Non-interactive Watermark Detection for a Correlation-Based Watermarking Scheme

André Adelsbach, Markus Rohe, and Ahmad-Reza Sadeghi

Horst Görtz Institute for IT Security, Ruhr-Universität Bochum, Germany
andre.adelsbach@nds.rub.de
{rohe, sadeghi}@crypto.rub.de

Abstract. Cryptographic techniques have been deployed to securely prove the presence of a watermark in stego-data without disclosing any security critical information to the detecting party.

This paper presents a detailed practical construction and implementation results of a correlation-based non-blind watermarking scheme in the non-interactive zero-knowledge setting. We extensively describe the modifications and hurdles that had to be overcome to transform a well-known watermarking scheme – whose general detection principle is applied in many other known schemes – into a two-party setting where the critical detection input, i.e. the watermark vector and the original data is cryptographically concealed from the verifying party using a commitment scheme. Our prototype implementation is very efficient and is an evidence of the practical feasibility of zero-knowledge watermark detection.

Keywords: Watermark, detection, implementation, zero-knowledge.

1 Introduction

When using watermarks as evidence in applications, such as fingerprinting, dispute resolving or direct authorship proofs, the presence of a watermark, embedded by some party (e.g., a merchant or the author) has to be verifiable by another, not fully trusted party (e.g., a judge, a dispute resolver or a customer). Unfortunately, verifying the presence of a watermark in given data by means of the watermarking system's detection algorithm requires knowledge of the watermark, the watermarking key and, in non-blind watermarking systems, additionally the original data. Once this information was disclosed to a malicious party, it enables this party to perfectly remove the watermark without any perceptible quality degradation.

Adelsbach and Sadeghi [1] suggest to conceal the critical detection input from the potentially dishonest verifying party in commitments and to apply a zero-knowledge protocol in which a prover $\mathcal{P}$ proves to the verifying party $\mathcal{V}$ that the committed watermark is detectable in the alleged stego-data.[1] The protocol

[1] Other protocols have been proposed before, but these do not achieve the same level of security or have documented security flaws [2].

J. Dittmann, S. Katzenbeisser, and A. Uhl (Eds.): CMS 2005, LNCS 3677, pp. 129–139, 2005.

is zero-knowledge which guarantees that the verifier gains no knowledge on the embedded watermark.

In this paper we present the concrete construction of a zero-knowledge proof system for the watermarking scheme proposed by Cox et al. [3] as *one example* for the class of correlation-based and non-blind detectable watermarking schemes. Furthermore, we give a precise quantisation of the computation and communication complexities of the protocol. After minor transformations, the correlation value can be computed as a polynomial expression such that the entire zero-knowledge watermark detection protocol can be composed from elementary zero-knowledge sub-protocols and by using the commitment scheme's homomorphic property.

We want to stress that this zero-knowledge watermark detection paradigm can be applied to *any* watermarking scheme whose detection criterion can be expressed as a polynomial expression. This also includes more advanced embedding and detection strategies to improve robustness and imperceptibility with respect to the HVS as cited in [4]. We have chosen the scheme of Cox et al. [3], because it is a widely known example of correlation-based watermark detection and convenient to demonstrate the practical feasibility of strong zero-knowledge watermark detection.

Outline: Section 2 recapitulates the technical basics, i.e. the applied watermarking scheme and the cryptographic primitives. Section 3 treats all considerations and modifications of the original watermarking scheme when it is transformed into an efficient zero-knowledge protocol. In Section 4 we estimate the computation and communication complexities and present results of a prototype implementation.

2 Technical Preliminaries

2.1 Watermarking Scheme by Cox et al.

Here we shortly recall the major facts from the watermarking scheme by Cox et al. [3] for black and white still-images as the basis for our detection protocol.

Generation: The watermark vector WM consists of m (in the order of 1000) independently chosen $N(0,1)$-distributed coefficients.

