Multimedia Security

Seminar Report

Submitted in partial fulfillment of the requirements
for the degree of

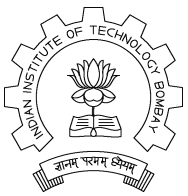
Master of Technology

by

M. Nithya
Roll No: 03305808

under the guidance of

Dr. Sharat Chandran



Department of Computer Science and Engineering
Indian Institute of Technology, Bombay
Mumbai

Acknowledgment

I would like to thank **Dr. Sharat Chandran** for his invaluable support, encouragement and guidance, without which my M.Tech seminar report would have been an exercise in futility. As a token of my esteem and gratitude, I honour him for his assistance towards this cause.

M. Nithya

Abstract

In the modern world, various media types such as text, audio, image and video have managed to enter the network arena. Reasons for this is quicker sharing and acquiring the data in a quicker period of time and this is amply supported by the considerable cheap availability of resources like high bandwidth. This step has raised many challenges opposing the former's success. One of the most challenging issues is the lack of security of the data.

Many techniques in digital watermarking were developed as and when required, to counter this challenge, but one single approach that would solve the security problems for all the media types has failed to exist till now. I have suggested some methods which meets the above mentioned demand. They are spread spectrum watermarking, non-repudiation oblivious watermarking, and attack characterization. The first technique is about embedding the watermark in few vital locations of the document so that watermark cannot be removed imperceptibly, whereas the second one proposes watermarking schemes for data distribution and the last method improves the robustness of the watermark by characterizing the attack using reference watermark.

Contents

1 Introduction

Information has become the source of billion dollar investment and multi billion dollar income. Enforcing proper security measures on the information has therefore become vital. It is following this idea that many security algorithms has mushroomed to the occasion in a considerable quicker period of time.

In the field of Multimedia, there is an increased requirement for security, due to various threats which includes replication of digital data without any information loss, and manipulations of the same without any detection. As the utility and usage of the Internet has grown considerably, multimedia documents are easily copied through un-authorized and illegal channels. It is following this reason that the authors of the work hesitate to publish their work electronically, fearing the threat that it can be pirated easily.

Once the information or the content is acquired by miscreants, it can be easily modified by employing some specific software tools and further, can be claimed by them as their own work.

This is definitely a horrendous act that has subsequently wasted the efforts and hard-work put in by the original author, depriving them from getting due rewards and honour for the same. In order to put things in its place, there should be some mechanism which would uniquely sort out the problem by identifying the document and its owner. The mechanism ought to be convincing enough in providing evidences of a document's ownership and substantiate it by providing credible proof. The mechanism should be able to trace the spots where leaks have occurred and to map out all the modifications that have been done to the document. It should also take into consideration the problem of modification of the data by any of the third party or un-authenticated clients, as it forms the root cause for fake ownership.

1.1 Purpose of the Report

Many new requirements and challenges have been mushroomed. One of chief among them is the security measure. As and when new type of media was available on net, corresponding security systems came to the fore that was specific to a particular media data. Though effective systems are available to ensure security in each type of media data, one common technique that would provide security to all multimedia documents failed to exist. Effective methods which solve this problem and provide evidences of robustness and effectiveness are introduced.

1.2 Threats

Threats are the conditions of possible specific actions that are enforced over the document that makes it counterfeit and illegal, as against the wishes of its owner or creator. The most important threats which ought to be handled to ensure the security of the multimedia system are :

Threat of confidentiality This threat represents the possibilities of accessing the data or document via unauthorized channels. With the growing usage of Internet, the chance of its occurrence is highly likely and is hard to get it dispelled out, unless effectively addressed.

Threat of integrity This is a threat to the content of the document by unauthorized entities, where the resource can be altered without any detection.

Threat of availability This threat highlights the condition where a person, who is not supposed to possess a document, actually has it, by obtaining the same through illicit channels.

