ON THE (IN)SECURITY OF TWO BUYER-SELLER WATERMARKING PROTOCOLS

Geong Sen Poh* and Keith M. Martin
Information Security Group, Royal Holloway, University of London, Egham, Surrey, TW20, 0EX, U.K.

Keywords: Buyer-Seller Watermarking, Security Protocols.

Abstract: A buyer-seller watermarking protocol deters dishonest buyers from illegally distributing bought content. This is achieved by giving the seller the capability to trace and identify these buyers, while also allowing the seller to prove illegal acts to a third party. At the same time, an honest buyer is prevented from being falsely accused of illegal content distribution by the seller. Many protocols have been proposed, with two recent proposals being the protocols proposed by Ibrahim *et al.* in IAS 2007 and SECRYPT 2007. We will show that these protocols are not secure, especially for the seller. We further put forward our thoughts on how it is possible to avoid the security weaknesses found in them.

1 INTRODUCTION

One of the many methods devised to address illegal distribution of digital content is copy deterrence through the use of fingerprinting schemes (Blakley et al., 1985; Wagner, 1983). A seller uses these schemes to embed a unique watermark into content before selling it to a buyer. If an illegal copy is found, the seller traces and identifies the buyer based on this watermark. However, since the seller knows the watermark, it is possible for the seller to frame an honest buyer by embedding the watermark into content and distributing copies of it. Conversely, a dishonest buyer can stage claims that illegal copies of content are actually distributed by the seller instead of the buyer. Hence asymmetric fingerprinting schemes (Camenisch, 2000; Pfitzmann and Schunter, 1996; Pfitzmann and Waidner, 1997) and buyer-seller watermarking protocols (Choi et al., 2003; Goi et al., 2004; Ju et al., 2002; Lei et al., 2004; Memon and Wong, 2001) were proposed to address these issues.

Many of the buyer-seller watermarking protocols focus on improving existing protocols by introducing new properties. However, additional features can come at the expense of existing safeguards. We show that this is precisely the case for two protocols proposed by Ibrahim *et al.* in IAS 2007 (Ibrahim et al., 2007a) and SECRYPT 2007 (Ibrahim et al., 2007b). While such an approach provides a straightforward design where new properties are added on top of the existing one, it may overlook potential security issues emerging from these additions. This is the case with Ibrahim *et al.*'s protocols, and our main objective is to examine the security issues in these two protocols.

We begin by identifying the basic security properties of a buyer-seller watermarking protocol in Section 2. We describe Ibrahim *et al.*'s protocols in Section 3. Using the notions and the protocols' description, we demonstrate the security weaknesses and how certain attacks can be mounted successfully on Ibrahim *et al.*'s protocols in Section 4. We then conclude our discussion in Section 5.

*This author is supported by the European Commission under contract IST-2002-507932 (ECRYPT).

2 BASIC NOTIONS

In this section we identify the main security properties required by a buyer-seller watermarking protocol.

2.1 Goals

The goals of a buyer-seller watermarking protocol are:

- A seller can *trace* a buyer from found copy of content.
- A seller should not be able to *frame* a buyer of illegal content distribution.
- A buyer who is guilty of distributing copies of

Sen Poh G. and M. Martin K. (2008).
ON THE (IN)SECURITY OF TWO BUYER-SELLER WATERMARKING PROTOCOLS.
In *Proceedings of the International Conference on Security and Cryptography*, pages 253-260
DOI: 10.5220/0001919702530260

content illegally, should not be able to *deny* this fact.

2.2 Security Properties

Before we can derive the main required security properties, we need to identify the threats. In brief, we expect an adversary to be able to observe and change all the information transmitted between all involved parties. This represents the threat commonly found in a security protocol (Boyd and Mathuria, 2003). Specifically for a buyer-seller watermarking protocol, we also need to guard against a non-trustworthy seller from framing an honest buyer. Similarly, we need to guard against a dishonest buyer from denying the act of illegal content distribution.

In common with many security protocols we require authenticated communication channels between the involved parties. There are many standard techniques for providing such channels (see (Boyd and Mathuria, 2003; Dent and Mitchell, 2004)). We will refer to this requirement as the need for *communication security*. The three main security properties that are required that are specific to buyer-seller watermarking are:

- *Traceability.* A legitimate, but dishonest buyer who illegally distributed purchased content can be traced to their identity by the seller.
- *Framing Resistance.* An honest buyer cannot be falsely accused of illegal distribution by the seller.
- *Non-repudiation of Redistribution.* A dishonest buyer found to have illegally redistributed purchased content cannot deny this fact by claiming that these copies were created and distributed by the seller. In other words, the seller obtains proof of illegal activity of the buyer.