Embedding: The discrete cosine transformation (DCT) is applied to the original image, resulting in $\hat{W}$. Let $\hat{W}^{[m]}$ denote the m coefficients carrying the watermark information which corresponds here to the m highest magnitude AC-coefficients in $\hat{W}$. Cox et al. originally propose three different equations to embed the watermark, yielding the m-dimensional vector $\hat{W'}^{[m]}$ of marked coefficients $\hat{W'}_i^{[m]}$ for $i = 0, \ldots, m-1$:

$$\hat{W'}_i^{[m]} = \hat{W}_i^{[m]} + \alpha \cdot WM_i \tag{1}$$

$$\hat{W'}_i^{[m]} = \hat{W}_i^{[m]} \cdot (1 + \alpha \cdot WM_i) \tag{2}$$

where the constant α denotes the strength of embedding. As the third equation $\hat{W'}_i^{[m]} = \hat{W}_i^{[m]} \cdot \left(e^{\alpha \cdot WM_i}\right)$ is practically not used, we omit its further discussion.

Substituting $\hat{W'}^{[m]}$ in $\hat{W}$ and applying the inverse discrete cosine transformation DCT^{-1} results in the watermarked image W'.

Detection: To decide whether a given watermark WM is contained in image W^* we extract a watermark candidate WM^* whose correlation value is computed against the watermark WM. In this extraction we first compute $\hat{W} = DCT(W)$ and $\hat{W}^* = DCT(W^*)$. Then set $\hat{W}^{[m]}$ to the m-highest magnitude coefficients of $\hat{W}$ and $\hat{W^*}^{[m]}$ to the corresponding coefficients (same position) of $\hat{W^*}$. Then WM^* is obtained by inverting the embedding equation (see Section 3.1). Finally, we compute the correlation

$$corr = \frac{WM \cdot WM^*}{\| WM^* \|} \tag{3}$$

and compare it to some given threshold S. If $corr \geq S$ then WM is considered to be present. Otherwise it is considered to be absent.

2.2 Cryptographic Primitives

Commitment Scheme. A *commitment scheme* is a cryptographic protocol that allows one party, the so-called *committer* $\mathcal{C}$, to commit himself to a message $s \in \mathcal{M}$ from the message space $\mathcal{M}$, such that the *recipient* $\mathcal{R}$ of the commitment C_s is assured that $\mathcal{C}$ is unable to change the value of s afterwards (*binding property*). At the same time s is kept secret from the recipient $\mathcal{R}$ (*hiding property*).

Protocols: A commitment scheme consists of two main protocol steps:

1. `Commit`(): To commit to a certain message $s \in \mathcal{M}$, $\mathcal{C}$ runs the algorithm $(C_s, sk_{C_s}) \leftarrow \mathtt{commit}(s)$ to obtain the commitment C_s to s and the corresponding *secret key* sk_{C_s} that allows $\mathcal{C}$ to open C_s correctly in the `Open`() protocol. The committer passes C_s to the recipient who saves it for further use.
2. `Open`(): To open C_s to $\mathcal{R}$, $\mathcal{C}$ sends the message s and the corresponding secret key sk_{C_s} to the recipient. With this information $\mathcal{R}$ is able to verify s regarding the previously received commitment C_s. If the verification has been successful, $\mathcal{R}$ outputs the message s, otherwise he rejects. We denote such a successful protocol run as $(\mathcal{C} : -; \mathcal{R} : s) \leftarrow (\mathcal{C} : s, sk_{C_s}; \mathcal{R} : -; C_s)$

We refer to [5] for a detailed introduction to commitment schemes.

The Concrete Commitment Scheme: We use the Damgård-Fujisaki (DF) integer commitment scheme [6] in our protocol. A commitment to a message $s \in \mathbb{Z}$ is computed as $C_s := g^s h^{sk_{C_s}} \mod n$, where n is the product of two safe primes, h is a random element of high order and its order has only large prime factors. g is a random element from $< h >$ and $\log_h g$ is unknown to $\mathcal{C}$. g, h

and n form together with some other public (security) parameters (cf. Section 4) the so-called *commitment description* $descr_{com}$. Instantiated in this manner, the DF commitment scheme is *statistically hiding* and *computationally binding* under the *strong RSA assumption.*

