Unified Model For Fingerprinting Code Applications

M.Sc. Thesis

in

Electronics and Computer Engineering

University of Gaziantep

Supervisor

Prof. Dr. Ergun ERÇELEBİ

by

Ibrahim Nasiru ALIYU

January 2017

REPUBLIC OF TURKEY

UNIVERSITY OF GAZİANTEP
GRADUATE SCHOOL OF NATURAL & APPLIED SCIENCES
ELECTRONICS AND COMPUTER ENGINEERING

Name of the thesis: Unified Model For Fingerprinting Code Applications

Name of the student: Ibrahim Nasiru ALIYU

Exam date: 19/1/2017

Approval of the Graduate School of Natural and Applied Sciences

Prof. Dr. Ahmet Necmeddin YAZICI

Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science

Prof. Dr. Ergún ERÇELEBİ

Head of Department

This is to certify that we have read this thesis and that in our consensus opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science

Prof. Dr. Ergun ERÇELEBİ

Supervisor

Examining Committee Members:	Signature
Prof. Dr. Ergun ERÇELEBİ	.....
Prof. Dr. İlyas EKER	.....
Assoc. Prof. Dr. Sema KAYHAN	.....

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Ibrahim Nasiru ALIYU

ABSTRACT

UNIFIED MODEL FOR FINGERPRINTING CODE APPLICATIONS

ALIYU, Ibrahim Nasiru

M.Sc. in Electronics and Computer Engineering

Supervisor: Prof. Dr. Ergun ERÇELEBİ

January 2017

49 Pages

The enormous popularity of digital contents contains unlicensed and illegal activities which are involved in different perspectives. Despite several efforts are made toward an illegal copy and redistribution (legal and technical), still the problems exist with growing use of digital application such as storing and distribution of contents, the problem is becoming more worrisome. Attaching a unique invisible marking to every unit of object, is viable trading with problems of unauthorized copying, that is if an illegal copy is released elsewhere, it is going to be practical to find the original owner of the object (traitor). This phenomenon of enforcement, in a copyright protection, is what is referred to as fingerprint code. Dealing with code, for fingerprinting in digital contents, in the context of several users combining (colluding) to create a new, unauthorized copy (pirate). Theoretical bounds on the performance in the code are realized, with general methods of theoretically observe whether certain users are involved in the illegal activity (copy and redistribution) is presented, through checkmate items. In this paper, we have introduces collusion–secure fingerprinting code have been similar to construction by aggeloss and serdar; but attain a shorter code in length, faster accusation by splitting the group of users in to partitions and allow for code over arbitrary alphabets, and asymmetric fingerprinting scheme out of symmetric scheme from the early idea work of tardos in his fingerprinting code.

Keywords: fingerprinting code, traitor, pirate, symmetric, asymmetric.

ÖZET

PARMAK IZI KOD UYGULAMALARININ BİRLEŞTİRİLMİŞ MODELİ

ALIYU, Ibrahim Nasiru

Yuksek Lisans Tezi, Elektronik ve Bilgiyasar Mühendisliği Bölümü

Tez Yöneticisi: Prof. Dr. Ergun ERÇELEBİ

Ocak 2017

49 sayfa

Dijital içeriklerdeki muazzam popülarite, farklı perspektiflerde yer alan ruhsatsız ve yasadışı faaliyetler içeriyor. Yasadışı bir kopyaya ve yeniden dağıtılmaya (yasal ve teknik) yönelik çeşitli çabalara rağmen, içeriklerin depolanması ve dağıtılması gibi dijital uygulamaların artan kullanımı ile ilgili sorunlar hala var, ve sorun gittikçe daha da endişe verici hale geliyor. Nesnenin her bir birimi için benzersiz bir görünmez işaretleme eklemek, yetkisiz kopyalama ile ilgili işlemlerle, yani yasadışı bir kopyanın başka yerlerde yayımlanmasıyla uygulanabilir ve böylece nesnenin orijinal sahibinin (hain) bulunması pratik olacaktır. Bir telif hakkı korumasında, bu uygulama fenomeni, parmak izi kodu olarak adlandırılır. Digital içeriklerdeki parmak izi için kodla uğraşmak, bir araya gelen birçok kullanıcı bağlamında yeni, yetkisiz (korsan) bir kopya yaratılabilir. Koddaki performans üzerindeki teorik sınırlar belirlenir, böylece bazı kullanıcıların yasadışı faaliyette (kopya ve yeniden dağıtılma) yer alıp almadığını genel kabul metotlarıyla genel usullerle gerçekleştirilir. Bu tezde, Aggeloss ve Serdar tarafından bulunana benzer parmak izi belirleme kodu tanıtılmıştır, ancak, kullanıcı gruplarını bölmelere bölmek ve rasgele alfabe üzerinden kod uygulamak, ve parmak izi kodunda Tardos'un fikrinden yola çıkarak, simetrik şemanın dışındaki asimetrik parmak izi şeması ile daha kısa bir kod elde edilmiştir.

Anahtar Kelimeler: parmak izi kod, hain, korsan, simetrik, asimetrik.

To my parents for their good guidance and support.

ACKNOWLEDGEMENT

This project report has been accomplished with the contribution of some people whom deserved my appreciation in all forms. I wish to appreciate my supervisor Professor Ergun Erçelebi for his guidance, advice, criticism, encouragements and insight throughout the research, despites his numerous tight schedules and the required resources and conducive atmosphere provided by Gaziantep University especially, the library whose contribution towards completing this report played much role.

