© 2015 IJSRSET | Volume 1 | Issue 2 | Print ISSN : 2395-1990 | Online ISSN : 2394-4099
Themed Section: Engineering and Technology

Watermark Detection In Compressive Sensing Domain To Preserve Privacy

Saranya S, Sangeetha B, Gayathri M
Computer Science and Engineering, Dhanalakshmi College of Engineering, Chennai, Tamilnadu, India

ABSTRACT

We propose a compressive sensing based privacy preserving watermark detection framework that involves secure multiparty computation and the cloud. There are three parties in the proposed framework, the data holders (DH) of the potentially watermarked images, the watermark owners (WO) and the cloud (CLD). The framework also requires a certificate authority (CA) to issue a Paillier public key pair to the DH and the DH's public key to the WO. For DH (e.g., media agencies), when it collects a large volume of multimedia data from the Internet and stores their encrypted versions in the CLD, it wants to make sure those multimedia can be edited and republished legally. Watermark owners (WOs) are also the content providers who distribute their watermarked content. In some scenarios, not only DH and WO care about the copyright of the multimedia data, certain CLD who offers storage services may also desire to initiate the watermark detection to check if the uploaded multimedia data is copyright protected. For example, a CLD may choose not to provide storage services to copyright protected data illegally owned. If DH would like to use a CLD for storage the encrypted multimedia data from another cloud to this CLD, it will require the CLD to perform watermark detection on the encrypted multimedia data before providing the storage services.

Keywords: Compressive sensing, watermark detection, secure signal processing, secure multiparty computation, privacy preserving.

I. INTRODUCTION

Image watermarking has become very important as the distribution of images in the cloud is increasing. It is preferred instead of hardware and software. Data's are stored and signal processing and data mining are performed under encrypted domain so as to maintain privacy and keep the data secured. As the usage of Internet and social networks is increasing, it is very easy for a third party to collect a large amount of multimedia data from different sources without knowing the copyright information of those data. The users can use the cloud for data storage and also use the copyright owners for watermark detection in order to keep the multimedia data private. The watermark pattern owner wants to keep their watermark patterns private during the watermark detection as well. A legal cloud offering storage services may also desire to participate in watermark detection initiated by the users, or initiate watermark detection itself without the involvement of the users, to check if the uploaded multimedia data is copyright protected. Another advantage of storing the multimedia data and facilitating watermark detection in encrypted version in the cloud is that those encrypted data can be reused if the image data holder needs to work with other watermark owners later for secure watermark detection.

The watermark pattern is embedded for each image so that an untrusted verifier cannot copyright it. This is ensured by some of the traditional secure watermark detection techniques. Asymmetric watermarking and zero-knowledge watermark detection are the two approaches used for secure watermark detection. However, most of the existing secure water-mark detection works assume the watermarked copy are publicly available and focus on the security of the watermark pattern, while the privacy of the target media on which water-mark detection is performed has received little attention. In some of the applications, the multimedia data's privacy is to be supported in the water-mark detection process. The existing secure watermark detection technologies such as zero-knowledge proof protocols ensure performing privacy preserving storage and secure watermark detection

simultaneously. These technologies also transform the multimedia data to a public key encryption domain. But they have limitations, such as complicated algorithms, high computational and communication complexity and large storage consumption in the public key encryption domain. They are theoretically good but their practical implications are difficult.

In our framework, the target image/multimedia data is possessed by the data holder only. A compressive sensing matrix is issued by a certificate authority (CA) server to the image holder. The DCT coefficients of the image data are transformed to a compressive sensing domain by the image holder before it is outsourced to the cloud. The watermark is transformed to the same compressive sensing domain using a secure multiparty computation (MPC) protocol so as to provide secure watermark detection and then sent to the cloud. The cloud has the data in the compressive sensing domain. The cloud cannot reveal the original multimedia data and the watermark pattern without the compressive sensing matrix. The Watermark detection is performed by the cloud in the compressive sensing domain. The image data in the compressive sensing domain can be stored in the cloud and reused for watermark detection from many other watermark owners.

