

Efficient Anonymous Fingerprinting with Group Signatures (Extended Abstract)

Jan Camenisch

IBM Research
Zurich Research Laboratory
CH-8803 Rüschlikon
jca@zurich.ibm.com

Abstract. Fingerprinting schemes enable a merchant to identify the buyer of an illegally distributed digital good by providing each buyer with a slightly different version. Asymmetric fingerprinting schemes further prevent the merchant from framing a buyer by making the fingerprinted version known to the buyer only. In addition, an anonymous fingerprinting scheme allows the buyer to purchase goods without revealing her identity to the merchant. However, as soon as the merchant finds a sold version that has been (illegally) distributed, he is able to retrieve a buyer's identity and take her to court.

This paper proposes a new and more efficient anonymous fingerprinting scheme that uses group signature schemes as a building block. A byproduct of independent interest is an asymmetric fingerprinting scheme that allows so-called two-party trials, which is unmet so far.

1 Introduction

Today's computer networks allow the trading of digital goods in an easy and cheap way. However, they also facilitate the illegal distribution of (copyrighted) data. Fingerprinting schemes are a method for supporting copyright protection. The idea is that a merchant sells every customer a slightly different "copy" of the good. For instance, in the case of an image, the merchant could darken or lighten some pixels. Of course, the fingerprint must be such that a buyer cannot easily detect and remove it. When the merchant later finds an illegally distributed copy, he can recognize the copy by its fingerprints and then hold its buyer responsible. A number of authors (cf. [8]) have studied methods to achieve this for various kinds of digital goods. Research is ongoing in this area.

Whereas fingerprinting as such is a technique that was already used in the previous century, security against colluding buyers was achieved only recently [2,3]. Such schemes tolerate a collusion of buyers up to a certain size, i.e., a collusion cannot produce a copy such that the merchant cannot trace it back to at least one of the colluders. The first such schemes that were proposed are *symmetric*, meaning that the merchant knows which copy a buyer gets [2,3].

Thus a malicious merchant could spread himself the version sold to some buyer and then accuse that buyer of having done so.

This problem is overcome by *asymmetric* schemes [1,12,14]. Here, the buyer sends the merchant a commitment to a secret she chose. Then the two carry out a protocol at the end of which the buyer possesses the desired digital good, fingerprinted with the chosen secret, whereas the merchant does not learn anything. Hence, whenever the merchant is able to present a sufficiently large fraction of the secret contained in a buyer's commitment, he must have found the copy a buyer bought (and distributed) and the buyer is therefore considered guilty.

In both symmetric and asymmetric schemes, the merchant needs to know a buyers' identity to be able to take her to court if she distributes the purchased copy. To protect buyers' privacy and match with anonymous digital payment systems, Pfitzmann and Waidner introduce *anonymous* asymmetric fingerprinting [13]. Here, a buyer must no longer identify herself for purchasing and remains anonymous as long as she keeps the purchased good secret, i.e., does not distribute it. More precisely, the merchant can learn a buyer's identity only if he obtains her purchased copy. This kind of scheme involves a further party, called registration center, at which all buyers are required to register prior to any purchase. Pfitzmann and Waidner also provide a general modular construction consisting of two building blocks. One handles the registration of buyers and the generation of the to-be-embedded information and the other building block is a method to embed committed information into the to-be-sold data. More precisely, the latter uses an error and erasure-correcting code together with an asymmetric fingerprinting scheme to guarantee that at least for one of the codewords *all* her committed secret bits can be extracted from a copy found.

The first building block uses general zero-knowledge proof techniques and renders the resulting scheme rather inefficient and hence it is merely considered a "proof of existence" [13]. The second building block can be realized efficiently in terms of computations [11]. However, the use of the error and erasure-correcting code prohibitively enlarges the number of bits that need to be embedded.

Recently, Pfitzmann and Sadeghi [11] presented an efficient replacement for the first part of this construction. It is derived from the anonymous e-cash scheme by Brands [4]. More precisely, it uses its property that coins are anonymous when spent once but reveal a user's identity when spent twice. However, the resulting scheme has the drawback that a buyer must register once for each purchase and that the merchant has to contact the registration center to retrieve the identity of a malicious buyer.

This paper presents an anonymous fingerprinting scheme that overcomes these drawbacks using group signature schemes as its main building block. A *group signature scheme* (e.g., [7,9]) allows a member of a group of users to sign a message on the group's behalf. The scheme protects the privacy of signers in that the verifier has no means to determine which member originated a signature or whether two signatures stem from the same signer. However, to handle special cases of misuse by some user, there is a designated *revocation manager* who can indeed find the identity of a signature's originator.