Homomorphic Property: The structure of the DF commitment scheme allows $\mathcal{R}$ to perform computations on secret values without knowledge of the corresponding opening information. This feature can be used to increase the efficiency of the watermark detection protocol. Let C_x and C_y be two commitments to the secret values x and y and γ be some publicly known integer. The committer $\mathcal{C}$, knowing sk_{C_x} and sk_{C_y}, can open the product $C_x \cdot C_y$ as $x + y$: $(\mathcal{C} : -; \mathcal{R} : x + y) \leftarrow (\mathcal{C} : x + y, sk_{C_x} + sk_{C_y}; \mathcal{R} : -; C_x \cdot C_y)$. Furthermore, $(C_x)^\gamma$ can be opened as $\gamma \cdot x$: $(\mathcal{C} : -; \mathcal{R} : \gamma \cdot x) \leftarrow (\mathcal{C} : \gamma \cdot x, \gamma \cdot sk_{C_x}; \mathcal{R} : -; (C_x)^\gamma)$ and $C_x \cdot g^\gamma$ can be opened as $\gamma + x$: $(\mathcal{C} : -; \mathcal{R} : \gamma + x) \leftarrow (\mathcal{C} : \gamma + x, sk_{C_x}; \mathcal{R} : -; C_x \cdot g^\gamma)$. Consequently, $\mathcal{R}$ can autonomously compute C_{x+y}, $C_{\gamma \cdot x}$ and $C_{\gamma+x}$, which can be opened accordingly by $\mathcal{C}$.

Elementary Zero-Knowledge Proof Systems. Interactive two-party proof systems involve a so-called *prover* $\mathcal{P}$ and a so-called *verifier* $\mathcal{V}$ where each of them has its own *private input* and both have access to some given *common input.* In our context, the common input consists of commitments of which $\mathcal{P}$ is aware of the secret messages and the corresponding secret keys as its private input. Applying such proof systems $\mathcal{P}$ convinces $\mathcal{V}$ that he is indeed able to open the commitments, provided as common input, correctly and that certain relations hold among their secret messages. There exist three security requirements for these proof systems: *Completeness:* If $\mathcal{P}$ and $\mathcal{V}$ act honestly, every run of the proof system will be accepted by $\mathcal{V}$. *Soundness* guarantees that a cheating prover (e.g. $\mathcal{P}$ has no opening information for the commitments) can trick $\mathcal{V}$ to accept the proof protocol only with a negligible probability. Finally, the *zero-knowledge* requirement guarantees that $\mathcal{V}$ gains no new knowledge from a protocol run beyond the assertion that has been proven.

We will make use of several elementary zero-knowledge proof protocols, which prove the multiplicative relation ($\texttt{PoK}_{mult}()$), the square relation ($\texttt{PoK}_{sq}()$) and the equality relation on committed values ($\texttt{PoK}_{eq}()$).[2] We use the multiplication protocol proposed by Damgård and Fujisaki [6], while the square and the equality proof are adapted from Boudot [7]. Finally, we use a proof system $\texttt{PoK}_{\geq 0}()$ which proves that a committed value is greater or equal to zero. An elegant proof system has been suggested by Lipmaa [8] and is based on a number theoretical result by Lagrange, which states that every positive integer x can be represented as a sum of four squares, i.e. $x = x_1^2 + x_2^2 + x_3^2 + x_4^2$. Hence, the proof system $\texttt{PoK}_{\geq 0}()$ can be composed by 4 square proofs, the homomorphic addition of $C_{x_1^2}, \ldots, C_{x_4^2}$ and the proof of an equality relation for $C_{x_1^2+x_2^2+x_3^2+x_4^2}$ and C_x.

Typically, zero-knowledge proofs are executed as interactive challenge-response protocols. However, there exists an efficient transformation to convert

[2] For example, $\texttt{PoK}_{mult}(C_c;\ C_a,\ C_b)$ denotes a zero-knowledge proof that $\mathcal{P}$ can open C_a, C_b and C_c, such that $a \cdot b = c$ holds.

the interactive version of a proof protocol into a non-interactive version in the random oracle model [9] where the prover performs a complete precomputation of a proof and passes it to the verifier. Our implementation uses the elementary sub-proofs in this efficient proof mode.

We will give the computational complexity of all protocols in terms of modular exponentiations (E), modular inversions (I) and modular multiplications (M), as these operations dominate their complexities. With $\mathsf{Comp}_{\mathcal{P}}$ we denote the *computational complexity* of the prover, whereas $\mathsf{Comp}_{\mathcal{V}}$ refers to that of the verifier. L denotes the computation expense of a 4-square Lagrange decomposition, for which an efficient probabilistic algorithm can be found in [10]. With $\mathsf{Comm}_{(\mathcal{P},\mathcal{V})}$ we denote the *communication complexity*, measured as the number of bits exchanged between $\mathcal{P}$ and $\mathcal{V}$.

