The study of various attacks on Digital watermarking technique

Mr.Mitesh Patel
Computer Department L.D.engg College Ahmedabad Gujarat India,
Gujarat Technological University

Ms swati Asst. professor L.D.C.E.

Mr.Alpesh Chauhan Asst.Professor KJIT-savli

Abstract— **Digital watermarking Technique is used to prevent Illegal copy and duplication in digital media. Here in this paper I present study of various attacks on Digital watermarking**

Index Terms— **attacks, Digital Watermarking,**

I. Introduction

The Internet is an excellent sales and distribution channel for digital assets, but copyright compliance and content management can be a challenge. These days, digital images can be used everywhere – with or without consent. Images that are leaked or misused can hurt marketing efforts, brand image and, ultimately, sales. With one click, your digital assets can be detached from your copyright information, so guarding brand and intellectual property assets is essential. Watermarking solutions let you add an extra layer of protection to your digital images.

Copyright Protection: Embed copyright, owner ID and other digital information into digital images, telling who owns it and how it can be used.

Robustness means that the watermark is able to withstand with some changes in the watermark-embedded signal; while imperceptibility represents the invisibility to human eyes, or for audio clips, the inaudibility to human ears. A good watermark algorithm should be by all means is simultaneously robust and imperceptible.

II Working of Digital water marking system

When combined, digital watermarking products and services form a complete copyright communication and image tracking system for digital images.

This system provides the tools and capabilities to:

- Embed digital watermarks into images
- Detect and read digital watermarks
- Link to complete contact details or a web site for the image creator or distributor (for inquiring about usage rights, licensing, etc.)
- Track instances of digitally watermarked images on the web. With the growth of numerical technologies, it became extremely easy to reproduce a data without any damage. For instance, any image taken on the Internet can be saved for a Personal usage and then be written on a CD-Rom or on another web page.

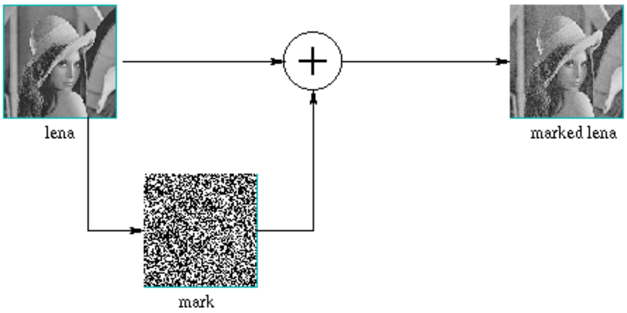

Fig1:Independent watermarking

III Attacks on watermarking system.

One categorization of the wide class of existing attacks contains four classes of Attacks:

- Removal attack
- Geometric attack
- Cryptography attack
- Protocol attack

REMOVAL ATTACKS

Removal Attacks aim at the complete removal of the watermark information from the watermarked data without cracking the security of the watermarking algorithm (e.g., without the key used for watermark embedding). That is, no processing, even prohibitively complex, can recover the watermark information from the attacked data. This category includes de noising, quantization (e.g., for compression), re modulation, and collusion attacks. Not all of these methods always come close to their goal of complete watermark removal, but they may nevertheless damage the watermark information significantly. Sophisticated removal attacks try to optimize operations like de noising or quantization to impair the embedded watermark as much as possible while keeping the quality of the attacked document high enough. Usually, statistical models for the watermark and the original data are exploited within the optimization process.

Collusion attacks are applicable when many copies of a given data set, each signed with key or different watermark, can be obtained by an attacker or a group of attackers. In such a case, a successful attack can be achieved by averaging all copies or taking only small parts from each different copy