1.3 Security Requirements

The threats that were focused on earlier were the cardinal reasons for the non-publishing of the digital data electronically by its authors. To overcome that, one has to provide and satisfy the security requirements, which addresses these threats. Some of the important security requirements that have to be implemented in ensuring an approximate fool-proof security are

Confidentiality This requirement emphasizes the permit of only authorized access to the document or content and prevents the unauthorized access of resources.

Data Integrity It is required by this security requirement that the data be identically maintained from its source to destiny, and has not been accidentally or maliciously modified, altered, or destroyed, and remain unchanged right through out the operations such as transfer, storage, and retrieval.

Data Origin Authenticity When a document is found, the origin of that resource should be traceable.

Entity Authenticity Entities participating in the communication should prove that they are the one they claim to be. This proves that the communication is carried out between the correct entities.

Non-Repudiation It should be possible to detect and prove the rightful ownership of that document. Many authors are worried about distributing their works in fear that it may be copied illegally or represented as another's work. Non-repudiation facilitates the identification of the end users who have copied the document. The rightful ownership of that document should be detected and proven.

2 Security Mechanisms

The convincing techniques that have come to the fore to meet this situation are cryptography and digital watermarking. Security measures are based on modern cryptographic mechanisms as well as on security infrastructures. The way in which the discussed security mechanisms can be applied to multimedia resource is difficult to analyze, due to the complexity of multimedia data. Cryptography has been widely used for access control mechanisms to prevent misuse and theft of the material. Though the theft can be prevented, the ease of with which perfect copies can be made without any information loss facilities unauthorized copying of multimedia documents like music and film in a large-scale. The situation demanded for the need of a better technique that would prevent copyright violation after authentication of the legitimate customer. Digital watermarking flourished for meeting the above stated demands.

2.1 Cryptography

Cryptography is a practice of converting the data that is to be transmitted over the network medium, into an un-recognizable format. This is usually achieved by employing computation of functions ranging in various levels of complexities. Apart from that, many of the modern cryptographic mechanisms rely on few unproven assumptions too. The complexity of the crypt-function, determines the security level and the reliability. The trade-off for the complexity of the function is the time period, i.e. more complex the procedure, more time it takes in encrypting and decrypting the data. Though the resources available in modern system are sufficiently large, it can not resolve the hard computations in an acceptable period of time. Therefore, though solution for achieving thorough security is possible, it is hardly reinforced owing to the reasons stated above.

The cryptographic functions are of two types: one-way functions and trapdoor one-way functions. The Cryptosystems by themselves are further subdivided into private-key and public-key cryptosystems.

The private-key cryptosystems are the ones in which the communicating entities share a common secret key K, which is supposed to be kept secret. The encryptor uses this key to encode the message before distributing it, while the same can be decoded using the same key, available at the receiver end. The length of the key is directly proportional to its chances of not being cracked.

Trapdoor one-way functions finds itself engaged in Public-key cryptosystems, where a key pair consisting of Public key and a Private key is held by each commuting entity. Most of the popular cryptosystems such as the RSA Encryption uses a key-pair for its successful operation. There is something called the trapdoor information, which should be known to both the entities. The encryption is done by one of the entities using the public key. This key need not necessarily be a secret one. The other entity who knows the trapdoor information, can decipher the code using his private secret key. It is analyzed that the computation of Private key from Public key is computationally not viable.

For the media data in particular, most of the time it is required that each bit of data of the original signal be taken as the input for the cryptosystem in calculating the encrypted data as output. But when these data are distributed over the network, due to errors in the channel, the data bits may get modified or garbled, though slightly. But this modification is far more than sufficient in upsetting the cryptosystem. It is therefore essential that such errors do not affect the cryptosystem as such, and for that to happen, the method such as the one explained above, should be avoided though it is efficient.