Finally, my family, friends and Kano state government of Nigeria deserve my profound appreciation for the tireless effort, prayers, and contribution in all forms. Thank you all for being there for me always.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF SYMBOLS AND ABBREVIATION

B	Buyer
M	Merchant
PIC	Picture
ARBITER	Arbitrary
Id	Identity
Len	length
Sym	Symmetry
Asym	Asymmetry
Log	Logarithm
OK	Okay
Seq	Sequence
No.	Number
Msg	Message
Sig	Signature
Com	Compute
ε	error
f	Function
K	Collusion size

CHAPTER 1

1.1 Introduction

The development of information in today's digital world, is configured in to the multimedia content for easy access, to pertinent information, for instance; Audio, Images, videos etc. And be sufficiently available. Despite the numerous advantages it provided, with the enormous growth, in the usage of digital means of storing and distributions. It is posed with detrimental challenges, to ensure the reliable and efficient contents. It is indeed important to find a safest mechanism to enhance a fruitful business, that is taking place in a digital network, and the intended contents are used by authorised users, who have the right for the access of such contents appropriately. In spite of the fact, the administrating to access of the contents, it is quite a significant element to make sure the content is used mainly for the receiver (authorised user). It is however, not enough for the protection of the content integrity. Some of the content protections such as encryption mechanism cease to be sighted, if the stored contents cease to be in the encryption domain. And this is seriously affecting the owner of original content distributor. The dubious act of unauthorised copy is a threat to various aspect as per as digital productions is concern. These contents of the data are copied illegally all over the world; as such the original owner of the product suffers Lots in losing by what it is caused. Initially, the illegal activities was under moderate control by the law enforcement agencies, to deal with redistribution problems and the fact that copies usually have a lower quality then the original, but due to the advancement of Internet technology and digital buffered data, the tune has changed, it makes such illegal copying so easy and cheap to the users in the Internet community. Knowing such illegal activity is taking place with minimum effectual hindrance to deter, only by making the risk of caught quite greater and the only by way this could be achieved is by usage of fingerprint code technology

(Schemes). The rationale behind the fingerprint code is that each user will receive a copy of the content data attached with a unique mark. The marking will serve to identify the uniquely content data and the particular user if his identity is linked to that data in anyway. Providing the defensive mechanism, in fighting against the illegal activity toward the digital contents (piracy) is "Traitor Tracing", where by the user who involved in the illegal activity is referred as a "traitor", and is traced back, whenever the illegal copy is identified by the tracer (distributor/owner). Traitor tracing is a technique used, in finding the possible culprit, who have unfaithfully responsible for the emergence of a newly compromised copy of a digital content, and redistribute it illegally. A user who is found guilty, of such illegal act is apprehend or in the situation where by the traitors colludes with the intention of comparing their authorised legal copy and able to produce illegal copy are referred to as "Pirates ". And those that are using such illegal copied contents are prominently called pirate users. These pirates' breaches the security of the privileges provided to them, with the intention of accessing a legal copy, and allowed the bad users (unauthorised) to have an illegal access. This process of traitor tracing, uses a very important handy mathematical tool, for a design of robust technique called "Fingerprint code", in achieving the fight against the traitors. Fingerprint code is an indicator of unique marking to each distributed digital contents, which can possibly takes various forms, depending on the context of the contents. There are actually, various models, in which the fingerprint code is quite applicable in traitor tracing phenomena, that will be discuss in application area of this work (Thesis). However, these models will be realistic in preparation with the anticipation that illegal copy and redistribution is always in the front line occurrence, and a good determination measures shall always be focused. It is assumes, a good countermeasures will be able to meet certain expectation below;

The capability of the scheme to track the initiation of the illegal activity (piracy). The pirate users should be revoke from the list of innocent users The innocent users shall not be distracted/harm. The tracing techniques should be able to provide a concrete evidence that, The identity of the pirate is purely distinguished. Thus, fingerprinting technique comprises the initial establishment of fingerprint code, Secondly the process of attaching the marks, thirdly identification of compromised copy and traceability. Today, research schemes in the field of information technology have came up with several fingerprinting techniques notably with its numerous

significances, especially in the collusion secured fingerprint; where a unique generated code (set of codeword's) are applied in to each of the user's copy. And can lead the tracer to identify at least one of the responsible users, in case, the content is manipulated under collusion attack and being distributed illegally. But lack of the comprehensively unified literature survey, on the applicable areas, where these phenomenal techniques are materialised and a new shorter way other then classically motivated approach. That is what this work (thesis) intended to focus on. This Thesis will focus on a survey on traitor tracing techniques, in which the fingerprinting codes on digital contents is applicable by identifying a shorter efficient Way, that satisfy the properties of identification capabilities in traitor tracing to checkmate the illicit usage of digital contents.

Considering the area of fingerprinting is a rather new and could not find standard question to ask or standard model to used, there are many ways to consider fingerprinting problems and many ways to describe it, which may give rise to other interesting research problems, some of the questions that may comes to mind naturally in fingerprinting research includes; What does the fingerprint system look like? Who are involved? How does it fit into everything together? What it is performance and how can we measure it? What are the parameters involved such as number of users, number of pirates colluding to create illegal copies, length of the fingerprint, some performance measures. Usually the performance measures is selected as some kind of combinatorial measures stemming from the fingerprint used, or the probability of failure to find the perpetrators. How can good performance be achieved? This question break down into two different ones; How to choose the fingerprint area and what method employed when trying to find the pirates. Both these questions are dependants on how the pirate behaviours when creating the illegal copy. This in turn may leads to curiosity on what the pirates could be to prevent the owner (tracer) from finding them. Some answers to these interesting questions were lead to a good platform in which further discussions may fits. The sections of this work (Thesis) are divided in to chapter(s);

Chapter 2 begins with the discussions about fingerprint which leads by a description of related research areas.