We also note that many of the protocols (Choi et al., 2003; Goi et al., 2004; Lei et al., 2004; Pfitzmann and Waidner, 1997) include a property known as *anonymity and unlinkability*, which allows a buyer to obtain content without revealing the buyer's real identity and the buyer's past preference. Since Ibrahim *et al.*'s protocols that we will examine do not provide this property, we will not discuss it any further.

2.3 Players and Trust Assumptions

A buyer-seller watermarking protocol involves two or more players. The *buyer* buys content and the *seller* sells content to the buyer. The buyer and the seller do not trust each other. There may be a third party that is involved in processing the buyers' watermarks. Whenever this third party is introduced, it is mostly assumed to be fully trusted. It is extremely important to clarify from the outset whether the trusted party is fully trusted.

Remark. We note that this section has set out a simple foundation for a buyer-seller watermarking protocol. We will soon see in our analysis that getting these notions right can help prevent weaknesses in the design of such a protocol.

3 IBRAHIM *ET AL.*'S PROTOCOLS

In this section we briefly describe two protocols due to Ibrahim *et al.*, beginning with the notation and building blocks for the various objects and parties, and the proposed additional "properties". More details can be found in (Ibrahim et al., 2007a; Ibrahim et al., 2007b).

Notation. Table 1 shows the objects and parties involved in the protocols.

Table 1: Notation.

B	Buyer.
S	Seller.
R	Reseller (who was a buyer in the first-hand market).
CA	Trusted Certificate Authority that issues digital certificates to protocol participants.
A	An arbiter settles dispute of illegal distribution between the buyer and the seller.
X	An original content.
V	Seller's watermark to uniquely identify a content buyer.
W	Buyer's watermark.
X'	Intermediate content where V is embedded into X.
X''	Marked content where W is embedded into X'.
Y	Illegal copy of a marked content.
(ek_I, dk_I), (vk_I, sk_I)	Public-private key pairs of party I.
$Cert_{CA}(I)$	A digital certificate issued by CA to party I.
OL	A digital object license issued by the seller to the reseller, containing the number of resells allowed for the reseller.

Building Blocks. The required building blocks are as follows:

- $X \oplus W$ means W is embedded into X based on an embedding operator $\oplus$.
- $\mathtt{Sig}_{sk_I}(.)$ is a signature generation algorithm with signing key sk_I.
- $\mathtt{Ver}_{vk_I}(.)$ is a signature verification algorithm with verification key vk_I.

- $\mathtt{HEnc}_{ek_I}(.)$ is a homomorphic encryption algorithm (e.g. RSA (Rivest et al., 1978) and Paillier (Paillier, 1999)) with encryption key ek_I. For our discussion, we require that given $\mathtt{HEnc}_{ek_I}(M_1)$ and $\mathtt{HEnc}_{ek_I}(M_2)$, we have $\mathtt{EncH}_{ek_I}(M_1 \oplus M_2) = \mathtt{HEnc}_{ek_I}(M_1) \oplus \mathtt{EncH}_{ek_I}(M_2)$. Here M_1 and M_2 are content or watermarks.
- $\mathtt{HDec}_{dk_I}(.)$ is a homomorphic decryption algorithm with decryption key dk_I.
- $H(.)$ is a cryptographic hash function.

Additional "Properties". Ibrahim *et al.* claim their two protocols, in addition to fulfilling the main properties (which are stated as "problems" in their protocols) defined in Section 2.2, also address the following "problems":

- *Conspiracy problem*, which refers to the possibility for a seller to *conspire* with a watermark authority in order to reveal the buyer's watermark. By revealing this watermark, it is then possible for the seller to frame the buyer by embedding the watermark into content and distributing copies of it.
- *Unbinding problem*, in which given a found illegal marked content, the seller can extract the watermark, re-embed this watermark into a more valuable content and accuse the buyer of illegally distributing copies of the found content and the more valuable content. This is possible when the watermark is not *bound* to the content itself.
- *Buyer's participation in the dispute resolution problem.* This issue refers to the assumption that during dispute of illegal distribution, the seller is solely responsible for proving the guilt of the buyer to a third party, and the buyer should not be required to participate in such a process.
- *Man in the middle attack.* This issue refers to the ability of an adversary to insert and modify the messages in transmission, without either of the seller or the buyer knowing that the communication channel has been compromised.
- *Practice applicability problem.* This issue refers to the need for the buyer to contact not just the seller, but also another party to obtain the content, which is inconvenient for the buyer.

3.1 Protocol in IAS 2007

This protocol (Ibrahim et al., 2007a) is intended for the secure selling of digital content between the seller and the buyer. In the following we described the protocol, which consists of two sub-protocols.