To overcome all of the above stated challenges and yet to maintain maximum confidentiality, one can use digital watermarking in combination with the cryptosystems, which can be used as a supplementary system to the former. Digital watermarking systems as such are very robust and nearly fool-proof. Yet, new attacks arise periodically that keeps challenging the security and fortitude of the system. So, as a measure of empowering safety to the document, one can use cryptography as an additional layer of security to

digital watermarking, thereby hoping for fool-proof security. Optimal security measures can be adhered to in both these systems, and when put together, can evolve into a robust technique.

2.2 Digital watermarking

2.2.1 Overview

Watermarking is the process of altering a work to embed a message (possibly a secret code or copyright information) imperceptibly. The hidden signal that is embedded in the multimedia data is the watermark and the resulting document after embedding the watermark is the watermarked document.

Watermark embedding is the process of embedding the watermark imperceptibly. Secret key can be used to encrypt the watermark and then embed in the document, so that extraction and manipulation of watermark can be prevented.

Watermark extraction is the process of extracting the embedded watermark from the possibly distorted watermarked document which have undergone some processing (may be common signal processing or intentional manipulations).

2.2.2 Classification

Watermarking can be broadly classified into two categories, fragile and robust watermarking. ***Fragile watermark*** detects even slight changes in the watermarked document addressing reorganization of manipulations. The sensitivity of fragile marks to modification leads to their use in authentication. As the name indicates, Fragile watermark becomes undetectable after even minor modifications of the work in which it is embedded. This cannot withstand most of the modifications, so fragile watermark as such cannot be used in many applications.

Robust watermark is the one which can withstand most of the common signal processing like lossy compression, filtering, and geometric scaling and some intentional manipulation. If the robust watermark is modified, it should introduce some perceptual artifacts to the work which discourages these alterations.

The different types of watermarking are highlighted as under.

Authentication Watermarking Authentication watermarking is a hidden data inserted into the data, in order to detect any alterations. Generally, authentication watermarking scheme goes hand in hand with cryptography theory and techniques, where the authentication code, which is also sometimes called as digital signature, is computed from the entire image. The resulting code is inserted into the image itself, but this may consequently affect and alter the image, thereby invalidating the watermark. As a partial solution to this problem, one considers only the Least Significant Bit of the multimedia data for inserting the authentication code, thereby minimizing the impact of distortion on the image.

Fingerprint Watermarking Fingerprint watermarking is the practice of extracting inherent feature vectors that uniquely identify the content of the document. An imperceptible message is embedded into the original document which slightly modifies the content. This is similar to the conventional technique. But additionally, there is some thing called as fingerprinting, which extracts the original data from the watermark embedded document, by comparing it with previously extracted fingerprints in a database. It is following this technique that a variety of applications including fraud detection, broadcast monitoring, etc. work.

Copy Control Watermarking This is a type of watermarking that is extensively and primarily used in the process of preventing copyright violations. The objective of this technique can be met by incorporating a system that would prevent pirating of original copyrighted data, i.e. stopping the recording process and also keeping at bay the process of playing back the pirated content.

Annotation Watermarking In this technique, a watermark is embedded in a document. The technique is supported by certain detection mechanisms, which gets activated whenever the document is accessed by a user and starts analyzing the type of media, its format, etc. The message in the media is searched for if the mechanism detects any evidences of watermark for that type. Once found, the message is translated following which, subsequent actions can be executed, depending on the requirement.

Integrity Watermarking Integrity watermarking is fragile watermarking technique which is similar to Data Integrity requirement, i.e. it is employed when modifications of the copyrighted content has to be detected, thereby making sure that data is unchanged right from its origin till its destination.

2.3 Security Measures

The security requirements should address these threats, which are met by security measures. Some of the important security measures that have to be implemented to ensure the security are :

Confidentiality Cryptographic mechanisms are used to prevent the unauthorized access of resources. Private-key and public key crypto systems are use to achieve confidentiality. However after decryption, unauthorized access cannot be prevented.

Data Integrity The modification of the information can be detected by using following techniques.

One-way hash function In one-way hashing function, encrypted information can be easily calculated from the given input by applying a hash-function. But in reality, it is virtually not possible to compute the actual information from the encrypted one. Also, the hash function is non-inverseable and hence this combined with other techniques can be used for maintaining data integrity.