Chapter 3 describes the classically motivated approach and provided with an efficient way which is shorter on the fingerprint code techniques, which is general possible to trace and identify the collusion of pirates users.

Chapter 4 Analyse the achieved aims of the traitor tracing technique with help of asymmetric fingerprint.

Chapter 5 Conclusion

CHAPTER 2

2.1 Background of Study

2.1.1 Definitions

A codeword is a sequence of marks that are placed in content. And fingerprint is the assembling of marks. A fingerprint can be understood as a term of length l over an alphabetic size q, where l is number of placed marks. Given an (l, n, q)-code C, each codeword $x \in C$ is considered as a unique fingerprint of a receiver that has access to an object/functionality through its codeword. We, further, suppose that any codeword in Q^l would also be useful in accessing the same object/functionality. In an application where the above assumption holds, an adversary (we will call pirate) is modelled as corrupting a number of receivers (we will call traitor) and retrieving their codeword's. The pirate then runs a Forging algorithm that produces a pirate codeword in $p \in Q^l$. In any reasonable setting, the Forging algorithm is contingent to a marking assumption which enforces the pirate to produce a codeword that is possible according to what the pirate has corrupted. More formally, we define the marking assumption as follows;

Definition 1. Marking assumption; in a fingerprinting where an individual codeword is assign to the users set over an alphabetic size. Every user is obtaining a copy of content with the marks that are assigning according to his codeword. When users colludes, they can find the difference of the marks in between their codeword (forging strategy), and can produce a new pirated copy. but the aimed property of the marks is that the user of the object cannot make any alteration on the object content ,if they cannot find the marks in the object in which subsequently makes it useless and these property leads to what is known to be a "Marking Assumption"

If the set of traitor in the number of users, and for all set of codeword

C = $\{c^1 \ldots c^n\}$ where a c^j is an element that is contain in a length *l* of an alphabetic size *Q*, for which **j** is an element in the number of users set. It hold that forging algorithm yields a codeword P out (descendant) of the valid user's codeword's $C_T = \{c^j | j$ which is the element of codeword's contain in a traitor codeword's. The descendant set is denoted by desc (C_T), and defined as; desc (C_T) is equivalent to a fingerprint code in a number of users with a fingerprint a_i over a length l of alphabetic size Q, found as an element in the traitor coalition, as such a colluding traitors with compromised copy is possibly link back to at least one traitor involved in the collusion. It is in high probability that the traitor will be traced back among the coalition members and will not overshoot the number of legal users.

Desc $(C_T) = \{x \in Q^{`} : x_i \in \{a_i : a \in C_T\}, 1 \leq i \leq `\}$ Where x_i , a_i are the i^{th} symbol of the related vectors.

CodeGen: is the code generating algorithm that input 1^n and a pair of typical (*C, tk*) sample; where in the sample, *C* is an (*l,n,q*)-code that is defined over an alphabetic Q and l as a function of (*n,q*) and identifying key *tk* as an auxiliary information be used for identification (which may be empty).

Identify: is the second algorithm (deterministic) that is been employed after codeGen, takes input a pair sample (*C, tk*) where C is an element contained in an alphabetic size with a length *l*, and the *tk* as an identification key, which yielded result to a codeword found in the number of traitors or the algorithm fails to identify. This fingerprinting code is referred as open, if *tk* is empty.

Definition 2. The fingerprinting code is define as (codeGen, Identify) and this pair of algorithm is also denoted as (α, ω)-identifier, in a condition on an event A, if the following condition holds;

I- CodeGen with given input 1^n is resulted in producing identification key *tk*

And (*l, n, q*)-code C=($C^1 \ldots C^n$). This satisfy the conditional event A

II- For any forging algorithm that satisfies the marking assumption, it hold that; For every traitor contained in a number of users, such that the number of traitors is less Then or equal to ω. The probability of not identifying *(tk, P)* contained in the number of traitor will be greater than equal to *α*, where the pirated copy over an alphabetic size *q* Is the result of forging algorithm in corresponding to the input $C_T = \{c^j \mid j \in T\}$.

Scheme

Initially, Tardos scheme [37], the binary fingerprinting code provided by trades in general sense of arbitrary alphabets, has to be given certain parameters; accumulate the system such as *n* to be the number of users. The scheme distribute the codeword length (*M*) of binary to each user, while certain parameter (*M*) is selected by distributor, then the number of users and length of codeword is assembled in a matrix *n x m*, where j^{th} row matches the fingerprint assigned to the j^{th} users; And C described a set of column users, *(c* = number of column and *Xc* the *C x m* matrix of codeword disseminate to colluders expectedly (possible to colludes non deterministic) procedure *p* to create a compromised (authorized) copy out of their original copies. The compromised copy convey a fingerprint $y \in \{0,1\}^m$ that depends among the procedures and received codeword, that is *y= p* (*Xc*).

Code Generation

The distributing authority generate the phase of matrix in a dual, a dual randomized approach, for the first approach, he choose *m* random variable $\{P_i\}^m_{i=1}$ over interval $P_i \in [t,1\text{-}t]$, where t is a fixed small parameter satisfying $C_0 t \ll 1$ the P_i variables are identical and independently disseminated in respect to probability density function *f*, *f*(p) is symmetrically around p = ½ while biased towered value of *p* close to *t* and *1-t* ,

$$f(p) = \frac{1}{2arcsn(1-2t)} \quad \frac{1}{\sqrt{p(1-p)}} \qquad 1$$

Second approach is filling the columns of matrix *x* by individually dragging random bit $X_{ji} \in \{0,1\}$ according to $P[X_{ji} =1]=P_i$. Afore to the release of content data to the various receivers *j*, was watermarked in j^{th} row in the matrix having coming across with authorised copy with attached watermarked *y*, the data distributing authority