Watermark Detection in the Compressive Sensing Domain

In this section, we first introduce the compressive sensing theory and related works on signal processing in the compressive sensing or random projection domain. We then introduce the statistical correlation based watermarking scheme proposed by Zeng and Liu, based on which we show that correlation based watermark detection in the compressive sensing domain is feasible. A theoretical performance analysis of watermark detection is performed in the compressive sensing domain.

A. Compressive Sensing

Restricted Isometry Property (RIP) is a required condition for the perfect reconstruction in the compressive setheory. Before presenting the compressive sensing theory, we first introduce the Restricted Isometry Property:

Restricted Isometry Property (RIP): A vector $x \in R^n$ is k-sparse if $|\{ j : |x_j| > 0\}| \leq k$. A matrix $\Phi \in R^{m \times n}$ is said to have the Restricted Isometry Property of order k and level $\delta \in (0, 1)$ (equivalently, (k, δ)-RIP) if

$$(1 - \delta) \_x \__2^2 \leq \_\Phi x \__2^2 \leq (1 + \delta) \_x \__2^2 \qquad (1)$$

for all k-sparse $x \in R^n$. The restricted isometry constant δ_k is defined as the smallest value of δ for which the above inequality holds.

Most of the literature of compressive sensing has focused on improving the speed and accuracy of compressive sensing reconstruction. Davenport et al take some initial steps towards a more general framework called compressive signal processing (CSP), which shows fundamental signal processing problems such as detection, classification, estimation, and filtering can be solved in the compressive sensing domain. Hsu et al use compressive sensing to learn to predict compressed label vectors and then reconstruct the learned compressed label vectors. It provides mathematical proof and experimental results that show prediction of sparse vec-tors could be done in the compressive sensing domain. Calderbank et al give some theoretical results and show that compressed learning, learning directly in the compressed domain, is possible. It gives the tight bounds demonstrating that the linear kernel SVM's classifier in the measurement domain, with high probability, has an accuracy close to the accuracy of the best linear threshold classifier in the original data domain.

B. Watermark Detection in the Compressive Sensing Domain

A watermark of an image is a pseudorandom bit sequence added to an image. Watermark is known only to the owner of that image. Since it is known to the watermark owners, others cannot copyright it or remove it. Therefore it can serve as proof that the owner really owns this image. There should be no perceptual degradation of the image, and the watermark should be detectable even after manipulation of the image.

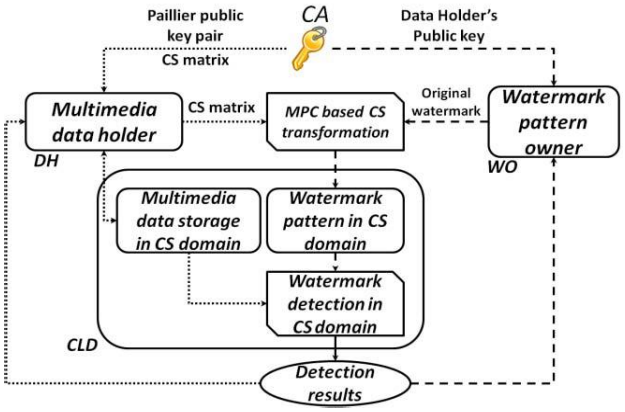

Figure 1: Architecture of the proposed framework

There are two main models for digital image watermarks:

1. The owner of the image stores some information about every image that he publishes. This information is needed for verification of the watermark bit sequence.
2. There is no extra information available except the watermark information added to the image. Obviously, the second model is more difficult to realize, since the user of the image can always try to destroy the complete watermark information by manipulating the image. In this paper, we use a slightly modified approach which needs very little extra information. This information can be shared between different images. The owner of the copyright has to remember a password used in the watermarking algorithm to check whether a watermark is embedded inside an image or not.

II. METHODS AND MATERIAL

The Proposed Framework

We first provide an overview of the proposed framework in Section III.A. The framework relies on a secure CS transformation protocol which is introduced in Section III.B. The complexity and security analysis of the framework is given in Section III. C.