The idea underlying our fingerprinting scheme is to have the buyer issuing a group signature on a message describing the deal. Opposed to an ordinary group signature scheme there is no (fixed) revocation manager. Instead, the buyer chooses a secret and public key pair for the revocation manager; this public key is then used for issuing the group signature, whereas the secret key gets embedded into the sold good. Thus, finding an illegally distributed copy puts the merchant in the position of the revocation manager for that particular group signature and he can retrieve the identity of the culprit. Due to the properties of group signature schemes, each buyer must register only once (registering basically amounts to join the group) and the merchant can retrieve a culprit's identity directly. One version of our scheme can even do without a registration center.

We also improve on the second building block: we exhibit a method for circumventing the use of error and erasure correction assuming a trusted third party (TTP). This TTP, however, needs only to be involved in the case that a malicious buyer is taken to court. This method can also be used to get a two-party trial for the asymmetric fingerprinting schemes [1,14]. We refer to Section 4 for an explanation of two- and three-party trials. Combining both our new building blocks gives an anonymous fingerprinting scheme that can tolerate larger collusions than previous ones and requires less administration from buyers and merchants.

2 Model of Anonymous Fingerprinting

Let $P_0 \in \{0,1\}^*$ denote some digital good (bit-string) that is fingerprintable, i.e., some of its bits can be changed such that (1) the result remains “close” to P_0 but (2) without knowing which particular bits were changed, altering “a good portion” of these bits is impossible without rendering the good useless. We refer to [3] for a formal definition of this “marking assumption”. Finally, let \mathcal{P} denote the set of all “close copies” of P_0 and ℓ be a security parameter (from now on we implicitly assume that ℓ is an input to all algorithms and protocols).

Definition 1. *An anonymous fingerprinting scheme involves a merchant, a buyer, and a registration center. Let c denote the maximal size of a collusion of buyers against which the scheme is secure. An anonymous fingerprinting scheme consists of the following five procedures.*

FKG-RC: *A probabilistic key setup algorithm for the registration center. Its output are the center's secret key x_C and its public key y_C , which is published authentically.*

FReg: *A probabilistic two-party protocol (FReg-RC, FReg-B) between the registration center and the buyer. Their common input are the buyer's identity ID_B and the center's public key y_C . The center's secret input is its secret key x_C . The buyer's output consists of some secret x_B and related information y_B . The center obtains and stores y_B and ID_B .*

FPri: A two-party protocol (FPri-M, FPri-B) between the merchant and the buyer. Their common input consists of y_C . The merchant's secret input is P_0 and a transaction number j and his output is a transaction record t_j . The buyer's secret input is x_B and y_B and her output consists of a copy $P_B \in \mathcal{P}$.

FRec: A two-party protocol between the merchant and the registration center. The merchant's input is a copy $\tilde{P} \in \mathcal{P}$, P_0 , and all transaction records t_i . The center's input consists of its secret key x_C and its list of y_B 's and ID_B 's. The merchant's output is a/the fraudulent buyer's identity together with a proof p that this buyer indeed bought a copy of P_0 , or \perp in case of failure (e.g., if more than c buyers colluded to produce \tilde{P}).

FVer: A verification algorithm, that takes as input the identity ID_B of an accused buyer, the public key y_C of the registration center, and a proof p and outputs 1 iff the proof is valid.

We require that the following conditions hold.

Correctness: All protocols should terminate successfully whenever its players are honest (no matter how other players behaved in other protocols).

Anonymity and unlinkability: Without obtaining a particular P_B , the merchant (even when colluding with the registration center) cannot identify a buyer. Furthermore, the merchant must not be able to tell whether two purchases were made by the same buyer. In other words, all data stored by the merchant and registration center and the merchant's view of a run of FPri must be (computationally) independent of the buyer's secret input ID_B , x_B , and y_B .

Protection of innocent buyers: No coalition of buyers, the merchant, and the registration center should be able to generate a proof \tilde{p} such that $FVer(ID_B, y_C, \tilde{p}) = 1$, if buyer ID_B was not present in the coalition.

Revocability and collusion resistance: There exist no polynomial-time algorithms $FCol$, $FPri-B^*$, and $FReg-B^*$ such that for any ID_1, \dots, ID_c we have $FRec(P_0, FCol(\tilde{P}, \mathcal{U})) \notin \{ID_1, \dots, ID_c\}$ with non-negligible probability, where $\mathcal{U} = \{FReg-B^*_{FReg-RC(i, y_C, x_C)}(y_C) \mid i \in \{ID_1, \dots, ID_c\}\}$ and $\tilde{P} = \{FPri-B^*_{FPri-M(P_0, y_C)}(y_C, \mathcal{U}, i) \mid i = 1, \dots, c\}$.