(owner)traces the identity of at least one colluder, though each user $1 \leq j \leq n$ with an accusation sum S_j as;

$$S_j = \sum_{i=1}^{m} y_1 \, u\,(X_{ji}, P_i) \text{ with } U(X_{ji}, P_i) = g_1\,(pi) \text{ if } X_{ji} = 1,\ g_0\,(pi) \text{ if } X_{ji} = 0 \qquad 2$$

Where g_1 and g_0 are the accusation function

$$g_1(p) = \frac{\sqrt{1-p}}{p} \quad , \quad g_0(p) = \frac{\sqrt{p}}{1-p} \qquad 3$$

The owner determine that user j is culprit (guilty) if $S_j > Z$, Z is referred as the threshold accusation is done all over the position i in y. All position with $y_i = 0$ are ignore, and for every position where $y_i = 1$, the sum S_j accusation is increment or decrement. It depends on how the suspicious evolved in that position; (if user j has 1 in that position, then the accusation is increase by positive weight $g_1(p_i)$). We will notice the suspicious decrease with higher probability p_i, since g_1 is a positive decreasing function, but if user j has 0, the accusation is précised by negative weight $g_0\,(P_i)$, which obtain more pronounced for large value of p_i as g_0 is negatively decreasing.

The selection of equation 3 of function g_1 and g_0, is quite significant, because of the property of fixed p_i, accusation $U(X_{ji}, P_i)$. In equation 2 for zero mean and unit variance, despites the fact that the variance did not depend on p_i but ease the analysis schemes.

2.2 Fingerprinting Applications

2.2.1 Watermarking

In the consideration of an important business transaction and distribution medium of digital content, on the Internet is pretty much excellent, however a serious issue is possible ways to counter numerous copyright violation and management. This has put the security of digital content to a high profile at stake. Today, the digital images can be found and used everywhere with permission or not. And this usage without permission causes a detriment to the market performance by the owner [17]. The innovative technique towards securing of digital content has proudly become a major concern in the process of a copyright protection. Seemingly, the owners (publisher) and relatives are

currently happy with this pressing problems, that is causing a major hindrance to their right to property. In considering how this issues is to be control, digital fingerprint can be smartly embedded/hidden in to the multimedia contents, in so many ways, by the application of "watermark " processes. Watermarking is a technique used to hide extra information in to the digital content, for the purpose of preventing an illegal copying and re-distributions. The watermarking also helps in identifying the original owner of the content. It is conventionally concerns with robustness (ability to withstand supplied change in the embedded watermarks), against the various attack, by the traitors. But this robustness is less effective when tackling the issues of coalition of attackers. The issue of coalition of attackers arises, by considering the average of several copies of the contents together, while extracting new manipulated copy [17]. The goal of watermarking here is; to trace and identify at least one traitor out of the colluders that participated in such illegality of forming a compromised copy [17, 19]. The coalition attacks does not hold, if there is only single copy of the watermarked content[19], but only the pirate users will be identified, while the responsible traitors will be trace free. it is base on this note, to collusion resistant and ability for identification, there by inculcating a deterrence to the collusion attempts by the users. Another signal processing attacks, which is not even producing an illegal copy, but the perpetrators publishes their own copies that are undistinguishable (closely same with the original copy). The scenario can be best understood, that watermarks are usually called a noise in a SS (spread spectrum) such as DCT. Especially when multiple watermarks copies are presented and certain number of traitors intends to conspire in combines copies, will attain a qualitative content deemed for distribution with general acceptance. The embedded hidden noise of digital watermark in spread spectrum (SS), has an energetic capability that is deemed in degraded acceptance of a content (Image) quality by the owner (watermark provider). The watermark (noise) may found everywhere within predetermined frequency band [19]. It is specifically noted that by considering the average signal that is differently represented in the copies. It is simply averaging the signal and finds their differences; one can obtain a combined signal for a number of watermarked

images. The noise is cancelled and original image(less noise) will hold. This is better for defeating an individual (watermark) scheme in a tracing.

There are several other schemes that enforces copy right protection, such as "Self enforcement", where the data content is encrypted with the uniquely different decryption key, and very vital information of user are all attached to the content. This will make the user not reveal his key, as he knew that his personal information is attached such as credit card information. It will serve as a simple and effective countermeasure for deterrence of key leakage by the user.

2.2.2 Key Distribution

In trying to make the information secured to the unintended users, especially in context of subscription channels, where an expected authorized user will be privileged to access a content of data, for example; the propriety channels (pay -tv). The contents of the digital data are protected against the unintended users by using the cryptographic encryption techniques.

A scenario; is when number of channel's subscribers will have to access same content of digitally broadcasted data, and authorized users should be able to use a key given/assigned to each one for the authentication, before one can be able to access the content (decryption process). The problem may arises, when the authorized user decided to reveal his decryption key given to him, purposely for him alone, but he share it to other, who are not suppose to access to access it. So, it is important to find a useful mechanism that will counter such leakage of the decryption key.

It is getting worse, if the users have came together to compare their decryption key and identify where it differs, and able to produce another new comprised key, that will serve/act as an authenticator key, and be used by illegal user (pirate users) to access the content. The major concern is to focus on how to securely distribute the assigned key and prevent the broadcasted content to be accessed, by preventing the traitor from redistributing the decryption key, that will an able the pirates user to have an illegal access to the content, that has been encrypted for security purposes.

Scheme and simple one level scheme) respectively. The simple one level; and security assumption of the scheme is of the personal keys base on symmetric encryption. In the general concept of the scheme; the supplier generates a meta keys which is a base set A of random keys and assign subset of these keys to the users, m keys per user(personal), then denotes the personal key for user U by $P(U)$ which is a subset of A; $(A_{P(U)})$ [1].