A. The Framework

There are three parties in the proposed framework, the data holders (DH) of the potentially watermarked images, the watermark owners (WO) and the cloud (CLD) as illustrated in Fig. 1. The framework also requires a certificate authority (CA) to issue the public keys and CS matrix keys to certain parties of the framework. For DH (e.g., media agencies), when it collects a large volume of multimedia data from the Internet and stores their encrypted versions in the CLD, it wants to make sure those multimedia can be edited and republished legally. Watermark owners (WOs) are also the content providers who distribute their watermarked content (the watermark embedding is performed by WO before the contents are published). WOs always want to know if their contents are legally used and republished

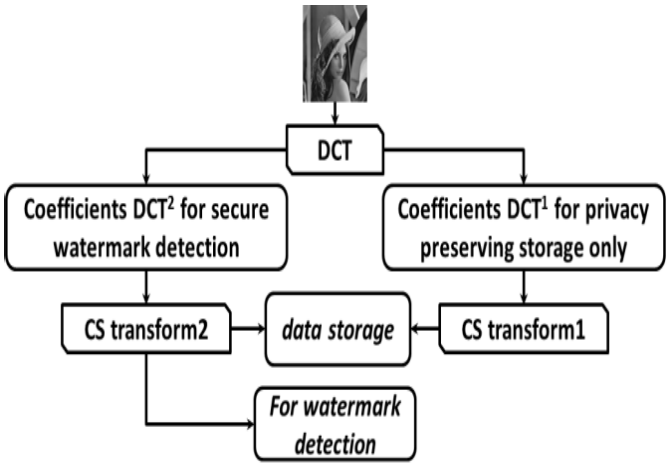

Figure 2: Some DCT coefficients (DCT1) are only used for storage;The other coefficients (DCT2) serve for both the storage and watermark detection purposes.

In our framework, initially, the CA needs to issue CS matrix suites to the DH. The CS matrix suites include the seeds and the random function used to generate the Gaussian CS matrix. We use the CA to issue the random function to guarantee the randomness of the generated Gaussian CS matrix. The CA also needs to issue a Paillier public key pair to the DH and the DH's public key to the WO.

The Discrete Cosine Transform (DCT) is an image transformation which is often used for compression of image. The practical importance of the DCT comes from the fact that a pixel matrix P with N2 values that may all be large is transformed into a result matrix H with only one large element in the upper left corner (which stores essentially all the "image energy"). All other matrix elements are close to zero. One further important aspect of DCT is the fact that the small values of H can be slightly changed without visible effect after using the inverse DC transformation. Therefore we can use all the matrix elements except the element in the upper left

corner for insertion of watermark information.

Watermarking using the RGB color model often inserts data into the least significant bit of each color component. This technique is appropriate for the RGB color representation since the least significant bit does not influence the "visible color" significantly, such that the human eye usually does not see any difference between the original and the watermarked image.

B. Secure CS Transformation Protocol

The CS transformation essentially is a scalar product between vectors, our secure CS transformation protocol is constructed from secure scalar product protocol.

1. Secure Scalar Product Protocol

Homomorphism based, commodity server based, secret sharing based techniques are the existing protocols. In homomorphism based techniques, there are three parties involved. The first two parties are for computation process and the third party is for final results.

2. Secure CS Transformation Protocol:

Based on Protocol 1, it is straightforward to give the secure CS transformation protocol.

3. Handling Real Values Through Scaling:

The Pailler public key system only takes positive integers as input, while our framework involves real number values. The floating point values are scaled into integer values by using scaling factor. Negative integers are represented by the upper half of the range [0, N − 1] (N is the modulo) in a modulo field, e.g., −1 is represented as N − 1.