Some fingerprinting schemes allow the merchant to recover the identity of a fraudulent user without the help of the registration center, i.e., FRec is not a protocol but an algorithm.

Realizations of the procedures FPri and FRec typically involve a pair of sub-protocols, one to embed some secret, committed to by the buyer, into the digital good and one to recover the embedded data again. Let Com be a commitment scheme, i.e., a (deterministic) function that takes as input the string x to commit to and an additional (randomizing) input string α . A buyer can commit to some x by $C = Com(x, \alpha)$, where α is randomly chosen. We require that the distributions of commitments to different x 's are (computationally) indistinguishable. A commitment C can be opened by revealing x and α . We require that it is (computationally) infeasible to open a commitment in two ways, i.e., to find

pairs (x, x') and (α, α') such that $\text{Com}(x, \alpha) = \text{Com}(x', \alpha')$. If the value of the parameter α is not essential, we drop it for notational convenience.

Definition 2. Let P_0 be a fingerprintable good known only to the merchant and let $y = \text{Com}(x)$ be the buyer's commitment to some secret $x \in \{0, 1\}^\ell$. An embedding method for P_0 , x , and Com consists of the following two procedures:

Emb: A two-party protocol (Emb-M, Emb-B) between the merchant and the buyer.

The merchant's secret input is P_0 , the buyer's secret input is x and their common input is $y := \text{Com}(x)$. The buyer's output is $P_B \in \mathcal{P}$.

Rec: An algorithm that takes as input P_0 and a fingerprinted copy \tilde{P} of it.

The algorithm's output is the data x embedded into \tilde{P} .

We require that the following properties are fulfilled.

Correctness: $\forall x, P_0 : x = \text{Rec}(P_0, \text{Emb-B}_{\text{Emb-M}(P_0, y)}(x, y))$, where $y = \text{Com}(x)$.

Recovery and Collusion-Resistance: There are no polynomial-time algorithms Col and Emb-B^* such that there is a set \mathcal{U} of at most c bit-strings of length ℓ for which $\text{Rec}(P_0, \text{Col}(\tilde{\mathcal{P}}, \mathcal{U})) \notin \mathcal{U}$ with non-negligible probability, where $\tilde{\mathcal{P}} = \{\text{Emb-B}^*_{\text{Emb-M}(P_0, y)}(x, y) \mid y = \text{Com}(x), x \in \mathcal{U}\}$.

Zero-Knowledgeness: For all Emb-M^* there exists a simulator such that for all $x \in \{0, 1\}^\ell$ the output of the simulator and the view of Emb-M^* are (perfect/statistically/computationally) indistinguishable.

3 Group Signature Schemes

Definition 3. A group signature scheme consists of the following procedures:

GKG-M: A key setup algorithm for the membership manager M that outputs her secret key x_M and public key y_M .

GKG-R: A key setup algorithm for the revocation manager R that outputs her secret key x_R and public key y_R .

GReg: A probabilistic interactive protocol (GReg-M, GReg-U) between the membership manager and a group member U . Their common input is the group member's identity ID_U and y_M . If both parties accept, the group member's output is her secret key x_U and their common output is U 's membership key y_U .

GSig: A probabilistic algorithm that on input of x_U , y_M , y_R , and a message m outputs a group signature s on m .

GVer: An algorithm that on input of the group public key Y , an alleged signature s , and a message m outputs 1 if and only if the signature is valid.

GTrace: A algorithm which on input of the revocation manager's secret key x_R , the group's public key Y , a message m , and a signature s on m outputs the identity ID_U of the originator of the signature and a proof V that ID_U is indeed the originator.

The following security requirements must hold:

Correctness of signature generation: All signatures on any messages generated by any honest group member using GSig will be accepted by the verification algorithm.

Anonymity and unlinkability of signatures: Given two signature-message pairs, it is only feasible for the revocation manager to determine which group member(s) generated any of the signatures or whether the signatures have been generated by the same group member.

Unforgeability of signatures: It is feasible to sign messages only to group members (i.e., users that have run the registration protocol with the membership manager) or to the membership manager herself.¹

Unforgeability of tracing: The revocation manager cannot falsely accuse a group member of having originated a given signature.

No framing: No coalition of group members, the revocation manager, and the group manager can produce a signature that will be associated with a group member not part of the coalition.

Unavoidable traceability: No coalition of group members and the revocation manager (but excluding the membership manager) can generate a valid signature that, when its anonymity is revoked, cannot be associated with a group member.