The complexities of the basic protocols mainly depend on the security parameters of the DF commitment scheme, namely $|n|$, k, T, B and $C(k)$ (which we will denote as F). Here, $|n|$ denotes the binary length of the strong RSA modulus. B is an estimator for the upper bound of the order of $< h >$, such that $ord(< h >) \leq 2^B$, while T specifies the message space $\mathcal{M} = [-T, T]$. We use the parameter k to limit the maximum statistical distance (statistical zero knowledge property) between an accepting real and a simulated protocol view which is less than 2^{-k}. F (aka $C(k)$ in [6]) determines the challenge size and therefore the security parameter for the proof's soundness. As such, it limits the probability that a cheating prover is able to carry out an accepting proof to $< 2^{-|F|}$. For further details regarding these parameters we refer to [6]. Table 1 gives an overview of $\mathsf{Comp}_{\mathcal{P}}$, $\mathsf{Comp}_{\mathcal{V}}$ and $\mathsf{Comm}_{(\mathcal{P},\mathcal{V})}$, including the communication complexity for reasonably chosen security parameters (cf. Section 4).

Technical Remark: Watermarking schemes require computations on real numbers, while the applied DF commitment scheme supports integers. However, by scaling all real values by an appropriate factor λ (e.g. $\lambda = 10^{10}$ or 10^{20}) we can perform all required computations in the integer domain. For instance, the relation $a \cdot b = c$ is scaled as $(\lambda_a a) \cdot (\lambda_b b) = (\lambda_a \lambda_b) c$.

Table 1. Communication and computation complexities in the non-interactive proof mode

Relation	$\mathsf{Comp}_{\mathcal{P}}$	$\mathsf{Comp}_{\mathcal{V}}$	$\mathsf{Comm}_{(\mathcal{P},\mathcal{V})}$	[KBytes]
$\mathtt{PoK}_{mult}()$	$6E + 9M$	$9E + 3I + 6M$	$6\|F\| + 3\|T\| + 8k + 3B + 5$	0.66
$\mathtt{PoK}_{sq}()$	$4E + 6M$	$6E + 2I + 4M$	$4\|F\| + 2\|T\| + 5k + 2B + 3$	0.44
$\mathtt{PoK}_{eq}()$	$4E + 5M$	$6E + 2I + 4M$	$4\|F\| + \|T\| + 5k + 2B + 3$	0.38
$\mathtt{PoK}_{\geq 0}()$	$38E + 42M + L$	$30E + 10I + 23M$	$8\|n\| + 20\|F\| + 9\|T\| + 25k$ $+10B + 15$	3.13

3 Transformation into the ZK-Setting

For each embedding equation (1) and (2) we consider its inversion, yielding WM^* and rate its usability. Furthermore, we address the problem how $\mathcal{V}$ chooses the detection coefficients $\hat{W^*}^{[m]}$ as he is only aware of the committed version of $\hat{W}$.

3.1 Embedding Equations

The first step of the detection algorithm is to extract WM^* from the alleged stego-image $\hat{W}^*$, which, in non-blind detection, additionally involves the original image $\hat{W}$. In zero-knowledge watermark detection, $\mathcal{V}$ is only aware of the committed version $C_{\hat{W}^{[m]}} := (C_{\hat{W_1}^{[m]}}, \ldots, C_{\hat{W_m}^{[m]}})$ of $\hat{W}$ such that, after the extraction, he has to be convinced in zero-knowledge that the content of $C_{WM^*} := (C_{WM^*{}_1}, \ldots, C_{WM^*{}_m})$ has been obtained correctly.

Equation 1: In case WM was embedded according to Equation (1) then WM^* is obtained[3] as

$$\Delta_i := \alpha \cdot WM^*{}_i = \hat{W^*}_i^{[m]} - \hat{W}_i^{[m]}. \tag{4}$$

such that Δ_i is a difference of committed values, which can be easily computed in the committed domain by taking advantage of the homomorphic property of the commitment scheme.