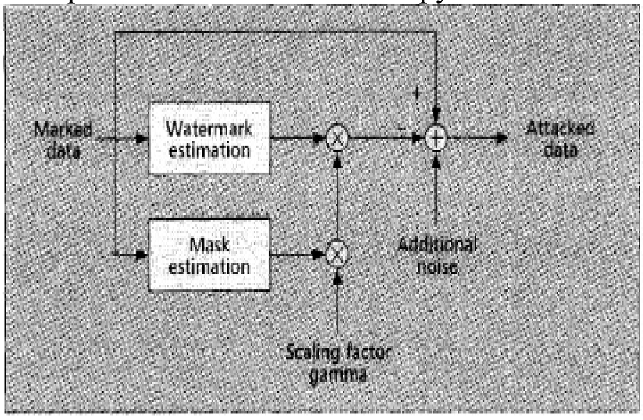

Fig2: re modulation attacks

GEOMETRIC ATTACKS

In contrast to removal attacks, geometric attacks do not actually remove the embedded watermark itself, but intend to distort the watermark detector synchronization with the embedded information. The detector could recover the embedded watermark information when perfect synchronization is regained. However, the complexity of the required synchronization process might be too great to be practical. However, most recent watermarking methods survive these attacks due to the use of special synchronization techniques. Robustness to global geometric distortions often relies on the use of either a transform-invariant domain (Fourier-Melline) or an additional template, or specially designed periodic watermarks whose auto-covariance function (ACF) allows estimation of the geometric distortions. However, as discussed below, the attacker can design dedicated attacks exploiting knowledge of the synchronization scheme. Robustness to global affine transformations is more or less a solved issue. Therefore, pixels are locally shifted, scaled, and rotated without significant visual distortion. However, it is worth noting that some recent methods are able to resist this attack.

CRYPTOGRAPHIC ATTACKS

Cryptographic attacks aim at cracking the security methods in watermarking schemes and thus finding a way to remove the embedded watermark information or to embed misleading watermarks. One such technique is brute-force search for the embedded secret information.

Another attack in this category is the so-called Oracle attack, which can be used to create a non-watermarked signal when a watermark detector device is available. Practically, application of these attacks is restricted due to their high computational complexity.

PROTOCOL ATTACKS

Protocol attacks aim at attacking the entire concept of the watermarking application. One type of protocol attack is based on the concept of invertible watermarks [7]

.

The idea behind inversion is that the attacker subtracts his own watermark from the watermarked data and claims to be the owner of the watermarked data. This can create ambiguity with respect to the true ownership of the data. It has been shown that for copyright protection applications, watermarks need to be noninvertible. The requirement of non-inevitability of the watermarking technology implies that it should not be possible to extract a watermark from a non-watermarked document.

A solution to this problem might be to make watermarks signal-dependent by using one-way Functions. Another protocol attack is the copy attack. In this case, the goal is not to destroy the watermark or impair its detection, but to estimate a watermark from watermarked data and copy it to some other data, called target data [8]. The estimated watermark is adapted to the local features of the target data to satisfy its imperceptibility. The copy attack is applicable when a valid watermark in the target data can be produced with neither algorithmic knowledge of the watermarking technology nor knowledge of the watermarking key. Again, signal-dependent watermarks might be resistant to the copy attack.

ESTIMATION- BASED ATTACKS

Here, we consider attacks that take into account the knowledge of watermarking technology and exploit statistics of the original data and watermark signal [5, 8-32]. In addition, we emphasize that for the design of attacks against watermarking schemes, the distortion of the attacked document and the success of watermark impairment has to he considered. Within the scope of these attacks, we present the concept of estimation-based attacks

.

This concept is based on the assumption that the original data or the watermark can be estimated - at least partially - from the watermarked data using some prior knowledge of the signals' statistics.

Note that estimation does not require any knowledge of the key used for watermark embedding. Furthermore, knowledge of the embedding rule is not required, but the attack can be more successful with it. Depending on the final purpose of the attack, the attacker can obtain an estimate of the original data or of the watermark based on some stochastic criteria such as maximum likelihood (ML), maximum a posteriori probability (MAP), or minimum mean square error (MMSE). We do not focus here on the particularities of the above estimation but rather concentrate on different ways to exploit the obtained estimates to impair the embedded watermark. Depending on the way the estimate is used, we can classify estimation-based attacks as removal, protocol, or desynchronization attacks

REMODULATION ATTACKS

Remodulation attacks aim at modification of the watermark using modulation opposite to that used for watermark embedding. Assuming the estimated watermark is correlated with the actual watermark, meaning a good estimate could be obtained, the estimated watermark can be subtracted from the watermarked data. Subtracting a very inaccurate estimate of the watermark might decrease the document quality without affecting the watermark too much. On the other hand, correlation-based detection can be defeated by subtracting an amplified version of the estimated watermark. For this reason, we introduced a gain factory 21, which gives us the possibility to trade off the distortion of the attacked document vs. the success of the attack. There are four basic variations of the remodulation attack. First, when y = 1, the attack yields the MMSE estimate of the original and reduces to the denoising attack.Second, for y > 1, the quality of the attacked document might be reduced, but correlation based detection might be defeated more successfully. The attack can even drive the correlation to zero so that the detector incorrectly decides that the watermark is not present in the attacked data. Third, when using a more sophisticated distortion measure than simple
MSE, a better compromise between success of the attack and introduced distortion can be
Obtained by weighting the re-modulated watermark by a perceptual mask. Fourth, the attacker can not only subtract the weighted, estimated watermark, but also add outliers to obtain a non-Gaussian noise distribution, which decreases the performance of correlation-based detection. Moreover, exploiting features of the human perceptual system, the attacker can efficiently embed a large amount of outliers in perceptually less significant parts of the data. For image data, this approach has been demonstrated to be successful in [9]. We refer to this attack as Perceptual re-modulation.