Message authentication codes One-way hash function is applied to the message and is appended to the actual message to check for integrity. If the actual message and the decrypted one don't tally, then the alteration of the message can be detected.

Digital Signatures Digital signatures are used to identify the owner. This can be implemented by encrypting the message by employing one-way hash function using sender's private key.

Fragile digital watermarking The embedded message, also known as the watermark can be detected only if it has been altered. This, as such doesn't suit for most of the applications because many a times, the data is modified in intermediate processing.

Robust digital watermarking In this technique, the embedded message is robust to alterations, i.e. even if the original document is altered; the watermark embedded in the document remains as such. If it is altered, then it will affect the actual information also.

Data origin authenticity The origin of the resource can be traceable by using message authentication code, digital signature and watermarking, from which the exact identity of the person, to whom the document belongs, can be identified.

Entity authenticity Challenge-response protocol is used to ensure that the communication is carried out between the correct entities. In this protocol, the time-variant challenge is provided to the participating entity, which proves its identity by giving proper response using the secrete key associated that entity.

Non-repudiation Non-repudiation facilitates the identification of the end users who have copied the document. The rightful ownership of that document can be detected and proven.

3 Challenges

Whenever a technique for creating something new is developed, it is most often paralleled by the creation of another technique that challenges its existence, i.e, a technique for defeating the purpose of creation of the previous technique is developed. This phenomenon can be aptly felt in the field of digital watermarking, where most of the watermarking techniques that were developed and believed to be fool-proof has been hacked. It is rightly said that "Every watermark can be hacked and every watermark will be hacked, if there is sufficient motivation and opportunity". So, the task of addressing the challenges and taking appropriate and sufficient measures in overcoming the possibilities of attacks is important than the creation of the watermarking technique itself.

3.1 Overview of Watermark Attacks

The term attack can be defined in short as "any type of illegal handling of a legal document of any media type". The attack can be of any type and its intensity also varies, but everything

ultimately results in loss for the owner. The main goals of an attack can be either removal of the watermark or making it undetectable; or modify it to a different valid one thereby creating fake copies; or create multiple claims of the document.

The main goal is to make the watermark undetectable.

A successful and efficient watermarked document should not only address the problems of attacks, but also resist various other types of legal operations that will be carried out on the media data. The watermark, for instance, should be able to resist all the image processing operations such as filtering, dithering, cropping, scaling and compression. All theses operations lawfully modify the contents of the document in one way or other; but the watermark's association with the document has to be good enough to resist these operations.

The attacks can be mainly classified under four headings. They are:

3.1.1 Robustness Attack

Robustness attacks tries to diminish or remove the presence of a digital watermark. This can be performed by using sensitivity analysis, which modifies a watermark slightly and then asks the mark detector whether the image is marked or not. The modifications are applied until the watermark is no longer detected. Some of the prominent examples are:

Stirmak Attack The most prominent tool for robustness testing of watermarking is the Stirmak. It applies several unnoticeable modifications to an image, like:

- JPEG Compression
- Geometric Transformations (horizontal flip, rotations, scaling, cropping, deletion of lines or columns)
- Enhancement Techniques (low pass filtering, sharpening, histogram modification, gamma correction)
- Noise Addition
- Printing-Scanning

Inversion Attack Inversion attack is an attack in which a watermarked document D + W (D is the original document and W is the watermark), is altered in such a way that it results in a document D' + W + W' where D'+ W' is D. The attacker will try to insert his own watermark into the already marked document. So, the expected situation is that that the copyright owner's original document will seem to have the watermark of the attacker. It leads to a situation in which, the attacker's original document contains the owner's watermark and the owner's original document contains the attacker's watermark. This type of attack is called inversion attack or dead lock attack.