Equation 2: In this case Δ_i is obtained as the quotient

$$\Delta_i := \alpha \cdot WM^*{}_i = \left(\hat{W^*}_i^{[m]} - \hat{W}_i^{[m]}\right) \,/\, \hat{W}_i^{[m]}. \tag{5}$$

To convince $\mathcal{V}$ in the committed domain that Δ_i in C_{Δ_i} has been computed correctly as $\Delta_i = \hat{W^*}_i^{[m]} \cdot \left(\hat{W}_i^{[m]}\right)^{-1} - 1$ an additional zero-knowledge proof has to be performed. Therefore, the computation of C_{Δ_i} at the beginning of the detection protocol described in Section 3.4 has to be extended by an additional multiplication subproof and a proof that $\left(\hat{W}_i^{[m]}\right)^{-1}$ was computed correctly.[4] Clearly, the entire detection protocol can be extended by the described subproofs, but this introduces additional overhead. Hence, embedding the watermark with Equation (1) yields a more efficient zero-knowledge watermark detection protocol.

3.2 How Verifier Determines $\hat{W^*}^{[m]}$

The original heuristic in Cox's watermarking scheme (see Section 2.1) requires to select the coefficients of $\hat{W}$ with the m-highest magnitudes to construct $\hat{W^*}^{[m]}$

[3] We invert to $\Delta_i := \alpha \cdot WM^*{}_i$ instead of $WM^*{}_i$, because with Δ_i in the detection inequality the construction of an efficient protocol is easier to achieve, cf. Sec. 3.3.

[4] This can be achieved by proving the multiplicative relation $\texttt{PoK}_{mult}(C_z; C_{\left(\hat{W}_i^{[m]}\right)^{-1}}, C_{\hat{W}_i^{[m]}})$ and that z is close enough to 1. The latter can be proven by an interval proof [7] that $z \in [1-\delta, 1+\delta]$ for a reasonable small δ.

and $\hat{W}^{[m]}$. In the context of zero-knowledge watermark detection, this heuristic cannot be done in a straightforward way, since $\mathcal{V}$ only knows the committed version $C_{\hat{W}^{[m]}}$ of the original transformed image $\hat{W}$. We describe two (among several other) viable solutions to overcome this problem:

Solution 1: This method provides a generic solution which is applicable to *every* correlation-based watermarking scheme whose detection criterion can be expressed as a polynomial. The general idea of this solution is that WM is chosen as large as the image size (e.g., $m = N \cdot N$) and that all positions i, not supposed to be marked, are set to value $WM_i := 0$. In this case no selection for $\hat{W}^{*[m]}$ and $\hat{W}^{[m]}$ is required at all and *corr* remains unaffected as well. Unfortunately, this general approach involves a significant overhead, as the number m of coefficients that have to be processed becomes quite large.

Solution 2: Here we consider the *special case* where the embedding positions are public parameters of the watermarking scheme and, therefore, can be given as common input to both parties, thus yielding more efficient detection protocols. One possibility to obtain these fixed embedding positions has been proposed by Piva et al [11] and works as follows: Embed WM along a zig-zag scan of the AC coefficients similar to the walk in the JPEG compression algorithm (but on the entire $N \times N$ DCT-transformed image). Embedding of the watermark begins at a predetermined diagonal l, which becomes part of the common input. l is chosen such that a sufficient number of low-frequency AC coefficients is used for embedding. This methodology matches the required choice of significant coefficients in $\hat{W}$ for embedding WM, since for most images the low-frequency coefficients mainly coincide with the highest magnitude coefficients of $\hat{W}$. The result is a compatible efficient zero-knowledge version of Cox's watermarking scheme.

3.3 Adaption of the Detection Inequality

We have to transform the detection criterion $corr \geq S$ respectively $corr - S \geq 0$ such that the computation of *corr* can be expressed as a polynomial term. Inserting Equation (3) into $corr - S \geq 0$ leads to $\sum_{i=0}^{m-1} WM_i \cdot WM^*{}_i - S \cdot \sqrt{\sum_{i=0}^{m-1}(WM^*{}_i)^2} \geq 0$. The detection threshold S is chosen as $S \geq 0$ [3]. If WM is present, then $\sum_{i=0}^{m-1} WM_i \cdot WM^*{}_i \geq 0$ also holds. In this case we are allowed to square the inequality in order to eliminate the root term which would require additional zero-knowledge subproofs. Otherwise, we are already assured in this stage that WM is not present and can omit further computations (cf. Section 3.4).