To use the group signature scheme for our construction in Section 5, we require that the key setup algorithm GKG-R for the revocation manager can be run after the algorithms GKG-M and GReg. That is, we require that the revocation manager can change her keys after the scheme has been set up and without requiring group members to reselect their key material. This property is provided by many group-signature schemes (e.g., [5,7,10]).

4 Previous Fingerprinting Schemes

All current anonymous and asymmetric fingerprinting schemes [1,12,13,14] are based on the symmetric scheme of Boneh and Shaw [3]. This section presents this scheme briefly, giving only those details that are needed to describe our results.

4.1 Symmetric Fingerprinting

A symmetric fingerprinting scheme consists of a set of binary codewords (or marking patterns) $W = w_1, \dots, w_n$ that can be embedded into the digital good [3]. Each time a copy is sold a different word is embedded and thereby

¹ The membership manager can always invent a fake identity and register it as a group member. It is understood that if a signature turns out to originate from a fake identity, the membership manager is considered guilty.

assigned to the copy's buyer. Let $\overline{W} \subseteq W$ denote all assigned codewords. If a redistributed copy is found later it must contain some word \tilde{w} that is a combination of words from \overline{W} due to the marking assumption. A scheme is called c -secure if there exists an algorithm A such that if a coalition C of at most c buyers generates a copy that contains a word \tilde{w} , then $A(\tilde{w}) \in \overline{W}$. It is said to have ϵ -error if the probability that $A(\tilde{w})$ outputs a codeword that was not assigned to any buyer in C (but might have been assigned to an honest buyer) is at most ϵ . Boneh and Shaw [3] show that $\epsilon = 0$ is not possible. They provide a binary code Γ_0 that is c -secure, has $n = c$ codewords, and whose length l is polynomial in n (i.e., $O(n^3 \log(n/\epsilon))$). Because the number of codewords equals the maximal number of colluding users tolerated, this code has the property that, no matter what a collusion does, the merchant will be able to extract a codeword assigned to one of the colluders with high probability (i.e., greater than $1 - \epsilon$, with $\epsilon = 2n2^{-l/(2n^2(n-1))}$).

Based on this code, Boneh and Shaw construct a random code Γ_1 over Γ_0 , i.e., each codeword in Γ_1 consists of the concatenation of, say, L randomly chosen codewords from Γ_0 . Extraction of an embedded word from an illegally distributed copy will now in general no longer yield a codeword assigned to one of the colluders but only a word whose components (codewords from Γ_0) stem from codewords assigned to the colluding users. Because at least L/c components of the extracted word must stem from a codeword assigned to one of the colluders, the extracted word must match that colluder's codeword in at least L/c positions, provided that the (malicious) buyers do not know any of the codewords. Therefore, a member of the collusion can be found by comparing all assigned codewords with the extracted word (provided that the number of codewords in Γ_1 is not too large, cf. [3]). The resulting code Γ_1 has length $c^{O(1)} \log(n)$, where n is the number of codewords (or, equivalently, the number of possible buyers).

Remark. As the amount of bits that can be embedded in a particular good is usually fixed, the length of a codewords translate into a maximum size of collusions that can be tolerated and how many buyers the good can be sold to.

4.2 Asymmetric Fingerprinting

In a nutshell, the idea behind an asymmetric scheme is as follows. First the buyer commits to some secret. Then merchant and buyer engage in a secure two-party protocol (henceforth called APri), at the end of which the buyer has obtained a copy of the good with her secret and some serial number (chosen by the merchant) embedded, whereas the merchant obtains the buyer's signature on a text describing their deal and on a commitment to the buyer's secret. Later, when the merchant finds an illegally distributed copy, he should be able to extract one of the colluding buyers' secret and the serial number from that copy. Being able to produce a buyer's secret will presumably convince a judge of her guilt. This approach is proposed by Pfitzmann and Schunter [12] for use with the code Γ_0 , in which case the protocol APri is reasonably efficient. However, when used with Γ_1 , the protocol APri is rendered prohibitively inefficient with this approach.

Pfitzmann and Waidner [14] solve this problem as follows (Biehl and Meyer [1] independently proposed a similar solution): During protocol APri, merchant and buyer construct on the fly a code similar to Γ_1 , that is, they together choose L random codewords w_1, \dots, w_L from Γ_0 such that the first half of each w_i consists of bits chosen by the merchant (but not known to the buyer) and the second half consists of bits chosen by the buyer (of which the merchant gets to know the only commitments C_i). At the end of the protocol, the buyer obtains a copy of the digital good with the codewords w_1, \dots, w_L embedded in it, whereas the merchant gets commitments C_1, \dots, C_L of the parts the buyer chose. When finding an illegally distributed copy, the merchant extracts the embedded word, and then can use a similar decoding strategy as the one described earlier for Γ_1 (restricted to the parts of the codewords known to him). Thus he will be able to identify one of the colluding buyers and also learn about L/c of the values committed to by C_1, \dots, C_L of the identified colluder and hence prove her guilt.