3.1.2 Presentation Attack

Presentation Attacks is the one in which the samples of a Work are scrambled prior to presentation to a watermark detector and then subsequently descrambled. It modifies the content such that the detector cannot find the watermark anymore. A well known example for this is the

Mosaic attack, a simple way of defeating the detection of a watermark. The attack consists of chopping an image into smaller pieces, which are embedded one after another in a web page. Common web browsers render juxtapose images stuck together as a single image. The mark detector however would regard each of the small images as a separate one and would thus get confused.

3.1.3 Interpretation Attack

Interpretation Attack is one in which an attacker can devise a situation which prevents assertion of ownership. Another idea exploits the fact that many schemes do not provide an intrinsic way of detecting which of two watermarks was added first. If the owner of the document d encodes a watermark w, publishes the marked version d + w and has no other proof of ownership then a pirate who has registered his watermark as w0 can claim that the document is his and that the original unmarked version of it was d + w - w0. Apparently, watermarking and fingerprinting are methods that must be used in the context of a larger system which also allows for time stamping and notarization. An example of this attack is ***Forgery Attack***, which is executed by means of forgery, i.e. based on the study of many existing marked images, approximate idea is gathered each time regarding the watermark, and the knowledge groomed by this rough idea, over a sufficient period of time becomes a handy resource in forging the authentication of the signal.

The attack can be countered by embedding the watermarks, containing image dependent data such as the image features, so that the watermarking policy differs each time on the image signal.

3.1.4 Legal Attack

A legal attack does not attempt to exploit the technical details of the watermarking system. Based on legal precedent, the reputation of the content owner, or some other information, doubt may be established in court as to whether a watermark actually constitutes the proof that its owner claims.

3.2 Krechoff's Law

According to krechoff's law, Watermarking system should not be based on the assumption that attacker does not know the algorithmic details. The lack of knowledge of the secret key should prevent the attacker from extracting, forgery or elimination of the watermark. This Law makes the watermarking system more secure and robust.

3.3 Digital Watermarking Requirements

The cardinal requirements of digital watermarking is to overcome the below stated problems, and these are precisely what is being addressed by my algorithm.

- Marks should not degrade the perceived quality of the work.
- Detecting the presence and value of a mark should require knowledge of a secret.

- Multiple watermarks which are embedded in a single object should not interfere with each other.
- The watermark should survive all attacks that do not degrade the work's perceived quality
- It should be self-authentication and self-recovery watermark, which will authenticate the correct unaltered part and recover the altered part.
- Even when clipped data is given, watermark should be extractable.

The important triple requirements of watermark that play an imperative role are Robustness, Imperceptibility, and Capacity

Robustness indicates the extent to which a watermark can withstand to attacks; Capacity is the size of watermark that can be embedded; and Imperceptibility is the measure that indicates the extent to which watermark can be embedded without introducing any perceptual changes.

Watermarking Algorithm should be designed to yield desired levels of capacity, robustness to distortion, and imperceptibility.

4 Approaches

4.1 Spread-spectrum Watermarking

For the watermark to be robust, it is essential that the distortions and tampering of the marked signal doesn't really effect changes in the embedded watermark. One of the effective methods for achieving this demand is by watermarking in the particular frequency domain of the signal. This method is called as frequency-based method of watermarking and is achieved using spread spectrum techniques.

The signal when traveled through any medium, undergoes few changes. The first and foremost one of them is the due to the effect of noise on the channel. Further, channel codes and standard encryption procedures may slightly modify the contents of the data, leading to partial degradation of the content, though they are generally information lossless techniques. The other sources of partial or complete data destruction are lossy compression, affine transformations, and other common signal distortions like re-sampling, conversion, re-quantization, etc.

If one wish to add resilience to the watermark against any of such attacks or operations, the watermark must be placed in the perceptually significant regions of the signal, i.e. it must not be placed in regions of least importance of the signal, as these are the components which are most likely to get eliminated by the lossy compression algorithms and other geometric processes. Further, it should not be placed in regions of high frequency owing to probable data loss among these components due to affine transformations like affine scaling and cropping. Though such affine transformations results in irreversible data loss in most of the spatial watermarking techniques, its impact is hardly felt in frequency-based schemes.