Watermark Generation/Insertion Protocol. This is the main protocol for purchase and watermarking (Figure 1).

1. Buyer B initiates the protocol by sending a purchase request to the seller S.
2. Seller S sends his/her certificate $Cert_{CA}(S)$ to B.
3. A purchase agreement AGR is agreed between B and S. This agreement states the rights, obligations and specifies content X.
4. Buyer B generates hash value $H(AGR)$ and a signature $\mathtt{Sig}_{sk_B}(H(AGR))$. This signature allows an arbiter A to confirm B's purchase during a dispute.
5. Buyer B generates a watermark W and signs it as $\mathtt{Sig}_{sk_B}(W)$. This is further encrypted using CA's encryption key as $\mathtt{HEnc}_{ek_{CA}}(\mathtt{Sig}_{sk_B}(W))$. During a dispute, S will send this encrypted object to CA so that B need not participate in the dispute resolution protocol.
6. Buyer B encrypts W, resulting in $\mathtt{HEnc}_{ek_B}(W)$.
7. Buyer B generates $\mathtt{Sig}_{sk_B}(H(H(W),H(AGR)))$. The purpose of this signature is to bind W to AGR.
8. Buyer B sends the signature $\mathtt{Sig}_{sk_B}(H(AGR))$, the encrypted watermark $\mathtt{HEnc}_{ek_B}(W)$, the encrypted signature $\mathtt{HEnc}_{ek_{CA}}(\mathtt{Sig}_{sk_B}(W))$, the signature $\mathtt{Sig}_{sk_B}(H(H(W),H(AGR)))$ and the buyer certificate $Cert_{CA}(B)$ to S.
9. After the message is received from the buyer B, the seller S sends the encrypted signature $\mathtt{HEnc}_{ek_{CA}}(\mathtt{Sig}_{sk_B}(W))$ and the buyer certificate $Cert_{CA}(B)$ to CA. Next CA decrypts the encrypted signature $\mathtt{HEnc}_{ek_{CA}}(\mathtt{Sig}_{sk_B}(W))$ to obtain the signature $\mathtt{Sig}_{sk_B}(W)$, which is then verified to obtain the watermark W. After that, CA re-encrypts W with B's encryption key as $\mathtt{HEnc}'_{ek_B}(W)$ and signs it to obtain $\mathtt{Sig}_{sk_{CA}}(\mathtt{HEnc}'_{ek_B}(W))$. This signature is sent to S. This is to prevent B from encrypting watermark W in $\mathtt{HEnc}_{ek_B}(W)$ while including a different watermark W' in $\mathtt{HEnc}_{ek_{CA}}(\mathtt{Sig}_{sk_B}(W'))$.
10. When the message from CA is received, the seller S verifies the signature $\mathtt{Sig}_{sk_{CA}}(\mathtt{HEnc}'_{ek_B}(W))$. After that, S generates hash value $H(\mathtt{HEnc}_{ek_B}(W))$ and hash value $H(\mathtt{HEnc}'_{ek_B}(W))$. These two hash values are compared, and if they are identical then S continues to run the protocol. If they are not identical, the protocol is halted.
11. Seller S generates a unique watermark V and embeds it into content X. The computation is $X' = X \oplus V$.

12. Seller S generates a marked and encrypted content as follows: $\mathtt{HEnc}_{ek_B}(X'') = \mathtt{EncH}_{ek_B}(X' \oplus W) = \mathtt{HEnc}_{ek_B}(X') \oplus \mathtt{HEnc}_{ek_B}(W)$.
13. Seller S stores in the database the following: V, AGR, $\mathtt{Sig}_{sk_B}(H(AGR))$, $\mathtt{HEnc}_{ek_{CA}}(\mathtt{Sig}_{sk_B}(W))$, $\mathtt{Sig}_{sk_B}(H(H(W),H(AGR)))$ and $Cert_{CA}(B)$.
14. Seller S sends the signature $\mathtt{Sig}_{sk_S}(\mathtt{HEnc}_{ek_B}(X''))$ and S's certificate $Cert_{CA}(S)$ to B.
15. Buyer B verifies $\mathtt{Sig}_{sk_S}(\mathtt{HEnc}_{ek_B}(X''))$ to retrieve $\mathtt{HEnc}_{ek_B}(X'')$ and then decrypts it to obtain X''.

Dispute Resolution Protocol. Whenever an illegal copy Y is found, the seller S runs this protocol to prove that a dishonest buyer B distributed this illegal copy Y.