All these schemes have the property that the judge will not be able to tell on her or his own whether the commitments indeed contain the values that the merchant presents (this is a property of every secure commitment scheme). Therefore, the accused buyer must take part in the trial (which seems a natural requirement) and will be found guilty only if she is not able to prove that most of her commitments do not contain the value presented by the merchant. This is called a three-party trial [14].

Subsequently, Pfitzmann and Waidner [13] improve on this and exhibit a new asymmetric fingerprinting scheme that allows two-party trials. This scheme has the property that the merchant can extract all secret bits of one of the colluding buyers. This is achieved by using an error and erasure-correcting code (EECC) on top of the scheme described in the previous paragraph [14]. In addition, the buyer now also signs the result of some one-way function applied to her secret bits, and thus the judge will be able to verify whether the merchant indeed presents a malicious buyer's secret bits by testing whether these bits are the function's pre-image of the value the buyer signed. Hence a trial could be held without the accused buyer. The price for this improvement is that the use of the EECC increases significantly the number of bits that need to be embedded because the code must be able to handle a large number of erasures. To give a rough idea of this increase, in this scheme the underlying code Γ_0 must have $n = O(1)c^2\ell$ codewords, whereas it is $n = O(1)c$ in the one described previously [14], where c is the size of the tolerated collusion and ℓ is the bit length of the buyer's (whole) secret to be embedded. (Recall that the bit length of codewords from Γ_0 is polynomial in n .) However, the purchase protocol for this new scheme can be realized quite efficiently [11]. Finally, we note that this asymmetric fingerprinting scheme in fact realizes the two procedures **Emb** and **Rec** of Definition 2, thus allowing the construction of an anonymous fingerprinting scheme as we will see.

4.3 Anonymous Fingerprinting

Anonymous fingerprinting takes asymmetric fingerprinting one step further in that the merchant no longer gets to know an honest buyer's identity. Of course,

if the merchant finds an illegally distributed copy, he must nevertheless be able to retrieve the identity of a malicious buyer.

Building on the asymmetric fingerprinting scheme proposed in the same paper, Pfitzmann and Waidner [13] construct an anonymous fingerprinting scheme as follows. They introduce an additional party, a registration center, at which a buyer has to register beforehand under her real identity. To do so, the buyer chooses a pseudonym and a public/secret key pair of any signature scheme and receives a certificate for the public key and pseudonym. When purchasing some digital good from the merchant, the buyer commits to the certificate, the public key, the pseudonym, and a signature under this public key on a text describing their deal and sends these commitments to the merchant. She then proves to the merchant (in zero knowledge) that what is contained in the commitment is sound. Upon this, the two parties use the embedding protocol *Emb* as realized by the asymmetric scheme [13] described last in the previous section. As a result, the buyer obtains a copy with all the committed information embedded into it, whereas the merchant learns nothing. Apart from the commitment, the merchant obtains no information about the buyer during the transaction, but is assured that if he later finds an illegally distributed copy he will obtain all identifying information. However, the only known realization of this approach requires general zero-knowledge proof techniques, which are rather inefficient and thus the resulting anonymous fingerprinting scheme is considered an existence result [13].

Pfitzmann and Sadeghi [11] replace this general construction by an explicit and efficient one derived from the digital payment system by Brands [4]. Coins in that payment system are anonymous, but contain some identifying information that can be extracted as soon as a user spends a coin more than once (and only then). This information will then allow the bank to obtain the double-spender's identity. Pfitzmann and Sadeghi exploit this property as follows. The registration center plays the role of the bank and issues anonymous coins to registering buyers. Then, when purchasing some digital good, the buyer presents such a coin to the merchant. If the coin is valid, the merchant will be convinced that it contains information that will allow the registration center to retrieve the buyer's identity. Finally, they use the asymmetric fingerprinting scheme [13] such that the identifying information contained in the coin will be embedded in the sold copy. Owing to the algebraic properties of the payment system, the resulting protocol is quite efficient. However, two disadvantages remain: (1) a buyer must register with the center before each purchase and (2) the merchant must contact the center to learn the identity of a malicious buyer.

In the following section we provide another replacement for the general construction by Pfitzmann and Waidner [13] that uses a group signature scheme, is efficient, and overcomes these two restrictions. That is, our construction allows the merchant to directly identify the buyer of an illegally distributed copy and the buyer needs to register only once (or even not at all, depending on the kind of group signature scheme used, as we shall see).