Common signal distortions that includes D to A and A to D conversion, dithering, recompression, re-sampling, etc. provide little clue in analyzing and getting rid of them, but based on the information available from the original document, the distortions can be roughly eliminated.

When said that the watermark should be placed in the most significant regions of the image or signal, the obvious question that is being asked is that how to achieve it without such operations being noticeable. This is pulled-off by spreading the watermark over large number of frequency bins so that the energy in any one bin is very small and certainly unnoticeable. But the owner or creator will be aware of the location and content of the watermark and so, is possible for him to concentrate these numerous weak signals into a single strong one. Whereas for an intruder, a noise of high amplitude is required to be added to all frequency bins for successfully destroying the watermark and this would ultimately deteriorate the signal itself.

A large measure of security can be attained when the watermark is spread throughout the spectrum of the signal at precisely selected frequency regions and it can be placed in the low energy component of the selected frequency coefficient. This leads to total imperceptibility of the watermark.

When frequency transformations are applied, the perceptual significant regions in the spectrum of the signal are highlighted by the perceptual mask. The watermarks can then be embedded in these regions by modifying the coefficients in that region to a specific magnitude, which is known only to the owner. This capacity of the frequency domain for accommodating additional quantity of data without much impact on the quality of the signal is called as perceptual capacity.

The document D is splitted into a sequence of values $V = v_1, ..., v_n$, into which the watermark $X = x_1, ..., x_n$ is inserted to get the adjusted sequence of values $V' = v'_1, ..., v'_n$. V' is substituted in the place of V in the original document D, to get the watermarked document D'. The Watermark X is scaled and inserted in V to get the altered sequence V'. Thus V' can be computed using any one of these formulae:

$$v'_i = v_i + x_i \tag{1}$$

$$v'_i = v_i(1 + \alpha x_i) \tag{2}$$

$$v'_i = v_i(e^{\alpha x_i}) \tag{3}$$

A single scaling parameter α may not be sufficient for all the values v_i, because different spectral components may have different tolerance to modification. Multiple scaling factors $\alpha_1, ..., \alpha_n$ are used to improve the robustness. The scaling factor α_i should be a relative measure of the alterations possible to the value V_i such that the changes are imperceptible.

It is expected that an attacker will attempt to modify or remove the watermark by intentionally modifying the watermarked document. The modified document D^* contains corrupted watermark X^*. In the corrupted watermark X^*, some of the values could be greatly distorted, these values can be ignored.

$$x_i^* \rightarrow \begin{cases} x_i^*, & \text{if } x_i^* > threshold \\ 0 & \text{Otherwise} \end{cases}$$

Watermark is placed in perceptually most significant component, it is robust to common signal processing and the intentional attack which alters the watermark introduces some perceptual changes, thereby damaging the document thereby providing the required level of security.

4.2 Non-repudiation Oblivious Watermarking

Digital watermarking method finds its usage in audio, image, video and multimedia data for the purpose of ensuring security by maintaining the confidentiality and data integrity of the contents.

It is also required that non-repudiation in the media document is effected during the distribution process, so that the end users who have copied or acquired the document illegally can be identified and proven. To achieve this, a technique that comes to the fore is ***Non-repudiation Oblivious watermarking***, whose usage can be highly felt in the distributed environment for multimedia data.

Non-repudiation oblivious watermarking is a technique for creating undeniable watermarks. The owner of the data distributes his contents to N number of distributors, who act as agents in supplying the content to the clients after suitably watermarking them using their own watermark-key. The content provider or the owner will also be able to identify which distribution agent watermarked the content. In this system, it is made sure that a particular distributor does not watermark the content that would appear to have been watermarked by another distributor using the same watermark key, for the same content. By this way, any locations where the pirated copies of the media are released (i.e., locations other than genuine distributors), can be detected.

