{WCA}(E_{pk^*}(W), pk^*, \mathrm{Sign}_{pk^*}(\mathrm{ARG}))$, stored in the matched entry and sends them along with $Y$ to **ARB** for arbitration.

Note that the watermarks embedded in $Y$ may have been distorted by signal processing techniques in the attempts made to remove or disrupt the watermarks. Although **S** may still be able to extract the first watermark by exploiting a robust watermarking scheme, it is not possible for her to detect or extract the second watermark without knowing what it is. It is also not feasible for **S** to detect or extract the second watermark in the encrypted domain, i.e., to alternatively detect or extract $E_{pk^*}(W)$ from $E_{pk^*}(Y)$, because doing so will require **S** to first encrypt $Y$ with every one-time public key in $Table_X$ before running the corresponding watermark detection and extraction algorithm.

Upon receiving $X'$, $\mathrm{Cert}_\mathbf{CA}(pk_\mathbf{B})$, $\mathrm{Cert}_{pk_\mathbf{B}}(pk^*)$, ARG, $\mathrm{Sign}_{pk^*}(\mathrm{ARG})$, $E_{pk^*}(W)$, $E_{pk_\mathbf{WCA}}(W)$, $\mathrm{Sign}_\mathbf{WCA}\big(E_{pk^*}(W), pk^*, \mathrm{Sign}_{pk^*}(\mathrm{ARG})\big)$, and $Y$ from **S**, **ARB** verifies the validity of the certificates and the signatures, and rejects the case if any of them is invalid. Otherwise, **ARB** requests **WCA** to decrypt $E_{pk_\mathbf{WCA}}(W)$ and obtains $W$. Then **ARB** performs another verification on the correctness of $E_{pk^*}(W)$ sent by **S** via encrypting $W$ with public key $pk^*$. If $E_{pk^*}(W)$ turns out to be incorrect, **ARB**
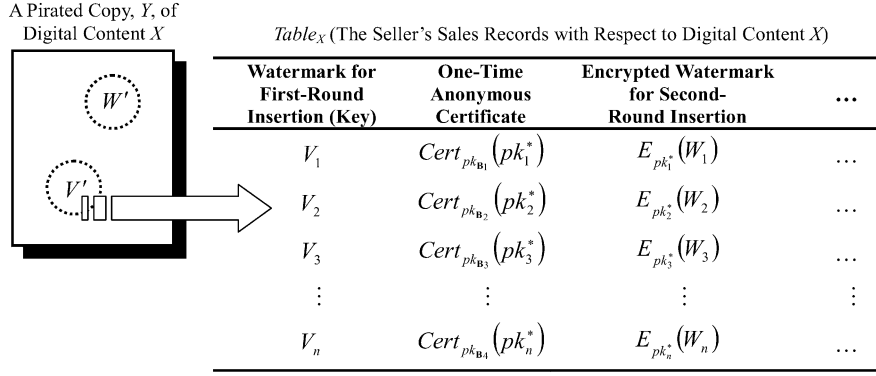
A Pirated Copy, $Y$, of Digital Content $X$

$Table_X$ (The Seller's Sales Records with Respect to Digital Content $X$)

| Watermark for First-Round Insertion (Key) | One-Time Anonymous Certificate | Encrypted Watermark for Second-Round Insertion | ... |
|---|---|---|---|
| $V_1$ | $Cert_{pk_{\mathbf{B}_1}}(pk_1^*)$ | $E_{pk_1^*}(W_1)$ | ... |
| $V_2$ | $Cert_{pk_{\mathbf{B}_2}}(pk_2^*)$ | $E_{pk_2^*}(W_2)$ | ... |
| $V_3$ | $Cert_{pk_{\mathbf{B}_3}}(pk_3^*)$ | $E_{pk_3^*}(W_3)$ | ... |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $V_n$ | $Cert_{pk_{\mathbf{B}_4}}(pk_n^*)$ | $E_{pk_n^*}(W_n)$ | ... |

Fig. 4. Seller searches her sales records for a matched watermark.

also rejects the case. If the case is undertaken, $\mathbf{ARB}$ continues to run the corresponding watermark detection and extraction algorithm (with $X'$, $W$, and $Y$ as inputs) to determine the existence of $W$ in $Y$. If $W$ is indeed found in $Y$, $\mathbf{ARB}$ turns to $\mathbf{CA}$ and asks for the real identity behind pseudonym $pk^*$. Once the identity of the buyer who owns $pk^*$ is revealed, $\mathbf{ARB}$ adjudicates the buyer to be guilty and closes the case. If $W$ is not detected in $Y$, the buyer is considered innocent, and his identity remains unexposed.