Now the resulting term has a polynomial form which allows us to apply the zero-knowledge protocol primitives. A multiplication with α^2 allows us to use $\Delta_i := \alpha \cdot WM_i$ directly from Equation (4) or (5) which leads to the detection criterion

$$\underbrace{\left(\sum_{i=0}^{m-1} WM_i \cdot \Delta_i\right)^2}_{=:A} - \underbrace{S^2 \cdot \sum_{i=0}^{m-1} \Delta_i^2}_{=:B} \geq 0. \tag{6}$$

An intermediate computation of C_{Δ_i}, C_{A^2} and C_B and a proof that $A^2 - B \geq 0$ in C_{A^2-B} convinces a verifier that a given committed watermark C_{WM} is present in W^*.

Certainly, the entire protocol becomes less sophisticated if one assumes a detection criterion $S \geq WM \cdot WM^*$ without any denominator. However, in the zero-knowledge setting, one cannot simply multiply Equation (3) by $\|WM^*\|$ because this value is obtained from $\hat{W}^{[m]}$, which is cryptographically concealed from the verifier by $C_{\hat{W}^{[m]}}$. Making it public as a new detection threshold $S \cdot \|WM^*\|$ would leak knowledge about WM^* and, hence, about $\hat{W}^{[m]}$.

3.4 The Entire Detection Protocol

The common input to the protocol (a graphical illustration can be found in [12]) consists of the commitments C_{WM}, $C_{\hat{W}^{[m]}}$, the commitment description $descr_{com}$, W^*, the watermark position l and the detection threshold S. Furthermore, $\mathcal{P}$ knows the plain-text version of WM and $\hat{W}^{[m]}$ as well as the corresponding secret opening information of the commitments.

First, $\mathcal{P}$ and $\mathcal{V}$ compute $\hat{W^*}^{[m]}$ according to the JPEG-analog zig-zag heuristic, starting at diagonal l. In several stages, $\mathcal{P}$ and $\mathcal{V}$ interactively compute the required committed intermediate results C_{Δ}, C_{A^2} and C_B. Finally, $\mathcal{P}$ proves to $\mathcal{V}$ that the detection equation (6) is satisfied.

$\mathcal{V}$ computes all m components C_{Δ_i} of C_{Δ} homomorphically as $C_{\Delta_i} := g^{\hat{W^*}_i^{[m]}} \cdot \left(C_{\hat{W}_i^{[m]}}\right)^{-1}$. The committed addends for C_A, i.e., $C_{WM_i \cdot \Delta_i}$, have to be provided by $\mathcal{P}$ and $\mathcal{P}$ initiates m subproofs $\mathtt{PoK}_{mult}()$ to convince $\mathcal{V}$ that the products contained in $C_{WM_i \cdot \Delta_i}$ are correct. Afterwards, $\mathcal{V}$ can compute C_A homomorphically on his own as $C_A := \prod_{i=0}^{m-1} C_{WM_i \cdot \Delta_i}$. Before the squaring step, $\mathcal{P}$ has to prove that A contained in C_A is greater or equal to zero. Otherwise, this would imply that *corr* in Equation (3) is < 0 and $\mathcal{V}$ would be assured already in this stage of the protocol that WM is not present in W^* and aborts the protocol. Finally, $\mathcal{P}$ generates C_{A^2}, sends it to $\mathcal{V}$ and proves in zero-knowledge that C_{A^2} indeed contains the square of the value A contained in C_A.

In the next protocol section, value B of Equation (6) is determined: $\mathcal{P}$ provides $C_{\Delta_i^2}$ and proves that $C_{\Delta_i^2}$ indeed contains the square of the value Δ_i contained in C_{Δ_i}. Then $\mathcal{V}$ can compute C_B and C_{A^2-B} by making use of the commitment scheme's homomorphic property. The watermark detection protocol is finished by a proof that the value $A^2 - B$, contained in C_{A^2-B}, is greater or equal to 0.

Completeness of the protocol follows from the completeness of all sub-protocols and the homomorphic property of the commitment scheme. The *soundness* of the entire protocol holds, because $\mathcal{P}$ would either have to break the soundness of at least one sub-protocol or the binding property of the commitment scheme. As both is assumed to be computationally infeasible, soundness of the overall protocol follows. The *zero-knowledge* property follows from the zero-knowledge property of the sub-protocols and from the fact that additional communication consists of commitments, which are statistically hiding.