1. Seller S detects the watermark V from the illegal copy Y using a watermarking detection algorithm corresponding to the embedding process. If the watermark V is detected, the seller S sends B's certificate $Cert_{CA}(B)$, the watermark V, two signatures, $\mathtt{Sig}_{sk_B}(H(H(W),H(AGR)))$ and $\mathtt{Sig}_{sk_B}(H(AGR))$, the agreement AGR, the encrypted signature $\mathtt{HEnc}_{ek_{CA}}(\mathtt{Sig}_{sk_B}(W))$ along with the illegal copy Y and the marked content X' to the arbiter A.
2. When the message is received from the seller S, the arbiter A sends the encrypted signature $\mathtt{HEnc}_{ek_{CA}}(\mathtt{Sig}_{sk_B}(W))$ together with his certificate $Cert_{CA}(A)$ to CA. When these are received, CA decrypts the encrypted signature $\mathtt{HEnc}_{ek_{CA}}(\mathtt{Sig}_{sk_B}(W))$ and encrypts the retrieved signature, $\mathtt{Sig}_{sk_B}(W)$, with A's encryption key, as $\mathtt{HEnc}_{ek_A}(\mathtt{Sig}_{sk_B}(W))$. This encrypted object is sent back to the arbiter A. The arbiter A decrypts this encrypted signature $\mathtt{HEnc}_{ek_A}(\mathtt{Sig}_{sk_B}(W))$ to retrieve $\mathtt{Sig}_{sk_B}(W)$ and verify it to obtain the watermark W.
3. Arbiter A detects the watermark W from the illegal copy Y.
4. If the watermark W is detected, arbiter A verifies the signature $\mathtt{Sig}_{sk_B}(H(AGR))$ based on the agreement AGR. If the verification is successful, the protocol continues. Otherwise it halts.
5. As the final step, the arbiter A verifies the signature $\mathtt{Sig}_{sk_B}(H(H(W),H(AGR)))$ using the watermark W and the agreement AGR given by the seller S. If the verification is successful, which proves the buyer B bought the content, then the buyer B is found guilty.

3.2 Protocol in SECRYPT 2007

This protocol (Ibrahim et al., 2007b) is similar to the previous protocol except that it involves a legitimate reseller R, who acts as an agent for S and sells content bought from S to a buyer B. The watermark generation/insertion protocol is illustrated in Figure 2. The main differences between this protocol and the previous one are:

- the creation of an object license OL by the seller S to monitor the selling of content by the reseller R. Each time the reseller R wants to sell content, the seller S generates a new object license OL' counting down the number of resells allowed.
- Instead of sending messages to the seller S, the buyer B sends messages to the reseller R, who then contacts the seller S.

We do not provide further details of this protocol since our analysis works on both protocols in a similar way.

4 OUR ANALYSIS

We now analyse these protocols by defining attacks and demonstrate how these attacks can be mounted successfully. We conclude this section by reflecting on why the attacks are possible and how they may be prevented.

4.1 Definition

- *Buyer-generate-watermark Attack.* In this attack the buyer B gives the watermark embedding party a specially-crafted watermark in order to make the watermark trivially easy to remove after it is embedded into content. This attack was first mentioned in (Memon and Wong, 2001).
- *Buyer-in-the-middle Attack.* In this active attack the buyer B monitors all communications between all parties involved. The buyer B is then allowed to insert new messages, delete messages and modify messages to his or her advantage. This is a customised attack on *communication security* (Section 2.2).
- *Seller-CA Conspiracy Attack.* In this attack the seller S and CA conspire with each other to reveal the buyer B's watermark W. The seller S then frames the buyer B by embedding W into content not bought by B and distributes copies of it. We exclude conspiracy between S and B, as if they are allowed to (or willingly) conspire, this defeats the very purpose of a buyer-seller watermarking

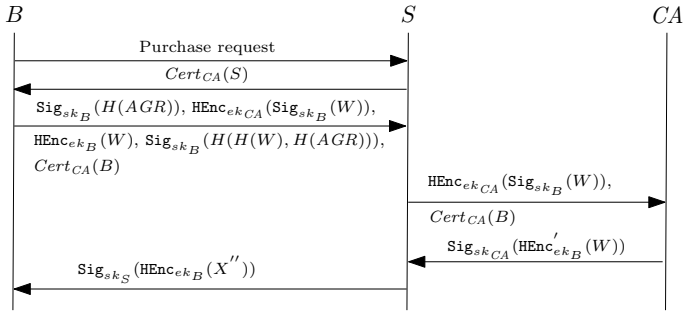

Figure 1: Ibrahim *et al.*'s Protocol (IAS 2007): Watermark Generation/Insertion.