## IV. DISCUSSIONS

In this section, we examine how the design goals are achieved and discuss some issues related to practicality, including the deployment of trusted third parties, the storage requirement of $\mathbf{S}$, and system availability. It has to be first pointed out here that the security of the proposed watermarking protocol essentially depends on the security and robustness of the underlying watermarking scheme as well as the security of the cryptosystem used to construct PKI.

### A. Accomplishment of Design Goals

Recall that we have set up four design goals in the beginning of Section III. All of them are achieved by the proposed watermarking protocol.

1) The proposed watermarking protocol is secure and fair to both $\mathbf{B}$ and $\mathbf{S}$. From the perspective of $\mathbf{S}$, the proposed watermarking protocol is secure and fair because $\mathbf{B}$ has no idea about the original digital content and the watermark embedded in the copy he purchased, and, hence, is unable to remove the watermark. The proposed watermarking protocol also provides mechanisms to unambiguously identify the guilty buyer once a pirated copy is found.

   From $\mathbf{B}$'s viewpoint, the proposed watermarking protocol is also secure and fair. Since $\mathbf{S}$ gets no access to the watermarked copy of the digital content in its final form, $\mathbf{S}$ cannot distribute illegal replicas and then lays her sins at $\mathbf{B}$'s threshold. That is, the *customer's right problem* is avoided in the first place. As for the *unbinding problem*, the signature $\mathrm{Sign}_{\mathbf{WCA}}(E_{pk^*}(W), pk^*, \mathrm{Sign}_{pk^*}(\mathrm{ARG}))$ explicitly binds $W$ to ARG, which, in turn, uniquely specifies a particular digital content $X$, so it is impossible for

$\mathbf{S}$ to play the trick of watermark transplantation. In addition, because the proposed watermarking protocol introduces the use of one-time key pairs, it is also impossible for $\mathbf{S}$ to fool $\mathbf{B}$ using outdated information taken from a previous transaction. If $\mathbf{S}$ tries to send $\mathbf{B}$ a watermarked copy encrypted with a wrong key, the probability for $\mathbf{B}$ to get any meaningful result from decryption is next to zero. In short, the *unbinding problem* is solved completely.

2) The operations of watermark insertion are not performed by $\mathbf{WCA}$. In the proposed watermarking protocol, $\mathbf{S}$ is the one in charge of inserting the watermarks (in the encrypted domain) and, therefore, $\mathbf{WCA}$ is relieved from excessive computations. There is also no need for $\mathbf{WCA}$ to hold any of the original digital contents.

3) $\mathbf{B}$ is not required to contact anyone but $\mathbf{S}$ in each transaction. As depicted in Fig. 2, $\mathbf{B}$ can safely delegate $\mathbf{S}$ to request $\mathbf{WCA}$ for a valid and robust watermark under PKI. From $\mathbf{B}$'s perspective, a single transaction consists of selecting the digital content to buy, sending out the purchase order, and waiting for the delivery. It very much resembles what people do when shopping in real life, and just cannot be any simpler.

4) $\mathbf{B}$'s privacy is well protected. The proposed watermarking protocol takes advantage of anonymous certificates to preserve the anonymity of $\mathbf{B}$ during transactions. As a matter of fact, the anonymity achieved in the proposed watermarking protocol is in its weakest form because it is actually asserted by $\mathbf{CA}$, a trusted third party. Under the assumption of $\mathbf{CA}$'s existence, $\mathbf{B}$ can keep his real identity unexposed unless he is adjudicated to be guilty by $\mathbf{ARB}$.

   In spite of $\mathbf{B}$'s anonymity, the transactions carried out by the same pseudonym are linkable to one another, and there are still risks for $\mathbf{B}$'s private information to be inferred through data mining techniques. A solution to this issue is to limit the maximum number of transactions for a single anonymous certificate to involve. An alternative way to reduce the risks is to let $\mathbf{B}$ apply for a number of anonymous certificates simultaneously and randomly choose one for each transaction. The exact policy depends on the degree of privacy needed.

Another achievement of the proposed watermarking protocol is the elimination of the need to ask for $\mathbf{B}$'s cooperation in the phase of arbitration. In the trial, as long as $\mathbf{S}$ provides enough

evidences, trusted third parties are capable of making appropriate adjudications collaboratively.

### B. Deployment of Trusted Third Parties

The cost of deploying trusted third parties directly impacts the practicality of most security protocols. In reality, a memoryless trusted third party, which does not keep records of information associated with requests received, is considered less expensive and is much more practical to implement. In the proposed watermarking protocol, each and every trusted third party can be memoryless.