4 Implementation Results

Theoretical Bounds: We discuss the communication complexity for the sequential composition of non-interactive elementary sub-protocols. The generation of $\hat{W}^{*[m]}$ will be neglected in computation complexity as it is not part of the very zero-knowledge watermark detection protocol. All in all, in addition to the protocol communication of the sub-proofs, $\mathcal{P}$ transfers $2m+1$ commitments (namely $C_{WM_i \cdot \Delta_i}$, C_{A^2} and $C_{\Delta_i^2}$; $i = 0, \ldots, m-1$), which corresponds to approximately $(2m+1) \cdot |n|$ bits of traffic. $\mathsf{Comm}_{(\mathcal{P},\mathcal{V})}^{WMCox} = m \cdot \mathsf{Comm}_{(\mathcal{P},\mathcal{V})}^{\mathsf{PoK}_{mult}()} + 2 \cdot \mathsf{Comm}_{(\mathcal{P},\mathcal{V})}^{\mathsf{PoK}_{\geq 0}()} + (m+1) \cdot \mathsf{Comm}_{(\mathcal{P},\mathcal{V})}^{\mathsf{PoK}_{sq}()} + (2m+1) \cdot |n| = (2m+17)|n| + (10m+44)|F| + (5m+20)|T| + (13m+55)k + (5m+22)B + 8m + 33$.

Next we consider $\mathcal{V}$'s computation complexity: The homomorphic computations which provide the intermediate committed results require the following operations: The computation of $C_{\Delta_i} : m \cdot E + m \cdot I + m \cdot M$, $C_A : (m-1) \cdot M$, $C_B : E + (m-1) \cdot M$, and the computation of $C_{A^2-B} : I + M$ such that we obtain a computation complexity of $(m+1) \cdot E + (m+1) \cdot I + (3m-1) \cdot M$. Together with the sub-protocols, we get $\mathsf{Comp}_{\mathcal{V}}^{WMCox} = m \cdot \mathsf{Comp}_{\mathcal{V}}^{\mathsf{PoK}_{mult}()} + 2 \cdot \mathsf{Comp}_{\mathcal{V}}^{\mathsf{PoK}_{\geq 0}()} + (m+1) \cdot \mathsf{Comp}_{\mathcal{V}}^{\mathsf{PoK}_{sq}()} + (m+1) \cdot E + (m+1) \cdot I + (3m-1) \cdot M = (16m+67) \cdot E + (6m+23) \cdot I + (13m+49) \cdot M$.

$\mathcal{P}$ is able to follow $\mathcal{V}$'s homomorphic operations directly on the secret values and secret keys of the corresponding commitments. Therefore, we obtain a computation complexity of $(4m+4) \cdot E + (4m+5) \cdot M$. Hence, $\mathcal{P}$'s computation complexity including all sub-protocols is: $\mathsf{Comp}_{\mathcal{P}}^{WMCox} = m \cdot \mathsf{Comp}_{\mathcal{P}}^{\mathsf{PoK}_{mult}()} + 2 \cdot \mathsf{Comp}_{\mathcal{P}}^{\mathsf{PoK}_{\geq 0}()} + (m+1) \cdot \mathsf{Comp}_{\mathcal{P}}^{\mathsf{PoK}_{sq}()} + (4m+4) \cdot E + (4m+5) \cdot M = (14m+84) \cdot E + (19m+95) \cdot M + 2 \cdot L$.

This leads to a total computation complexity of $\mathsf{Comp}_{(\mathcal{P},\mathcal{V})}^{WMCox} = (30m+151) \cdot E + (6m+23) \cdot I + (32m+144) \cdot M + 2 \cdot L$.

Practical Results: A prototype implementation was done in JAVA to achieve a proof of concept of the practicability of zero-knowledge watermark detection. Table 2 shows the results for different numbers of coefficients while the security parameters were chosen as follows: $|n| = 1024$, $B = 1024$, $T = 2^{512}$, $|F| = 80$ and $k = 40$. The runtime was measured for a prover and a verifier process, running simultaneously on one Athlon 1200 desktop PC. The estimated lower bound for the communication complexity $\mathsf{Comm}_{(\mathcal{P},\mathcal{V})}$ – without any implementation or network overhead – is obtained by summarising the theoretical results from Table 1 together with the transmission of the supplementary commitments. The last column of Table 2 shows that if the communication traffic exchanged by our implementation is compressed by a zip-packer, we come very close to the expected theoretical bound $\mathsf{Comm}_{(\mathcal{P},\mathcal{V})}$.