Most of the watermark scheme in distributed systems depends on the trusted third party (TTP) to verify the authentication of the watermark itself. This, many a times, results in a series of problems, like, locating the third party, relying on that third party, etc. and thus causing inconveniences in satisfying the security requirements. To overcome this challenge, oblivious watermarking can be employed, which does not need trusted third party to judge, and can automatically trace the source of the watermark embedding.

4.2.1 Distribution Model with Secure Copy Monitoring

This mechanism of watermarking allows non-repudiation of watermarked content in the distributed environment and finds itself useful when the source or authenticity of the media data needs to be known or checked. The content owner supplies the digital information to the distributors or agents, to be distributed to all eligible clients, and simultaneously, hoping to monitor the usage or security of each copy of the data at the client sites. There will be a standard procedure or understanding in place between all the genuine clients and the agents regarding the watermarking of the digital content. Based on this understanding, each client put the watermarks on the data, by which the owner of the content can monitor its usage without the assistance of the trusted third party. It assumes that both the parties at either ends has adhered to the procedures as agreed earlier and it also does not account for the proper application of the watermark on the content.

4.2.2 Non-repudiation Watermark Schema for Distribution

This technique works through a handshaking operation between the content owner and the agent, and which requires the usage of public and private key encryption algorithms. There will be a unique key, private to the content owner and the agent, which is subsequently essential in the identification of the media data watermarked by the agents or distributors. The entire activity is initiated when the owner of the contents sends them to the distributor. Once the content has been sent to the distributor, the distributor must contact the provider. The provider or owner responds by sending a random number to the distributor, which corresponds to the watermark-style. This random number also forms the basis in watermarking the content and its subsequent identification or verification by the owner. The distributor or agent generates a key pair (public and private key) for this content. The distributor generates the watermark-key by encrypting the random number using the private key that was previously generated. The distributor then watermarks the content with this key and sends across the public encryption key back to the provider or owner along with information that will allow him to obtain the watermark key given the watermarked content. Using this key, the provider can monitor or verify the authenticity of the content for a particular client and eventually, locating the leak locations.

4.3 Attack Characterization

One approach for improving the performance of general class of watermarking schemes is through attack characterization. This can be combined with most of the watermarking techniques. However it cannot be applied for spread-spectrum watermarking. In this technique, instead of a single watermark, two different watermarks, ***Robust Watermark*** and ***Reference Watermark*** can be embedded into the document. Robust watermarking is used to watermark the original signal whereas Reference watermarking is used to check for modifications in the watermarked signal. The channel efficiency can be analyzed through attack characterization by estimating the reference watermark embedded in a particular sub-channel.

It is an obvious and anticipated condition that an attacker will attempt to modify or remove the watermark by intentionally modifying the watermarked document. To defeat his intentions, one must strive to embed the watermark in such a way that it is difficult to be removed unless the watermarked document is considerably damaged and distorted. The cost of robust watermark removal by an attacker should be much greater than the value of the multimedia document itself.

In digital watermarking, when robust watermarking is carried out, the document is transformed to a watermark domain for embedding the watermark. The data at the modified spots of the document containing the watermark are called as watermark coefficients. Once the document is watermarked, there will be some distortions in it, which is often imperceptible. The attacker will try to extract the watermark by different methods and in the process, will damage the embedded watermark which will cause imperceptible changes. This distortion introduced by the attacker can be estimated by embedding a known reference watermark with the robust watermark.

The importance and working of the dual watermarking is that the problem is seen from

an attacker's perspective, i.e. regarding information concerning the attacker's actions. The effectiveness of the technique is most felt at the time of incorporation of the reference watermark for attack identification, prior to robust watermark extraction. The extent of the attack can be easily estimated from the embedded reference watermark, which gives the extent to which the robust watermark is damaged. The correct watermark can be estimated from this channel estimation.