C. Analysis

1. Complexity Analysis

The CLD has the image and the watermark pattern in the CS domain, so watermark detection in the CS domain only involves linear correlation, hence only introducing negligible computational overhead. The computational and the communication complexity of Protocol 2 are based on Protocol 1. When the watermark size is n and the CS matrix size is m*n, WO performs m*n exponentiations and m encryptions in the public key domain, and DH performs m*m encryptions and m decryptions in the public key domain. For communication complexity, DH sends WO m*n public key encrypted values and WO sends DH m public key encrypted values.

2. Security Analysis

It has been proven in the original paper that Goethals's secure scalar protocol is secure under the semi-honest model. It is straightforward to see that the MPC protocols (Protocol 1&2) are also secure under the semi-honest assumption that all the parties follow the protocol strictly and no two parties will collude to attack a third party. After running the secure CS transformation protocol, DH and WO do not leak their private values to other parties. Only the CLD has the image data and watermark pattern in the CS domain.

The security of using compressive sensing transformation as an encryption has been explored in [20] and it was concluded that it is computationally secure under the brute force and structured attacks when each CS matrix is used only one time. So if the data holder encrypts different images with different CS matrix keys, the CS domain data are secure in the cloud.

3. Comparison to Previous Works and Complexity Evaluation

When compared to previous works, our framework has the following advantages:

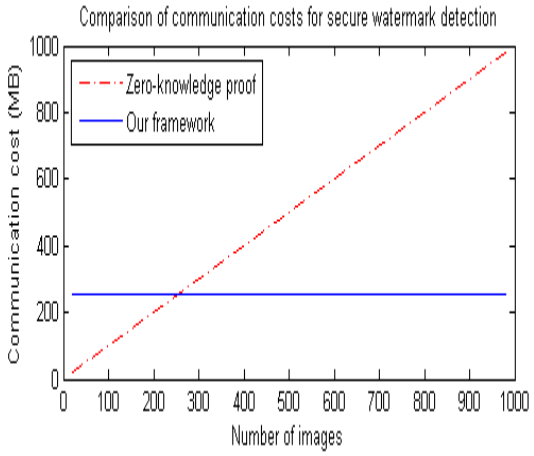

Figure 3: Comparison of the communication costs between our framework and the Zero-knowledge proof protocol [6] when the total number of target images increases.

1. Our framework utilizes the computing and storage resource of the cloud simultaneously and provides better efficiency and flexibility as the encrypted image data (and the encrypted watermark pattern under some circumstances, if so chosen) can be reused for multiple watermark detections in the cloud.
2. Most of the existing secure watermark detection works paid little attention to the privacy of the multimedia data, while our framework protects the privacy of the self-collected data.

III. RESULTS AND DISCUSSION

A. Experimental Settings and Notations

We tested the proposed system using some standard 512×512 images. For the watermark detection, there is several detection methods proposed in [8]. We choose the one in which the watermark pattern used for watermark detection is directly generated from a Normal distribution N(0, 1).Given a CS matrix Φm×n, m/n will be referred to as the compressive sensing rate (CS rate). Since the CS matrix size will be extremely large if we convert the 512×512 image to a vector for CS transformation. Instead, we cut the image into pieces and each piece contains 64 8×8 DCT blocks. Selective DCT coefficients of each piece will form a vector and be transformed to a CS domain with the same CS rate but using different CS matrixes. The data in the CS domain from all pieces is treated as { pi }. Similarly, we get {ri} from the 512×512 original watermark pattern. We test the watermark detection performance when different numbers of DCT components are transformed to the CS domain as DCT2 in Fig. 2. In the rest of this section, "Top AC 20" means top 20 AC coefficients in the zigzag order are selected as DCT2.

B. Scaling Floating Point to Integer Error Analysis

Since the MPC protocol is based on the Paillier public key system which requires integers as input, we scale the floating point values to integers with certain scaling factors. We test the error introduced by the conversion by comparing the result from secure CS transformation protocol to CS transformation with the original CS matrix and the watermark pattern. As shown in Table I, the MSE decreases significantly as the scaling factor increases. In the following experiments, the scaling factor is set to 1.0e8.