5 Anonymous Fingerprinting Using Group Signatures

In this section we show how the asymmetric fingerprinting scheme [13] and any suitable group signature scheme can be combined to achieve an anonymous fingerprinting scheme. Suitable means that the key setup of the revocation manager can run after the registration of group members. This is true for many known group signature schemes (e.g., [5,7,10]).

The idea underlying our construction is that the registration center in the anonymous fingerprinting scheme plays the role of the *membership* manager in the group signature scheme. Every user that registers at the center then becomes a member of the group in the group signature scheme, i.e., the group consists of all registered buyers. When a user wants to buy some digital good P_0 , he first runs the group signature scheme's key-generation protocol for the *revocation* manager and gets a key pair, say y_R and x_R . Then the buyer signs a document describing the deal using the group signature scheme, where y_R is used as the revocation manager's public key. Note that a different revocation manager's key pair is used for every instance of the purchase protocol. Finally, the merchant and the buyer carry out the asymmetric fingerprinting protocol with respect to x_R , i.e., such that x_R is embedded into P_0 . Whenever the merchant obtains an (illegally distributed) copy \tilde{P} , he can extract² x_R —the secret key corresponding the y_R . This puts him into the position of the revocation manager for the instance of the group signature scheme that used y_R , and hence he can revoke the buyer's anonymity and identify her.

More formally, our anonymous fingerprinting scheme is as follows. Let (GKG-M, GKG-R, GReg, GSig, GVer, GTrace) be a suitable group signature scheme and (Emb, Rec) be an embedding protocol and a recovery algorithm for a commitment scheme Com as provided by the asymmetric fingerprinting scheme [13]. Let P_0 denote the digital good for sale.

FKG-RC: The registration center runs GKG-MM to get the key pair (y_M, x_M) and publishes y_M .

FReg: The center and the buyer run (GReg-M, GReg-U). The buyer gets y_U and x_U . The center gets and stores ID_U and y_U .

FPrint: Let m be the text that describes the deal. The buyer first runs GKG-R to obtain a key pair (y_R, x_R) , signs m by computing $\sigma := \text{GSig}(y_U, (y_R, y_M), m)$, and sends the merchant σ , y_R , and $y = \text{Com}(x_R)$. The buyer proves to the merchant that y indeed commits to the secret key corresponding to y_R . The merchant verifies σ using GVer and, if it was successful, the two parties engage in the protocol Emb, where the merchant's input is P_0 and y , the buyer's input is x_R and y , and the buyer's output is a copy P_B of P_0 .

FRec: Let \tilde{P} be a copy of P_0 produced by at most c dishonest buyers. Running Rec on \tilde{P} , the merchant obtains some x_R . This allows him to compute y_R and find the group signature σ in his database. Running GTrace($x_R, (y_R, y_M), m, \sigma$),

² Here, we assume that fewer than c buyers colluded to generate the distributed copy.

the merchant learns the identity of one of the buyers in the collusion that produced \tilde{P} .

Theorem 1. *Given a secure group signature scheme where the key setup of the revocation manager can be run after the registration of group members and a secure and collusion-resistant embedding method, the above construction is a secure anonymous asymmetric fingerprinting scheme.*

Proof. Correctness: By inspection.

Anonymity and unlinkability: All information the merchant obtains during a purchase is a group signature on a message that describes the deal. Owing to the properties of the embedding scheme, the merchant gets no information about the secret key x_R corresponding to y_R . Because group signatures are anonymous and unlinkable for everybody but the one knowing x_R , purchases are anonymous and unlinkable.

Protection of innocent buyers: In order to frame an innocent buyer a coalition would either have to produce a group signature with respect to some public key y'_R they can choose, or they would have to come up with a fingerprinted copy containing the secret key for some y_R that the buyer used in a purchase. The first attack is prevented by the “no-framing” property of the group signature scheme. The second attack is infeasible due to zero-knowledge property of the embedding protocol.

Revocability and collusion resistance: Given the collusion resistance and the correctness of the embedding scheme, the merchant can recover at least the secret key for one of the y_R 's that was used by a member of the collusion if it contains fewer than c buyers. Knowing the secret key of some y_R places the merchant in the position of the revocation manager in the group signature scheme and hence he can revoke the anonymity of the buyer/group member.

5.1 Discussion and Comparison with Previous Solutions

It is easy to see that buyers in our anonymous fingerprinting scheme need to register only once and can then buy many goods without these transactions being linkable. Whether the merchant is able to retrieve the identity of a malicious buyer on his own depends on the group signature scheme chosen. We discuss this briefly as well as other properties the fingerprinting scheme will have as a function of the type of group signature scheme that is applied.