Upon receiving a request for an anonymous certificate, $\mathbf{CA}$ may encrypt the real identity of the supplicant with a nondeterministic encryption algorithm and deposit the result in the extension field of the generated anonymous certificate. When later asked by $\mathbf{ARB}$ to reveal the real identity behind a pseudonym associated with certain anonymous certificate, $\mathbf{CA}$ simply decrypts the data item stored in the extension field of the anonymous certificate and derives the real identity. There is no need for $\mathbf{CA}$ to memorize the associations between the real identities of the supplicants and the anonymous certificates issued.

$\mathbf{WCA}$ is also not required to maintain a database of all watermarks generated because once generated, the watermarks are encrypted with its public key and handed over to $\mathbf{S}$. When requested by $\mathbf{ARB}$ to disclose a specific watermark, $\mathbf{WCA}$ restores the watermark by decrypting the ciphertext provided by $\mathbf{ARB}$, which, in turn, receives the ciphertext from $\mathbf{S}$. As for $\mathbf{ARB}$, it simply reacts upon the requests from $\mathbf{S}$ and has nothing to remember.

### C. Storage Requirements of the Seller

In the proposed watermarking protocol, the burden of storing necessary information has been put on $\mathbf{S}$. It is reasonable because real-world vendors are very likely to have their own sales records, and, in most cases, they are even unwilling to grant others access to such information. The cost of keeping sales records can also be regarded as a part of an investment in the business. In addition, considering the ever-decreasing prices of various storage devices, the overhead introduced is actually negligible in running a business.

### D. System Availability

As illustrated in Fig. 2, $\mathbf{WCA}$ is an additional role involved in each transaction between $\mathbf{B}$ and $\mathbf{S}$. Consequently, the availability of $\mathbf{WCA}$ is critical to the availability of the whole system. In general, the availability of $\mathbf{WCA}$ requires the reliability of $\mathbf{WCA}$ itself and the stability of the underlying communication network. To improve the reliability of $\mathbf{WCA}$, cluster-based solutions can be used to construct a single node of $\mathbf{WCA}$. On the other hand, deploying a distributed system with multiple nodes of $\mathbf{WCA}$ over the Internet solves the problem of network failures that may cause temporary or permanent disconnection from a particular node.

Even if the availability of $\mathbf{WCA}$ is assured by the distributed-system techniques mentioned above, it may also be necessary for the party who must communicate with $\mathbf{WCA}$ to possess multiple physical connections to the Internet so that the party will have less chance to suffer from complete disconnection from the network. In the proposed watermarking protocol, only $\mathbf{S}$ is required to contact $\mathbf{WCA}$ during transactions. We argue that it is more practical for $\mathbf{S}$ (i.e., a corporate) to have multiple network connections rather than for $\mathbf{B}$ (i.e., an individual) to do so. On the contrary, in the watermarking protocols proposed previously, $\mathbf{B}$ is also obliged to contact $\mathbf{WCA}$ during transactions, and therefore those protocols are considered less practical.

## V. CONCLUSION

In this paper, we propose a buyer-seller watermarking protocol derived from Memon and Wong's protocol and solve both the *customer's right problem* and the *unbinding problem*. In the proposed watermarking protocol, the operations of watermark insertion are performed by the seller rather than by the watermark certification authority. In addition to identifying and addressing the issue of the *unbinding problem* in Memon and Wong's protocol, we also achieve a number of improvements over their scheme, including the following.

1) The buyer is required to contact only one party, the seller, during transactions. Specifically, the seller requests the watermark certification authority for the watermark on behalf of the buyer, while the buyer's right is still protected.
2) The buyer remains anonymous during transactions via the help of a trusted third party.
3) In the phase of arbitration, the cooperation from the buyer is not required.
4) The underlying watermarking scheme adopted does not need to be *linear*, since the seller is not going to further permute the watermarks generated by the watermark certification authority. In particular, watermarking schemes that do not tolerate permutation are not excluded from the choices.
5) The watermark certification authority is given at least a profile describing the digital content of interest, so it has a greater chance to generate a more robust watermark tailored for the particular digital content. It is impossible to do so in Memon and Wong's protocol because the watermark generated is going to be further permuted by the seller.

As mentioned previously, the security of the proposed watermarking protocol essentially depends on the security and robustness of the underlying watermarking scheme. Unfortunately, all of the known watermarking schemes in the literature have been broken. To develop a truly robust watermarking scheme remains an open research topic.

## REFERENCES

[1] G. Depovere, T. Kalker, and J.-P Linnartz, "Improved watermark detection reliability using filtering before correlation," in *Proc. IEEE Int. Conf. Image Processing*, vol. 1, Oct. 1998, pp. 430–434.
[2] J. J. Eggers and B. Girod, "Blind watermarking applied to image authentication," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, vol. 3, May 2001, pp. 1977–1980.
[3] J. R. Hernández, F. Pérez-González, J. M. Rodríguez, and G. Nieto, "Performance analysis of a 2-D-multipulse amplitude modulation scheme for data hiding and watermarking of still images," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 510–524, May 1998.