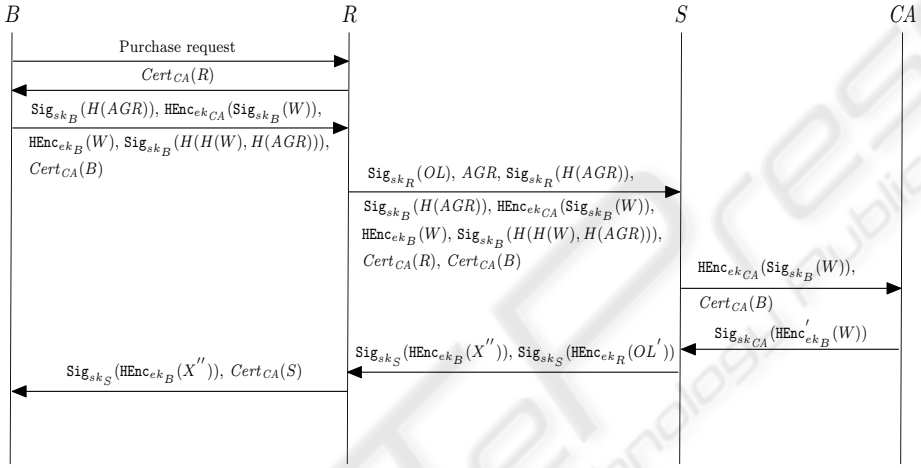

Figure 2: Ibrahim *et al.*'s Protocol (SECRYPT 2007): Watermark Generation/Insertion.

protocol. This attack was first mentioned in (Choi et al., 2003).

4.2 Attacks

We analyse the protocols by first giving the main idea of our exploitation, and then how the attacks defined in Section 4.1 can be performed.

Attack 1: Buyer-generate-watermark Attack on Both Protocols. The idea behind this attack is for B to remove the watermark W from the marked content X'' that B received from S. The attack will be successful on Ibrahim *et al.*'s protocols since in these protocols *B generates watermark* W. In fact, such an attack was mentioned by Memon and Wong (Memon and Wong, 2001), who recommended that B should not generate W due to the possibility of the buyer-generate-watermark attack. This is the main reason that a trusted third party called a watermark authority is used in (Memon and Wong, 2001) to generate W. However, Ibrahim *et al.*, in an attempt to prevent the conspiracy problem (Section 3), remove this watermark authority without giving a solution as to how to prevent a dishonest B from generating an ill-formed watermark.

Attack 2: Buyer-in-the-middle Attack on Both Protocols. The idea behind this attack is for B to modify protocol messages and generate a different watermark for a different message, so that S will fail to prove B's act of distributing content illegally, even when B generates a proper watermark W.

In the following we demonstrate how the attack works on Ibrahim *et al.*'s protocols by modifying the protocol steps given in Section 3.1, starting from Step 5. We note that the same attack can be deployed for the second protocol. Figure 3 further illustrates this attack on both protocols.

5′. Different from the original proposal, the buyer B generates three watermarks W_1, W_2 and W_3 instead of one watermark. The buyer B then generates two signatures, $\mathtt{Sig}_{sk_B}(W_1)$ and $\mathtt{Sig}_{sk_B}(W_2)$. After that, the buyer B encrypts $\mathtt{Sig}_{sk_B}(W_1)$ and $\mathtt{Sig}_{sk_B}(W_2)$, resulting in $\mathtt{HEnc}_{ek_{CA}}(\mathtt{Sig}_{sk_B}(W_1))$ and $\mathtt{HEnc}_{ek_{CA}}(\mathtt{Sig}_{sk_B}(W_2))$.

6′. Buyer B encrypts W_2 as $\mathtt{HEnc}_{ek_B}(W_2)$.

7′. Buyer B generates hash value $H(W_3)$ and signature $\mathtt{Sig}_{sk_B}(H(H(W_3),H(AGR)))$.

8′. Buyer B sends the signature $\mathtt{Sig}_{sk_B}(H(AGR))$, the encrypted signature $\mathtt{HEnc}_{ek_{CA}}(\mathtt{Sig}_{sk_B}(W_1))$, the encrypted watermark $\mathtt{HEnc}_{ek_B}(W_2)$ and the signature $\mathtt{Sig}_{sk_B}(H(H(W_3),H(AGR)))$, together with $Cert_{CA}(B)$ to S (to R for the protocol proposed in SECRYPT 2007).