2.2.3 Pirate Rebroadcast

The digital contents are exposed to the possibility of contents such as (Images, Video, and Audio). With no exception, the rebroadcast is also affected with the pirate attack called "pirate rebroadcast attack", usually in this type of attack, it is difficult to trace back the traitor's key in the pirate decoder, as the pirate issued the copy of the digital content as his own productions, and be able to revoke the user's key. What normally happens from the source (broadcast publisher) is encrypting the content to be send to the number of users (receivers), it is expected that each user will receive a unique key, which means that the total number of user will have to obtain a key as such the publisher will need to make an effective use of medium, that the content will be transmitted, and at the same time to be able to cancel (revoke) any compromised user found in the mist of legitimate (authorized) users. It is quite difficult for the publishers to handle these keys as same to every user. There are some minor solutions to the problem of encryption of content, from the required user's storage and the length of the content.

In the beginning instance for the minor(trivial) solution, private/personal key is obtain by the user, while the publisher will send/communicate to each single user with a length of the content through a communication medium [2]. Notably, this way causes much squandering of bandwidth, but enhances optimality on a user's side. The second solution; the publisher prepares a different keys to assign it to each subgroup of users (receivers), and subsequently each user obtained the key assigned to his subgroups. This enhances the optimality of the content length, and causes the user to unexpectedly store a keys X^{y-1} (X=2, y= N), which is so big to be handle by the user.

In the case where the users give out their key, that will allow the unauthorized user (Pirate) can have access to the encrypted content, and then the pirate can possibly produce a new compromised key that will enable them also to access the content. The user who happened to leak the key is referred as "traitor". This issue can be resolve by employing the traitor tracing scheme [1], where the publisher will engage with the pirate decoder, and carefully observe the responses supplied. The identity of the traitor can be fished out, then immediately revoke (cancel) the identified user (traitor).

2.2.4 Trace and Revoke Technique

It is of no use for trace and revoke scheme in respect to pirate rebroadcast attack, as the attacker publishes the key, and the tracing authority will never find any information regarding the pirated publisher (traitor) [2]. The significant measures in curtailing this type of attack (pirate rebroadcast), is to employ the watermarking technique [2]. And each user receives a portion of the content, but this result in the consuming high rate of bandwidth. This types of technique is quite deterministic; that a trace and implicated user will be caught while in some cases, it tends to be probabilistic; where the technique seems to have a minor false in its implications. In this process of trying to apprehend at least one of the traitor. There are some strategies employed; one is sequential traitor tracing and the other is dynamic traitor tracing.

It is important to the note that tracing mechanism does not provide a revocation possible, but the significant to permit and add revocation to the techniques for either dynamic and sequential traitor tracing. It is not a simple and straightforward scheme, because it fails when one decide to cascade a broadcast encryption at decoder level, considering the sequential traitor tracing scheme, it must compose two encryption function which means the original decoder (authorised user) will acquires two set of keys; one set of keys is used for encryption/decryption that bind the marked object/content to the user (receiver) and the other set of key is for encryption/decryption of the broadcasted encryption. It is simple to witness unauthorised user acquiring the key material as fewer as two traitor users can evade revocation at the decoder level by easily supplying the keys of the one user for decrypting the sequential traitor tracing layer while the other is for decrypting layer.

In this type of attack, the sequential traitor tracing scheme will capture the identity of the traitor, but revoking the captured user identity will have no effect in terms of capability of decrypting the pirate decoder, in which will continuously using the key of the second unidentified user.

There is need for the provision of better ways to provide solution to this kind of problem, as such design scheme should be able to trace and revoke every illegal act on the digital content. Considering the result suggested by aggelos and serdar [5], present a primitive of tracing and revoking in to their construction, which has the ability to trace unlimited number of users with security and performance analysis on the previous techniques used. They were able to find the maximum number of revocation is bounded by the receiver storage and maximum traitor coalition will e track without false accusation, more over the setting is quite flexible especially in the choice of parameter[5][20].

The basic parameter (trace and revoke) scheme in pirate rebroadcast is communication overhead, that is the amount of replication necessary to transmit a key, rebroadcast bound which is the maximum number of transmission can survive before been entirely revoked. Lastly the marking alphabet which refers to the number of different variation of data contents created by the distributors. The communication overhead is linear on the number of revoke users R and number of traitors *t* grow in linear with number of the pirate users ($2R+ 4t$ in the worst case), and also, the communication overhead is quadratic; in which depend on logarithmic number of users. This scheme is constructed to keep on tracking and revoke arbitrarily number of users with rebroadcasting bound as revocations accumulate.

The binary alphabetic impose a bound W on the size of the maximum traitor collusion and its been improved to the maximum number of revoke users. Finally on the larger marking alphabet size of $2t+1$, where t is the number of traitors.

Traitor and revoke scheme T in the pirate rebroadcast consist of some procedures (Init, Transmit, Receive, Revoke) where N is a number of users, q is number of symbol in the transmission, alphabet $\sum$, the scheme T is stateful (set of state);

Init: is a probabilistic procedure that is given 1^n , which produces a set (N. J $\{I_U\}_{U \in N}$) where N= {1.....N} $\subseteq 2^N$ and $I_U \subseteq$ J for all U $\in$ as well as an initial state $\beta_0 \in$ state x 2^J .

Transmit: It is probabilistic procedure that given a state $\beta \in$ state x 2^J and (optionally) feedback symbol $f \in$ {1.....q}, it produce a new state $\beta^1 \in$ state x 2^J and subset J X $\sum$.