Most newer group signature schemes (including [6,7]) can be used for our construction. These schemes have the property that the group's public key and the length of signature are independent of the group's size. A signature in those schemes typically contains a randomized encryption of identifying information under the revocation manager's public key. If a group signature scheme is used that allows the revocation manager to trace a signature without any interaction with the membership manager, it follows that the merchant need not interact with the registration center to identify a malicious buyer. This is possible for

instance with the recent group signature scheme by Camenisch and Michels [6]. There, a group member chooses her own RSA modulus that upon signing is encrypted by the revocation manager's public key. Thus, when the membership manager (aka registration center) enforces that the most significant bits are set to the identity of the group member (aka buyer), a direct identification is possible. The efficiency of the resulting fingerprinting scheme is governed by the embedding protocol **Emb**. Because this is the same for the scheme by Pfitzmann and Sadeghi [11], the two schemes have about the same efficiency. Thus, the main advantage of our scheme is that it overcomes the latter's drawback that a buyer must register prior to each purchase and that the merchant needs to contact the registration center to identify a malicious buyer.

If we apply the group signature scheme described in [5] and assume a public key infrastructure, we do not even need a registration center. This group signature scheme works for any semantically secure public key encryption scheme for which the revocation manager can be used and the buyer can have any public key signature or identification scheme that fulfills certain properties [5]. This includes for instance the RSA, DSS, or Schnorr signature schemes. The group's public key in this scheme consists of a list of users' public keys and certificates on them. Thus, using this scheme, a buyer can simply present the merchant with any list of public keys and certificates among which she would like to hide and chooses some public and secret key of an encryption scheme. Then, using the group signature scheme, the buyer signs the purchase contract and engages with the merchant in protocol **FPrint**. The resulting fingerprinting scheme will not need a registration center at all and the merchant is able to identify a malicious buyer on his own. However, the merchant needs to store the list of all the public keys and certificates the buyer presents as well as the group signature, which is about the same size as this list. As long as the number of public keys presented by the buyer is not too large (i.e., much smaller than the number of bits of the sold good), the scheme's efficiency is governed by the embedding protocol **Emb**.

6 Replacing Error and Erasure Correction by TTPs

As described in Section 4, the asymmetric fingerprinting scheme [13] underlying our (and all other known) anonymous fingerprinting scheme uses an error and erasure-correcting code to guarantee the full recovery of one of the colluders' committed secret from a found copy. As mentioned earlier, this error and erasure correction significantly increases the number of bits that ultimately need to be embedded.

To be able to base our anonymous fingerprinting scheme on the more efficient asymmetric fingerprinting schemes [1,14] and thereby circumvent error and erasure correction, we extend the model by a trusted third party (TTP). This TTP will be responsible for identifying malicious buyers. Of course, the TTP must not be involved in normal operations but only in the case that it comes to a trial. Moreover, the trust to be put in the TTP shall be minimal, i.e., buyers need to trust the TTP only that it does not reveal identities at will and the

merchant needs to trust only that the TTP cooperates for identifying malicious players. Other than that, the TTP is not trusted, e.g., a coalition of the TTP and the merchant must not be able to frame an honest buyer. Serving as a TTP could for instance the judge who has to take part in the trial anyway. To reduce the risk of fraudulent behavior, the TTP could be distributed (techniques for this are standard).

The general idea for our scheme with a TTP is to use the group signature scheme in the way originally conceived: the TTP plays the role of the revocation manager and the role of the membership manager is assumed by the registration center. Then, we use one of the asymmetric fingerprinting schemes [1,14] but without the buyer identifying herself and with her signature scheme being replaced by the group signature scheme. Now, if the merchant finds an illegally distributed copy and extracts the embedded information, the trial can take place as in the asymmetric scheme with the difference that the TTP must first identify the accused buyer via the revocation mechanism of the group signature scheme. Thus, we have an anonymous fingerprinting scheme.

Owing to the three-party-trial nature of efficient asymmetric fingerprinting schemes [1,14], the merchant cannot provide evidence to the TTP (aka revocation manager) that the buyer he wishes to identify is indeed malicious. A dishonest merchant could take advantage of this and learn the identity of an honest buyer simply by accusing her. This can be prevented by doubling the length of the parts of the L codewords from I_0 that the buyer chooses and then requiring the anonymous buyer to verifiably encrypt (see, e.g., [5]) the first half of each of her (secret) parts under the TTP's public key. Then, the merchant stores these encryptions as part of his transcript. The rest of the scheme remains unchanged. Later, when finding an illegally distributed copy, extracting the embedded information and thereby linking the copy to a purchase transcript, the merchant sends the transcript together with the first half of the extracted buyer parts of codewords to the TTP. Receiving this, the TTP decrypts the verifiable encryptions and compares the result with the corresponding parts that the merchant claims to have extracted from the copy. If most of these match (the merchant must be allowed a certain error rate, see [1,14]), the TTP reveals the identity of the buyer (who can then be taken to court); otherwise the TTP refuses. After finding out the identity of one of the colluders, the merchant can take her to court as before.