9′. **When the seller S sends $\mathtt{HEnc}_{ek_{CA}}(\mathtt{Sig}_{sk_B}(W_1))$ to *CA*, the buyer B intercepts the message and sends $\mathtt{HEnc}_{ek_{CA}}(\mathtt{Sig}_{sk_B}(W_2))$ instead**. So *CA* decrypts $\mathtt{HEnc}_{ek_{CA}}(\mathtt{Sig}_{sk_B}(W_2))$ to obtain the signature $\mathtt{Sig}_{sk_B}(W_2)$, which is then verified to obtain W_2. Next *CA* re-encrypts W_2 with B's encryption key as $\mathtt{HEnc}'_{ek_B}(W_2)$ and signs it to obtain $\mathtt{Sig}_{sk_{CA}}(\mathtt{HEnc}'_{ek_B}(W_2))$. This signature is sent to S.

10′. Seller S verifies $\mathtt{Sig}_{sk_{CA}}(\mathtt{HEnc}'_{ek_B}(W_2))$ to obtain $\mathtt{HEnc}'_{ek_B}(W_2)$. This is identical to $\mathtt{HEnc}_{ek_B}(W_2)$ given by the buyer B. So for the seller S, the comparison $H(\mathtt{HEnc}_{ek_B}(W_2)) = H(\mathtt{HEnc}'_{ek_B}(W_2))$ will be true. Seller S continues the protocol since both hash values are identical.

12′. **Subsequently the buyer watermark that is embedded into content is W_2**, which is different from W_1 in $\mathtt{HEnc}_{ek_{CA}}(\mathtt{Sig}_{sk_B}(W_1))$ possessed by S.

Following from the above steps, recall from Section 3.1 that when an illegal content $Y = X \oplus V \oplus W_2$ is found, one of the objects sent by S to *CA* is $\mathtt{HEnc}_{ek_{CA}}(\mathtt{Sig}_{sk_B}(W_1))$. For B to be found guilty, *CA* retrieves W_1 from $\mathtt{HEnc}_{ek_{CA}}(\mathtt{Sig}_{sk_B}(W_1))$, signs it and re-encrypts it as $\mathtt{HEnc}_{ek_A}(\mathtt{Sig}_{sk_B}(W_1))$ and passes this new encrypted object to the arbiter A. Upon receiving the encrypted object, the arbiter A decrypts and retrieves W_1. The arbiter A then runs the detection algorithm, expecting to detect W_1 from Y. However, due to the interception by B in Step 9′ above, the watermark that is embedded in this content is W_2, instead of W_1. Hence A will fail to detect B's watermark and will declare B as innocent. We also note that the reason that this attack can be deployed is due to the protocols introduction of Step 8 and Step 9 (Section 3.1) in order to avoid the *Buyer's participation in the dispute resolution problem*.

Furthermore, due to Step 5′ and Step 7′, a third watermark W_3 is used by B to generate the signature $\mathtt{Sig}_{sk_B}(H(H(W_3),H(AGR)))$. Arbiter A will not be able to match the watermark W_2 extracted from the illegal copy Y to the purchase order based on this signature, and thus cannot be certain that B bought this content.

Attack 3: Seller-CA Conspiracy Attack on Both Protocols. The main idea behind this attack is based on the fact that *CA knows* the buyer watermark W, and essentially has the similar responsibility of a watermark authority (which generates buyer's watermark) in the existing protocols they observed (e.g. (Lei et al., 2004)).

This can be seen from Step 9 in Section 3.1. In this step the seller S sends *CA* an encrypted signature $\mathtt{HEnc}_{ek_{CA}}(\mathtt{Sig}_{sk_B}(W))$ that contains W, and *CA* is tasked to retrieve the watermark W, re-encrypt it and sign it with *CA*'s signing key. Hence *CA* can store a copy of W when retrieving it from the signature, and then sends this copy to S. Thus the claim of avoiding the *conspiracy problem* fails.

4.3 Comments

We now comment on what we have learnt from our analysis in the previous section.

- *Should the buyer be allowed to generate the watermark?* In most buyer-seller watermarking protocols (Choi et al., 2003; Goi et al., 2004; Ju et al., 2002; Lei et al., 2004; Memon and Wong, 2001) a trusted third party is introduced to generate the buyer's watermark. It is tempting to remove this trusted third party so that a protocol can be more efficient and involve only the buyer and the seller, as has done in Ibrahim *et al.*'s protocols. However, from **Attack 1**, if we prefer the buyer to generate the watermark, then we need to have a mechanism to prevent the buyer from generating a weak watermark which, when embedded, can later be easily removed. Such a mechanism was proposed in asymmetric and anonymous fingerprinting schemes (Pfitzmann and Schunter, 1996; Pfitzmann and Waidner, 1997). In these schemes the buyer needs to prove in zero-knowledge (Fiat and Shamir, 1987) to the seller that the generated watermark is well-formed. While these schemes are elegant, there have been suggestions (Goi et al., 2004; Ju et al., 2002) that they are more complicated than a protocol with a trusted third party. In summary, unless new and simpler mechanisms are found, we should try to avoid watermark generation by the buyer.
- *Provide secure communication channel through well-established methods.* In **Attack 2** we see that the buyer can intercept and replace messages transmitted between the seller and the *CA*. This happens mainly because the protocols fail to safeguard the communication between the seller and the *CA*, although much care has been taken to ensure secure transmission between the buyer and