Receive: it is procedure that given I_U for some U $\in$ N and subset JX $\sum$, it returns the symbol $\sum$.

Revoke: It is also procedure that given a state $\beta \in$ state x 2^J and R $\subseteq$ N

Initiation; procedure for trace and revoke technique T , in pirate rebroadcast is important in the actual system instantiation. The Init procedure produce J which correspond to the set of keys in the system, I_U determine the key set of keys assign to each user U, that is each user U $\in$ N will get a set of keys corresponding to the set I_U and Initial state β_0 = (state 0 V_0). The set of V_0 is the set of key that has been initially revoked in the system (V_0 = 0). The procedure transmit possibly received some feedback symbol $f \in \sum$ (originating from pirate rebroadcast) and produce a subset j $\in \sum$ which determine way the content is transmitted for each (j, s) $\in$ j will transmit within encryption key j$\in$ J for every version of the of the content marked with s $\in \sum$.

Receive procedure will produce an accessible marking symbol given the content transmission with the key of the users.

Revoke update the state of the system taking in to account of set of revoked user R.

2.2.5 Round Complexity

The round complexity is another form of very significant traitor tracing technique, where a transaction of interaction is engaged between the tracing entity and pirate user (adversary). And observing the amount of communicated-interaction held [3], in a period of time, for the tracer to be able to identify the identity of the adversary. These techniques employ the

fingerprint code application for an efficient tracing operation. The administration of multimedia content to the various numbers of users is sufficiently distributed by meagre enrolling the encryption scheme [1], [3]. Considering this setting In which numerous users are expected to portion the same decryption capability, it is logical to expect some bad (Adversary) could have get some legitimate user keys and disseminate to others. In any way of stopping this illegal act of emanating the users secrete key, the curtailing operation can be constructed in to the encryption scheme. This type of context where the number of users are more, the efficient capability of tracing entity to interact with the generated pirate decoder and able to identify at least one of the users identity who take part in the illegal act of releasing his key to the unauthorized pirate decoder, this ultimately leads to the scheme best known as traitor tracing [1], [3], [2].

2.2.6 Fingerprint Ideology Concepts

In trying to deter people from illegal digital activities; the scenario could be achieved in one of the ways is by making the punishment for those caught. On the other hand is by technical (schematic) means, in which the "fingerprint" employed. The rationale behind the fingerprint is to mark each copy uniquely, so that every distributed copy is differentiated from each other, by doing so, it is possible to distinguishes between the copies. Thus, if the distributed copy copies only to person (legal) that identify themselves, it might be possible if an illegal copy is found, to identify the person who has the legal copy from which the illegal copy was made. After identifications, it is hope to deter the users from trying the illegal acts, since after they leak the illegal copy they no more have control over it and could possibly be traced and be identified and be prosecuted for such act. So the risk of been identified ought to make it less popular to engage in spread of illegal copies (compromised). Fingerprint code has to be in such a way that, it is not possible to be visualized to any user and could not remove it, else tried it; the fingerprint has a unique properties that if the object is copied, the fingerprint code will be also be copied so that the new copy contain same previous fingerprint code as the original copy. It is quite

interesting, but trivial in putting a fingerprint code where it does not affect the data, for instance in text documents, by coding a fingerprint as small alteration in the space between the letter or rows. This is bad because by copying only the interesting data, it is possible to remove the markings. In the text document case is just by retyping the whole text document which may be tedious but possible; the solution is to put marking in the data that forms the products. The most interesting part is the way of marking fingerprint(encoding digit individually by marking alteration in the object) is what distinguishes the fingerprint problem in to units which can be properly treated; the embedding and coding problem. The embedding is a way to make the alteration by encoding single digit in to the contents while the coding is how to choose the fingerprint in their abstract representation (vector of digit) as such the fingerprint will be robust against different form of attacks.

2.2.7 Fingerprint Goals

Some of the targeted goals expected to achieved in fingerprinting is; (1) Ability to provide distinct copies of content.

(2) When traced, the scheme should be able to identify the traitor(s)/pirate(s) users and as many as possible.

(3) The scheme should be able to observe and revoke several group of users who are guilty of creating compromised copy.

(4) Ability to fits various forms in a digital context.

In these cases, first is receiving the fingerprint code of the illegal copy and used the black-box to observe the situation by running of some identification algorithm.

2.3 Discussion on Related areas

Some of the related area to be describe and referenced could not be considered as a whole but a mere considering a bigger picture of further reading and how this work encompass in widened the horizon, considering the work of hurting and kutler 1990 and new edited version by ketzenbeisser and pititcolas [4] in fingerprinting, watermarking and stenography.

2.3.1 Stenography

The stenography; is a scheme in which the message is embedded with hidden in- formation, the embedded can be parameterized by a key that is used when retrieving the hidden information. The secrete message may also be encrypted for the purpose of additional security, but they are two different schemes. In general with several further references conducted by anderson and petitcolas [5]. A theoretical analysis of stenography has been presented by ettinger [6]. The practical bound of stenography by Adam strizel [7], uses the embedding of secrete message (hiddentext), seemingly innocuous objects (covertext) to produce a stegotext. When the receiver of the messages (stegotext), obtained a copy then he uses the known techniques employed from the stenography for the successful recovery of the hidden messages. Actually, the main aim of stenography is to communicate between parties in a disguise manners, so that the unintended third party cannot understand what are the actual meaning of the message in a whole communication. It is not exactly like cryptographic (secrete) form, despite the fact that, it provide a very private communication which is so difficult for the attacker to navigate and extract the meaningful information.