For this technique to be used effectively, it is preferable to have the watermark (robust watermark) as a binary data stream which is repeatedly embedded throughout the document at distinct localized spots in the watermark domain. The different repetitions of these robust and reference watermark are extracted to estimate the correct watermark. The reference watermark is used to estimate the probability bit error of the corresponding robust watermark in that channel.

The host signal is embedded with both robust and reference watermarks and both of them are placed orthogonally so that they do not interfere with one another. The available space in the document watermark domain is shared between two watermarks, and because of the presence of reference watermark, the number of repetitions of the robust watermark gets reduced.

Robust watermark sequence is denoted by w_i, where i = 1,2,...,M and M is the number of repetitions. The corresponding reference watermark is denoted by v_i. In the embedding process alterative bits of w_i and v_i are embedded, so if it is attacked, it will be reflected on both robust and reference watermark to the same extent and thus the probability of bit error will be same for both watermarks.

This model is similar to transmitting the watermark simultaneously though M independent binary symmetric channels. The accuracy of transmitted message depends on the reliability of the communication channel. So the channel is estimated using the embedded reference watermark.

Reference watermark wi and the associated watermark vi are embedded in a localized region D_i of the watermark domain. From the extracted reference watermark v_i', the probability of bit error, P_{E_i} associated with the channel D_i can be estimated as,

$$P_{E_i} = \frac{1}{N} \sum_{k=1}^{N} v_i(k) \oplus v_i^{'}(k) \tag{4}$$

where exclusive-or $\oplus$ is performed between each bits of the reference watermark and extracted watermark to find the number of bits changed. Thus the average of bit error occurred is computed which gives the probability P_{E_i}.

The watermark is extracted by weighing each repetition according to the estimated bit error probability. It is weighed in such a way that less weightage is given for higher error rates, and those watermarks with lesser errors are given more weightage. Thus the embedded watermark will be deemed to be more accurate. The estimated watermark is calculated as follows:

$$w_i^{'} = \text{round} \sum_{k=1}^{M} \alpha_i w_i(k) \tag{5}$$

For each watermark bit i=1,2,...,N. a minimizes the bit-error rate of overall estimated watermark, which is given by

$$\alpha_i = \frac{\log\left(\frac{1-P_{E_i}}{P_{E_i}}\right)}{\sum_{j=1}^{M}\log\left(\frac{1-P_{E_j}}{P_{E_j}}\right)} \tag{6}$$

There are important advantages in using this model of the watermark channel. The model is simple and the parameter P_{E_i} is easy to be estimated accurately using the associated reference watermark. In addition, a different parameter P_{E_i} for each w_i is incorporated, which provides a localized assessment of the attack in the watermark domain. In most watermarking schemes, the extracted watermark repetitions w_i' are averaged to produce the overall extracted watermark. The attack characterization allows us to combine these repetitions based on a measure of their reliability to minimize the probability of watermark bit error. It should be emphasized that degradations such as filtering additive noise and lossy compression are reliably modeled using the Binary Symmetric Channel. This characterization, however, is not appropriate for geometric transformations on the signal such as rotation and scaling.

Improved performance for robust watermarking can be achieved through assessment of attacker's tampering of the signal. Watermark repetition throughout the signal provides diversity in combating a broad class of degradations. Characterization of the attacks can be used to optimally combine the extracted watermark repetitions to minimize the probability of error in watermark extraction. Future work involves extending the method to detect and identify geometric transformations on the marked signal to decrease the computational load required for watermark synchronization.

5 Summary

In this report, glimpses of the usage of cryptographic schemes along with few of the digital watermarking techniques are given. Watermarking is not a new concept and has been a while in the field of security in distributable digital data; but its robustness, reliability and imperceptibility for all types of media has remained un-explored and un-addressed in the past. This condition has to be explored, where a single watermarking algorithm is employed for all possible multimedia signals. Further, it has been proposed that crypto mechanisms can be used as an additional security layer in digital watermarking, which in turn combines the effectiveness of three different approaches. All these goodness, when incorporated into a single procedure, will prove to be a robust and healthy approach in resisting any type of attacks.