COPY ATTACK

The estimated watermark can be exploited to implement a copy attack, as already described. Of course, the copied watermark has to be adapted to the target data to keep the quality of the falsely watermarked target data high enough. There are many practical ways to adapt the watermark to the target data based on perceptual models. For images, contrast sensitivity and texture masking phenomena of the HVS can be exploited. The estimation-based copy attack is most successful when the same perceptual model is used as in the original watermarking algorithm. Note that the copy attack in its described version is mainly applicable to additive watermarking schemes. In the case of quantization-based watermarking schemes, even a perfectly estimated watermark signal w cannot be copied since it is highly unlikely that the copied signal w is a valid watermark in the target signal

OPTIMIZED ATTACKS

So far, we have developed different attacks based on watermark estimation and discussed how an embedder can make watermark estimation as difficult as possible. However, the embedder has another chance to react to estimation based attacks, especially to the re-modulation attack. The detector can first estimate the re-modulated watermark and try to invert the re-modulation, and thus retain reliable watermark detection. For instance, when watermark estimation is based on Wiener filtering, the detector can apply inverse Wiener filtering. This is of course not wanted by the attacker. Thus, he/she has to add noise to the attacked data, which would be amplified by the inverse Wiener filtering and thus impair watermark detection. Of course, the additive noise further degrades the attacked signal. Now, we have the situation where the attacker has to find a good combination of using the estimated watermark and the noise to be added.

The embedder no longer simply tries to make watermark estimation as hard as possible, however; he .also has to get a power advantage over the additive noise an attacker might introduce. This problem can be formulated in an even more general way: the attacker attempts to minimize the watermark capacity under a constraint on the distortions introduced by the attack. The embedder attempts to maximize the watermark capacity under a constraint on the embedding distortions. This situation can be regarded as a game between the attacker and embedder. To solve this problem, we followed a game-theoretic approach, assuming that the embedder and attacker know each other's behavior.

CONCLUSION

The idea behind this paper is to provide basic information about Digital watermarking technique and possible attack on it. So any beginner can start their work on it.

REFERENCES

[1] M. Kutter and F Petitcolas, "A Fair Benchmark for Image Watermarking Systems," Electronic Imaging '99:Security and Watermarking
of Multimedia Content, SPIE Proc., vol.
3657, San Jose, CA, Jan. 1999.

[2] I. Cox et al., "Secure Spread Spectrum Watermarking for Multimedia," /E€€ Trans.
Image. Proc., vol. 6, no 12,Dec. 1997,
pp. 1673-87.

[3] B. Chen and G. W. Wornell, "Dither Modulation: A New Approach to Digital
Watermarking and Information Embedding," Security and Watermarking of Multimedia
Contents, Proc SPIE. vol 3657, San Jose, CA, Jan. 1999.

[4] J. J. Eggers, J.K. Su, and B. Girod, "A Blind Watermarking Scheme Based on Structured
Codebooks," Colloq: Secure Images and Image Authentication. London, UK, Apr. 2000.

[5] S.Voloshynovskiy et al.. "Attack Modeling: Towards a Second Generation Watermarking Benchmark," Sig. Processing. Special
Issue on Information Theoretic
Issues in Digital Watermarking, 2001,
vol. 81, no. 6, pp. 1177-214.

[6] Sviatolsav Voloshynovskiy, Shelby Pereira, and Thierry Pun, University of Geneva Joachim 1. Eggers and Jonathan K. Su, University of Erlangen - Nuremberg,"Attacks on Digital Watermarks: Estimation-Based Attacks, and Benchmarks", 0163-6804/01/0 2001 IEEE Communications Magazine August ZOO1

Web references

1]http://academic.mu.edu/phys/matthysd/web226/L0205.htm
2]http://www.profc.udec.cl/~gabriel/tutoriales/rsnote/cp10/cp10-9.htm
3]http://www.cg.tuwien.ac.at/~theussl/DA/node36.html
4]http://www.cs.umsl.edu/~sanjiv/classes/cs5420/lectures/spatial.pdf