Another concept with regards to the stenography was illustrated G.J Simmons [8] in a communication between the ”prisoners” and the medium for the to pass information is through the prison guard(ward) presence. There are so many instances in which the technique of stenography is quite important, especially in the intelligence agencies for security purposes. Secured Stenography form; also enhances the security measurement against attacks, just like public key encryption, as illustrated by hopper, Langford and Ahn [9]. It is design in a form of game between prison guard (ward) and an oracle, in which;

many samples of covertext is allowed to ward, as he can develop sample distribution over covertext.

-The oracle will respond to any queries required by the ward for a chosen hidden text.

-The stegnotext form oracle or covertext provided by the oracle will be distinguished by the ward. It is noted that random sampling enquiries in guessing by the war disadvantage.

2.3.2 Watermarking

Digital watermarking is another viable method for the protection of ownership, right of digital audio, images, Video. It can be applied to different application include copy protection, device controls, broadcast monitoring fingerprinting, proof of ownership. It is much like stenography only that the schemes should be robust against active attackers, even if they know not only that an object contain watermark, but also if they know the algorithm principle of the method. The proposed application for watermarking is usually copyright information/validation. In [10], by hurting and kutters there is computer application of watermarking and several concept papers on watermarking could be found on communication of the ACM [11]. Recent techniques with cryptographic protocol by minouru Kuribayashi [12] elaborated the significance of implementations spread spectrum watermarking on encrypted domain. And conventional methods by kuribayashi and tanaka [12]. The trace and Revoking pirate rebroadcasting explained and analyzed by Aggelos kiayias and Serdar pehlivanoglu [3].

In the multimedia development in the field of communication infrastructure a lots of progress has been realisation of some mechanism to make possible for the contents (Images Video Documents) are legally used to the intended users (authorized user). As suggested by Wade, Min et al [13] that the techniques of watermark robustness in retaliation to numerous attacks, does not often addresses the issue of its robustness especially by the colluding users (coalition) in a particular same context of content. Simply in the attacking the digital content by collusion attacks, user colluder average as many as possible copies [11].

The digital watermarking scheme resistance in collusion attacks was analysed [17], in reviewing prominent plan creation (design). In collusion resistant of digital content with some unique fingerprinting technology. These goals are to provide a large picture on the well advancement in fingerprinting, where the multimedia requires a method to provide a robust embedding marks that is going to be able to withstand

any form of adversary attack especially when trying to remove the fingerprint from the content of data, And at the same time to be able to track those who are behind the illicit act. Considering this phenomena, it has many techniques proposed by numerous ways [17], was able to used broad spectrum. The simplified way to apply spread spectrum watermarking to fingerprint will identify each context, while encoding and inserting (embedding) permit them for identification application in the mist of users. And another way (option) in using spread spectrum watermarking is by employing code modulation, that will enable the designer to design on how several users collusion held, but the main due behind the digital right is existing of the fraudulent contents various different perceptive arises under different situation, but the fingerprinting techniques is provide appropriate performance such as ;

1. **Individual Catch:**

In this scenario, the paramount aim of fingerprint design is to have a certain possibility of catching at least one user that is involved in the coalition and at the same time not involve the innocent user while trying to catch the culprit. In this phase the tracer (detector), tracing approach fails when either the tracer to recognize and catch any among the users as in coalition group (false positive). It is very important to note this as in providing a sufficient evident to the court of law.

2. **Multiple Catch:**

In this type of design, the main aim for the fingerprinting is to cat as many as possible coalition team, despite the fact that, it is possible to accuse innocents users. The set in performing criteria consist of the expected fraction of coalition team that has been captured, at the same time with expected number of innocent users that are falsely suspected

3. **All Catch:**

This scenario, the fingerprinting is design to possibly in maximizing the possibility of identifying all the coalition team that have been involved in the illicit activity of unauthorization of digital contents. In this scheme consist of performance criteria for measuring the probability of identifying and capturing all involved. It is important to note, when constructing collusion-resistant fingerprinting on how detection will take place, strength of the fingerprinting and the efficient in tracing [17].

2.3.3 Information Theory

The technique of combinatorial tool in information theory which has been in used long enough in the content of data (information) hidden. A description of the certain application of methodology employed in information hiding [], for instance assuming that U= $X_1, X_2.....X_n$ is a string in which X_i are driven from the alphabet $\sum = \{1, 2, 3,.....C\}$. Let $N(i|U)$ be the number of occurrence of symbol i in U, while the type of U, denote T, is a C-turple of rational numbers $T= N(i|U)/N,.....N(C|U)/N$, From the above description, the standard representation of type is a turple of rational number. The summation of each one we will note an individual coordinates of the turple by T(a) for $a \in \sum$, the significant alternate view of a type is a probability distribution over alphabet, that is type T is defined as a probability distribution over alphabetic size $\sum$; $i, N(i/U)$ may have $n+1$ value differently to the type of string length n at most $(n+1)^c$. This is also so important as the upper bound of $(n+1)^c$ on the number of type is small (polynomial in n) when compared to C^n , the number of all string with exponentiation of n.

Lemma 1 Let $P_{c,n}$ be the number of type of string length n drawn from an alphabetic size $CP_{c,n <=} (n+1)^c$.

In a situation which the generation of string is random, when the symbols in a strings are independent and distributed identically then the string in a type setting are in the same probability. Below theorem is used to prove some result type with the establishment of relation among the entropy and its relatives.

Theorem 1: Let $X_1,X_2,X_3.....Xn$ be selected individually from the alphabetic size $\sum$, while a given a sequence $U, U = X_1,X_2, X_3.....Xn$ depend only on the type of U and $2^{-n(H(TU)+D(TU||Q))}$ where Tu donate type of U and is considered a distributing on the alphabet.