C. Secure Watermark Detection in the Compressive Sensing Domain

1. Assertions Validation

We can see that the covariance is very small and close to zero. However, it is interesting to see that under H1, the covariance term is concentrated around a very small negative value.

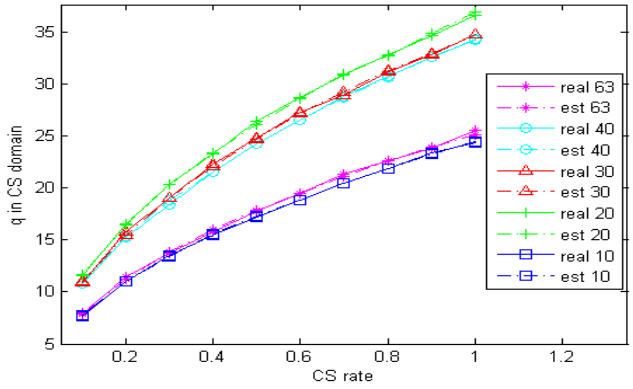

Figure 4: Watermark detection in the CS domain under different CS rates when different DCT components are selected.

2. Watermark Detection in the CS Domain

We evaluate the watermark detection performance in the CS domain when both the watermark signals and certain noises are transformed to the CS domain simultaneously.Fig.6 shows the watermark detection performance in the CS domain when Gaussian noise (generated by the Gaussian random value generator in Mat lab) is inserted into the test image. The figure shows that the watermark detection performance decreases only slightly even when the zero-mean Gaussian noise has a standard deviation of 40.

D. Compressive Sensing Encryption

A different CS matrix is used for the CS reconstruction, the reconstructed image is totally random. The block effect is due to the inverse-DCT operation on the 8×8 DCT block. If watch closely, it can be observed that still preserves the spatial contour roughly. The reason is that the CS transformation in our experiment is performed piece-wisely as mentioned in Section IV.

E. Compressive Sensing Reconstruction

For privacy preserving storage, since the DCT coefficients are not perfectly sparse, the CS reconstruction will introduce distortion to the reconstructed image, especially when CS rate is low. The CS reconstruction error has been studied in many other works. Here we present our CS reconstruction experimental results when all AC components are transformed to a CS domain. In order to have a good quality image after the CS reconstruction, the CS rate needs to be high.

IV. CONCLUSION

This paper brings out the idea that watermark detection and privacy preserving can be done simultaneously. The semi-honest adversary model is used to protect the private data. For example, collusion between WO and CLD will cause the leakage of DH's CS matrix. More theoretical analysis of the covariance term will be conducted in the future work. In addition to watermark detection, our framework can also be extended for other secure signal processing algorithms. Future work also includes further evaluation of the robustness of the watermark detection in the CS domain under some other attacks. In addition to secure CS transformation, developing MPC protocols for secure CS reconstruction is part of our future work too.

V. REFERENCES

[1] T. Bianchi and A. Piva, "Secure watermarking for multimedia content protection: A review of its benefits and open issues," IEEE Signal Process. Mag., vol. 30, no. 2, pp. 87–96, Mar. 2013.

[2] Z. Erkin, A. Piva, S. Katzenbeisser, R. Lagendijk, J. Shokrollhi, G. Neven, et al., "Protection and retrieval of encrypted multimedia content: When cryptography meets signal processing," EURASIP J. Inf. Security, vol. 7, no. 2, pp. 1–20, 2007.

[3] J. Eggers, J. Su, and B. Girod, "Public key watermarking by eigenvectors of linear transforms," in Proc. Euro. Signal Process. Conf., 2000.

[4] S. Craver and S. Katzenbeisser, "Security analysis of public-key water-marking schemes," in Proc. Math. Data/Image Coding, Compress., Encryption IV, Appl., vol. 4475. 2001, pp. 172–182.

[5] A. Adelsbach and A. Sadeghi, "Zero-knowledge watermark detection and proof of ownership," in Proc. 4th Int. Workshop Inf. Hiding, vol. 2137. 2001, pp. 273–288.