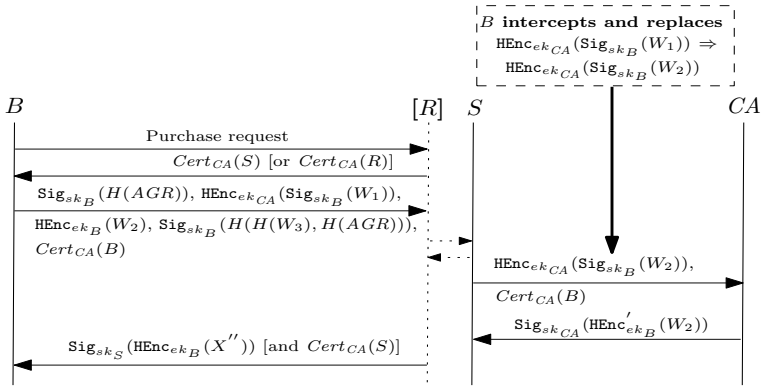

Figure 3: Ibrahim *et al.*'s Protocols: Attack.

the seller. To avoid this pitfall, we think a standard and safer approach should be followed. This is to secure the communications between all involved parties based on well-established protocols in the literatures (Bellare et al., 2000; Boyd and Mathuria, 2003; ISO, 1998), or based on a standard protocol such as SSL/TLS (Dierks and Rescorla, 2006). We can then construct the main part of buyer-seller watermarking on top of this secure channel.

- *Make explicit the trust assumptions on the TTPs.* In their protocols, Ibrahim *et al.* mentioned that *CA* is fully trusted. This means *CA* will not conspire with the seller. However, we can still run **Attack 3** for two reasons.
  - The first is that the watermark authority in the other protocols should also be fully trusted and the claims in (Ibrahim et al., 2007a; Ibrahim et al., 2007b) that they face conspiracy problems is rather controversial.
  - The second reason is that *CA* in these protocols processes the watermark, which means that *CA* has access to the watermark and hence is essentially similar to the watermark authority in other protocols.

  Ultimately, this bring us to conclude that if we properly define the trust assumptions on the TTPs (i.e. *CA* and the watermark authority), the *conspiracy problem* of Section 3 may not be an issue at all.
- *Differentiate between properties and other "problems".* In Section 3, it can be observed that the "problems" stated by Ibrahim *et al.* are rather disparate in nature. If we examine them more carefully, we see that the *conspiracy problem* and the *unbinding problem* are attacks on the main property known as framing resistance, defined in Section 2.2. Similarly, the *man in the middle attack* relates to an attack on communication security. However the *buyer's participation in the dispute resolution problem* and the *practice applicability problem* actually reflect the *framework* of a buyer-seller watermarking protocol. This confusingly merges the security properties required by such a protocol with the practical considerations and operations in an instantiation of it. We suggest that it is preferable to properly categorise these "problems" so that a protocol can be studied in a more systematic way, or by clearly identifying the operational environment (*framework*) and then well defining the security properties. Ambiguity in this can lead to attacks of the type we have described.

5 CONCLUSIONS

We have analysed two buyer-seller watermarking protocols, and showed that both protocols contain flaws. Especially crucial is that these protocols build upon existing ones by changing the design to gain additional properties, but in the process new weaknesses (some of which were previously known) are introduced into the protocols. In particular our buyer-generate watermark attack, exploits the fact that the buyer is allowed to generate the watermark; the buyer-in-the-middle attack, exploits the failure to contact the buyer during dispute resolution, and the Seller-CA conspiracy attack, exploits the ambiguity concerning the trust assumption of these protocols. We have also commented on how such design flaws should be avoided in future protocols.

REFERENCES

Bellare, M., Pointcheval, D., and Rogaway, P. (2000). Authenticated Key Exchange Secure against Dictionary

Attacks. In Preneel, B., editor, *Advances in Cryptology - EUROCRYPT 2000*, volume 1807 of *Lecture Notes in Computer Science*, pages 139–155. Springer-Verlag.

Blakley, G. R., Meadows, C., and Purdy, G. B. (1985). Fingerprinting Long Forgiving Messages. In Williams, H. C., editor, *Advances in Cryptology - CRYPTO 1985*, volume 218 of *Lecture Notes in Computer Science*, pages 180–189. Springer-Verlag.