Since the same bases g and h are used in all subproofs and intermediate commitments, the use of fixed-base exponentiation algorithms (see Chapter 14 of [13]), achieved a speed up of factor 3 for the modular exponentiations. The precomputation required by these exponentiation algorithms took 4 : 20 minutes

Table 2. time [min:sec] and $\mathsf{Comm}_{(\mathcal{P},\mathcal{V})}$ [Bytes], precomputation time excluded

Coeffs	time	$\mathsf{Comm}_{(\mathcal{P},\mathcal{V})}$	measured $\mathsf{Comm}_{(\mathcal{P},\mathcal{V})}$	zip ($\mathsf{Comm}_{(\mathcal{P},\mathcal{V})}$)	$\frac{\mathsf{Comm}_{(\mathcal{P},\mathcal{V})}}{\mathsf{zip}(\mathsf{Comm}_{(\mathcal{P},\mathcal{V})})}$ in %
100	0:58	152,360	221,614	161,879	5.6
200	1:53	290,560	413,808	303,825	4.4
400	3:42	566,960	801,080	587,493	3.5
800	7:19	1,119,760	1,572,529	1,154,695	3.0
1000	9:09	1,396,160	1,958,554	1,438,377	2.9

and can be done during the setup of the commitment scheme and has to be done only once for all further executions of watermark detection protocol.

5 Conclusion

We presented the entire technical details how to construct a non-interactive zero-knowledge watermark detection protocol for the watermarking scheme by Cox et al [3] chosen as an established correlation-based scheme for many similar derivatives. The obtained results of a prototype implementation state that this secure methodology is indeed applicable in practice and not just a theoretical construction.

References

1. Adelsbach, A., Sadeghi, A.R.: Zero-knowledge watermark detection and proof of ownership. In: Information Hiding, IHW 2001. Volume 2137 of LNCS., Springer, Germany (2001) 273–288
2. Adelsbach, A., Katzenbeisser, S., Sadeghi, A.R.: Watermark detection with zero-knowledge disclosure. ACM Multimedia Systems Journal, Special Issue on Multimedia Security **9** (2003) 266–278
3. Cox, I., Kilian, J., Leighton, T., Shamoon, T.: A secure, robust watermark for multimedia. In: Information Hiding—First International Workshop, IH'96. Volume 1174 of LNCS., Springer Verlag (1996) 175–190
4. Hernandez, J., Perez-Gonzales, F.: Statistical analysis of watermarking schemes for copyright protection of images. In: Proceedings of the IEEE. Volume 87. (1999) 1142–1166
5. Damgård, I.: Commitment schemes and zero-knowledge protocols. In Damgård, I., ed.: Lectures on data security: modern cryptology in theory and practise. Volume 1561 of LNCS. Springer Verlag (1998) 63–86
6. Damgård, I., Fujisaki, E.: A statistically-hiding integer commitment scheme based on groups with hidden order. In: ASIACRYPT. Volume 2501 of LNCS., Springer (2002) 125–142
7. Boudot, F.: Efficient proofs that a committed number lies in an interval. In: Advances in Cryptology – EUROCRYPT '2000. Volume 1807 of LNCS., Springer Verlag (2000) 431–444
8. Lipmaa, H.: On diophantine complexity and statistical zero-knowledge arguments. In: ASIACRYPT. Volume 2894 of LNCS., Springer (2003) 398–415

9. Bellare, M., Rogaway, P.: Random oracles are practical: A paradigm for designing efficient protocols. In: Proceedings of the ACM CCS, ACM Press (1993) 62–73
10. Rabin, M.O., Shallit, J.O.: Randomized Algorithms in Number Theory. Communications on Pure and Applied Mathematics **39** (1986) S 239– S 256
11. Piva, A., Barni, M., Bartolini, F., Cappellini, V.: Dct-based watermark recovering without resorting to the uncorrupted original image. In: Proceedings of ICIP97. Volume I., Santa Barbara, CA, USA, IEEE (1997) 520–523
12. Adelsbach, A., Rohe, M., Sadeghi, A.R.: Full version of this paper. `http://www.prosec.rub.de/publications.html` (2005)
13. Menezes, A.J., van Oorschot, P.C., Vanstone, S.A.: Handbook of Applied Cryptography. CRC Press series on discrete mathematics and its applications. CRC Press (1997) ISBN 0-8493-8523-7.