[4] N. Memon and P. W. Wong, "Protecting digital media content," *Commun. ACM*, vol. 41, no. 7, pp. 35–43, July 1998.

[5] G. Voyatzis, N. Nikolaidis, and I. Pitas, "Digital watermarking: An overview," in *Proc. 9th Eur. Signal Processing Conf.*, Sept. 1998, pp. 9–12.

[6] S. Katzenbeisser, "On the design of copyright protection protocols for multimedia distribution using symmetric and public-key watermarking," in *Proc. 12th Int. Workshop Database and Expert Systems Applicat.*, Sept. 2001, pp. 815–819.

[7] N. Memon and P. W. Wong, "A buyer-seller watermarking protocol," *IEEE Trans. Image Processing*, vol. 10, pp. 643–649, Apr. 2001.

[8] L. Qiao and K. Nahrstedt, "Watermarking schemes and protocols for protecting rightful ownership and customer's rights," *J. Vis. Commun. Image Representation*, vol. 9, pp. 194–210, Sept. 1998.

[9] S. Craver, "Zero knowledge watermark detection," in *Proc. 3rd Int. Workshop Information Hiding*, vol. 1768, LNCS, Sept. 1999, pp. 101–116.

[10] K. Gopalakrishnan, N. Memon, and P. L. Vora, "Protocols for watermark verification," *IEEE Multimedia*, vol. 8, pp. 66–70, Oct.–Dec. 2001.

[11] D. Chaum, "Security without identification: Transaction systems to make big brother obsolete," *Commun. ACM*, vol. 28, pp. 1030–1044, Oct. 1985.

[12] R. Housley, W. Polk, W. Ford, and D. Solo, "Internet X.509 public key infrastructure certificate and certificate revocation list (CRL) profile," in *RFC 3280*, Apr. 2002.

[13] R. L. Rivest, A. Shamir, and L. Adleman, "A method for obtaining digital signatures and public key cryptosystems," *Commun. ACM*, vol. 21, no. 2, pp. 120–126, Feb. 1978.

[14] D. R. Stinson, *Cryptography: Theory and Practice*, 1st ed. New York: Chapman & Hall, 1995.

[15] J. D. Cohen and M. J. Fischer, "A robust and verifiable cryptographically secure election scheme (extended abstract)," in *Proc. 26th Annu. Symp. Foundations of Computer Science*, Oct. 1985, pp. 372–382.

[16] C. Ellison, "SPKI requirements," in *RFC 2692*, Sept. 1999.

[17] C. Ellison, B. Frantz, B. Lampson, R. Rivest, B. Thomas, and T. Ylonen, "SPKI certificate theory," in *RFC 2693*, Sept. 1999.

**Pei-Ling Yu** received the B.S. degree in computer science from National Chiao-Tung University, Taiwan, R.O.C., in 1996. He is currently pursuing the Ph.D. degree at the Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan.

His research interests include computer and network security, cryptography, and electronic payment systems.



**Pan-Lung Tsai** received the B.S. degree in mechanical engineering and the M.S. degree in electrical engineering from the National Taiwan University, Taipei, Taiwan, R.O.C., in 1997 and 1999, respectively. He is currently pursuing the Ph.D. degree at the Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan.

His research interests include software agents, mobile computing, network security, and distributed processing.

Mr. Tsai is a member of the Association for Computing Machinery.



**Chin-Laung Lei** received the B.S. degree in electrical engineering from the National Taiwan University, Taipei, Taiwan, R.O.C, in 1980, and the Ph.D. degree in computer science from the University of Texas, Austin, in 1986.

From 1986 to 1988, he was an Assistant Professor in the Computer and Information Science Department, The Ohio State University, Columbus. In 1988, he joined the faculty of the Department of Electrical Engineering, National Taiwan University, where he is now a Professor. His current research interests include computer and network security, cryptography, parallel and distributed processing, design and analysis of algorithms, and operating system design.

Dr. Lei is a member of the Association for Computing Machinery.



**Ming-Hwa Chan** received the B.S. degree in atmospheric sciences and the M.S. degree in electrical engineering from the National Taiwan University, Taipei, Taiwan, R.O.C., in 1995 and 2002, respectively.

From 2002 to 2003, he was with AscenVision Technology, Inc., Taiwan, as a Senior Software Engineer. He is currently a Research Assistant in the Department of Atmospheric Sciences, National Taiwan University.