It is easy to see that, as long as the TTP is honest, the merchant is guaranteed to learn the identity of a malicious buyer, whereas an honest buyer's anonymity is protected. Finally, the probability that a collusion of the TTP and the merchant can frame a buyer is the same as for the merchant in the underlying asymmetric scheme. With respect to efficiency, the number of bits that are embedded is at most a factor of 2 greater for the original asymmetric scheme.

We briefly describe how a TTP could also be used to achieve an asymmetric fingerprinting scheme with a two-party trial. The drawback of the asymmetric fingerprinting schemes [1,14] with a three-party trial is that a (malicious) merchant can accuse any buyer of misconduct, causing the buyer the inconvenience

of going to court to prove her innocence. This is because in these schemes, it is not possible for the judge to check whether the evidence provided by the merchant is real. Thus the judge must always start a trial.

This can be overcome by using a TTP (which could be the judge himself) in the same way as described earlier in this section, neglecting the group signature scheme entirely. This results in an asymmetric fingerprinting scheme where the judge could use the TTP to check the evidence before opening a trial. Hence, a merchant can no longer accuse an honest buyer, as long as the TTP remains honest. Moreover, if the buyer trusts the TTP, then she can also discard her purchase transcript.

References

1. I. Biehl and B. Meyer. Protocols for collusion-secure asymmetric fingerprinting. In *Proc. 14th STACS*, vol. 1200 of *LNCS*, pp. 213–222. Springer Verlag, 1997.
2. G. R. Blakley, C. Meadows, and G. B. Prudy. Fingerprinting long forgiving messages. In *Advances in Cryptology — CRYPTO '85*, vol. 218 of *LNCS*, pp. 180–189. Springer-Verlag, 1986.
3. D. Boneh and J. Shaw. Collusion-secure fingerprinting for digital data. In *Advances in Cryptology — CRYPTO '95*, vol. 963 of *LNCS*, pp. 452–465, 1995.
4. S. Brands. Electronic cash systems based on the representation problem in groups of prime order. In *Preproceedings of Advances in Cryptology — CRYPTO '93*, pp. 26.1–26.15, 1993.
5. J. Camenisch and I. Damgård. Verifiable encryption and applications to group signatures and signature sharing. Technical Report RS-98-32, BRICS, Department of Computer Science, University of Aarhus, Dec. 1998.
6. J. Camenisch and M. Michels. Separability and efficiency for generic group signature schemes. In *Advances in Cryptology — CRYPTO '99*, vol. 1296 of *LNCS*, pp. 413–430. Springer Verlag, 1999.
7. J. Camenisch and M. Stadler. Efficient group signature schemes for large groups. In *Advances in Cryptology — CRYPTO '97*, vol. 1296 of *LNCS*, pp. 410–424, 1997.
8. G. Caronni. Assuring ownership rights for digital images. In *Verlässliche IT-Systeme — VIS '95*, pp. 251–263. Vieweg Verlag, 1995.
9. D. Chaum and E. van Heyst. Group signatures. In *Advances in Cryptology — EUROCRYPT '91*, vol. 547 of *LNCS*, pp. 257–265. Springer-Verlag, 1991.
10. L. Chen and T. P. Pedersen. New group signature schemes. In *Advances in Cryptology — EUROCRYPT '94*, vol. 950 of *LNCS*, pp. 171–181, 1995.
11. B. Pfitzmann and A.-R. Sadeghi. Coin-based anonymous fingerprinting. In *Advances in Cryptology — EUROCRYPT '99*, vol. 1592 of *LNCS*, pp. 150–164. Springer Verlag, 1999.
12. B. Pfitzmann and M. Schunter. Asymmetric fingerprinting. In *Advances in Cryptology — EUROCRYPT '96*, vol. 1070 of *LNCS*, pp. 84–95. Springer Verlag, 1996.
13. B. Pfitzmann and M. Waidner. Anonymous fingerprinting. In *Advances in Cryptology — EUROCRYPT '97*, vol. 1233 of *LNCS*, pp. 88–102. Springer Verlag, 1997.
14. B. Pfitzmann and M. Waidner. Asymmetric fingerprinting for larger collusions. In *4th ACM CCS*, pp. 57–66. ACM press, Apr. 1997.