Boyd, C. and Mathuria, A. (2003). *Protocols for Authentication and Key Establishment*. Information Security and Cryptography Series, Springer-Verlag.

Camenisch, J. (2000). Efficient Anonymous Fingerprinting with Group Signatures. In Okamoto, T., editor, *Advances in Cryptology - ASIACRYPT 2000*, volume 1976 of *Lecture Notes in Computer Science*, pages 415–428. Springer-Verlag.

Choi, J.-G., Sakurai, K., and Park, J.-H. (2003). Does It Need Trusted Third Party? Design of Buyer-Seller Watermarking Protocol without Trusted Third Party. In Zhou, J., Yung, M., and Han, Y., editors, *Applied Cryptography and Network Security - ACNS 2003*, volume 2846 of *Lecture Notes in Computer Science*, pages 265–279. Springer-Verlag.

Dent, A. and Mitchell, C. (2004). *User's Guide to Cryptography and Standards*. Artech House.

Dierks, T. and Rescorla, E. (2006). The TLS Protocol Version 1.1. *RFC 4346*.

Fiat, A. and Shamir, A. (1987). How to prove yourself: Practical solutions to identification and signature problems. In Odlyzko, A. M., editor, *Advances in Cryptology - CRYPTO 1986*, volume 263 of *Lecture Notes in Computer Science*, pages 186–194. Springer-Verlag.

Goi, B.-M., Phan, R. C.-W., Yang, Y., Bao, F., Deng, R. H., and Siddiqi, M. U. (2004). Cryptanalysis of Two Anonymous Buyer-Seller Watermarking Protocols and an Improvement for True Anonymity. In Jakobsson, M., Yung, M., and Zhou, J., editors, *Applied Cryptography and Network Security - ACNS 2004*, volume 3089 of *Lecture Notes in Computer Science*, pages 369–382. Springer-Verlag.

Ibrahim, I. M., El-Din, S. H. N., and Hegazy, A. F. A. (2007a). An Effective and Secure Buyer-Seller Watermarking Protocol. In *Third International Symposium on Information Assurance and Security (IAS 07), IEEE Computer Society Press*, pages 21–26.

Ibrahim, I. M., El-Din, S. H. N., and Hegazy, A. F. A. (2007b). An Effective and Secure Watermarking Protocol for Digital Rights Protection Over the Second-Hand Market. In *SECRYPT 2007 - International Conference on Security and Cryptography*, pages 263–268.

ISO (1998). Information Technology - Security Techniques - Entity Authentication Mechanisms - Part 3: Entity Authentication Using a Public Key Algorithm ISO/IEC 9798-3. *ISO/IEC International Standard, 2nd Edition*.

Ju, H. S., Kim, H. J., Lee, D. H., and Lim, J. I. (2002). An Anonymous Buyer-Seller Watermarking Protocol with Anonymity Control. In Lee, P. J. and Lim, C. H., editors, *Information Security and Cryptology - ICISC 2002*, volume 2587 of *Lecture Notes in Computer Science*, pages 421–432. Springer-Verlag.

Lei, C.-L., Yu, P.-L., Tsai, P.-L., and Chan, M.-H. (2004). An Efficient and Anonymous Buyer-Seller Watermarking Protocol. *IEEE Trans. on Image Processing*, 13(12):1618–1626.

Memon, N. and Wong, P. W. (2001). A Buyer-Seller Watermarking Protocol. *IEEE Trans. on Image Processing*, 10(4):643–649.

Paillier, P. (1999). Public-key Cryptosystems Based on Composite Degree Residuosity Classes. In Stern, J., editor, *Advances in Cryptology - EUROCRYPT 1999*, volume 1592 of *Lecture Notes in Computer Science*, pages 223–238. Springer-Verlag.

Pfitzmann, B. and Schunter, M. (1996). Asymmetric Fingerprinting. In Maurer, U. M., editor, *Advances in Cryptology - EUROCRYPT 1996*, volume 1070 of *Lecture Notes in Computer Science*, pages 84–95. Springer-Verlag.

Pfitzmann, B. and Waidner, M. (1997). Anonymous Fingerprinting. In Fumy, W., editor, *Advances in Cryptology - EUROCRYPT 1997*, volume 1233 of *Lecture Notes in Computer Science*, pages 88–102. Springer-Verlag.

Rivest, R. L., Shamir, A., and Adleman, L. (1978). A Method for Obtaining Digital Signatures and Public-Key Cryptosystems. *Commun. of the ACM*, 2(2):120–126.

Wagner, N. R. (1983). Fingerprinting. In *IEEE Symposium on Security and Privacy*, pages 18–22.