Proof: Let $U = U_1,U_2,U_3.....Un$ then $Pr (U = X_1,X_2,X_3....Xn) =$

Conditional type

We have to remember that type of a string is referred as distribution on the corresponding alphabet. For $U \epsilon \sum^n$ and V $\epsilon \prod^n$, the conditional type of U given U

will be considered to a set of distribution on $\sum$ indexed by element of $\prod$. In trying to understanding the definition of conditional type, there should be a defined joint type.

For string $U \epsilon \sum^n$ and $V \epsilon \prod^n$ defined string $U \times V \epsilon (U_j V_j)$, it is then jointly type of U and V define as the type of string U x V;

A string $U \epsilon \sum^n$ has conditional type R, given $V \epsilon \prod^n$ for any $a \epsilon \sum$ and $b \epsilon \prod$,

$N(a,b|U,V) = N(b|V)\ R(a\ |b)$ if for some $b \epsilon \prod$, $N\ (b|V) = 0$, then R is not determine in unique but the set of all $U \epsilon \sum^n$ with conditional type R, given V is unique as it is noted that for each $b \epsilon \prod$, $R(.|R)$ is a distribution on $\sum$, so R can be represented by a matrix that consist of row and column. The set of all string U that have conditional type R, given V called R-shell of V denoted by $C_R\,(V)$.

Theorem: Let $V \epsilon \prod^n$ and $X_1,X_2,X_3.....Xn$ be chosen independently form an alphabetic $\sum$, according to a stochastic matrix $W(.|.)$ that given a distributions on $\sum$ conditional on the string V, and the probability of $U=X_1,X_2,X_3.....Xn$ depend on T_V and $T(U,V)$ and is equal to

$2^{-n(H(TU|TV)) + D(TU||W\backslash TV))}$ where Tv denotes type of V and $T_{(UV)}$ is the joint type of U and V. Given Tv and $T_{(UV)}$, T_U is uniquely determines.

Context

Cryptography is preferably interested in hiding the message content of information. For the instance, we consider a spy agent who is working in another country and make to sure he is safe from any intimidation and interrogation; he will hide the gathered information to be transmitted in exchange to his country. In this situation a cipher messages will only make it worse and attract suspicion, so in another form, it has application, the act of hiding the message has some qualitative properties especially in a copyright protection for picture which embed a serial number in each content copy of the picture, so in such situation the message should be hidden and cannot be percept. It also be long enough so that each content copy has individual serial number, it will be good suggestion if embedding the serial number to have much complication that embedded a unique message in a picture or some other media.

It is not new to note that this kind of situations since long ages ago people were kin to find a tangible method that they will be hiding very top important information especially when passing it to others. In the year 440 B.C. Hastianeus, the tyrant of miletus ancient city Aydin province of turkey, tattooed an information on a shave head of his trust slave. After some time hair grew , the message were hidden, so this method/technique have been used recently in 20^{th} century by German spy [29]. In the mid of 1980s magret thather [29] acquire a word processor be programmed to be embed a certain unique key for the spacing of word document, so that when a staff of the cabinet expose a certain document and found elsewhere, he can be easily traced back. Despite the significant important of hiding information since form the ancient time, it is until some decade that the academia have stated serious research as a field of study, where the first international workshop about information hide was initiated in 1996 at Cambridge, UK [29]. The essence of the gathered workshop was for the placement of more researchers who are interest and work on information hiding and were to know each other, and some other important issues regarding standardisation of terminology in the field of information technology .

Looking at the research situation issues studying in this field of information technology especially hidden technique, we will have to significantly focus on one of the most important aspect; the fingerprinting under the marking assumption which this thesis adhere to. It is cryptographically natural, for this technique play a role in a problem of traitor tracing. The traitor tracing is the father of fingerprinting with the rest of issues conveyed in the information hiding. For us to look closely to fingerprinting with the siblings of fingerprinting code which are identifiable parent property (IPP)-code. They are result of research on traitor tracing problems.

2.3.4 Lower bound capacity of fingerprinting

The fingerprinting under marking assumption, were probabilistic phenomena and yielded on lower bound on codeword length, lately the fingerprinting have used the theoretic technique of information as a concept of capacity and rate. In considering of boneh-shaw [15] on traitor tracing, the scheme of code constructed in [15] length $0(t^4 \log N \epsilon \log 2tl/\epsilon)$ where $L = 2t \log 2tl/\epsilon$. This can be understood as a scheme with rated $0(1/t^4)$, in which the construction convey two steps; the fist is easier and simple

but less in efficiency but still works, and then the concatenation of code in a random error correcting parameters.

Also tardos [44] biased generation of code, where it simplifies the technique of non construction of fingerprinting scheme base on error correcting code. The construction he used yield in code of length $100t^2$, this achieved an optimal order of magnitude with the possibility of factor optimization, but this construct *100* of the length $100t^2 \log(N|\epsilon)$ was improved severally. Most of these improved papers went for proof in [37] to improve the code length without alteration to some improvement in (*t*-secure scheme) rate. The most significant achievement was by skoric and celik [36], where they make a reasonable assumption with non justification in mathemetics. The improvement of smaller factor of skoric, vladimirova, celik and isra [38] and Blayer and tassa [32] was mostly base on experimental evidence, despite the fact that some more research latter extracted some formulas for the parameters in the proof of [37].

In a various category of approach to the problem, alot have tried to find a high rate in construction of t-fingerprint for small t. The hope will be that the technique will generalize to arbitrary t. An early research paper regards to this [41] achieved the rating 0.026 in pirates by using *(2,2)-code* separation. The rate was improved to 0.2075 of 2-fingerprinting by the used of random generated code and its base on this Anthapadamanabhan, Barg [30] initiated the construction of t-fingerprinting scheme and achieved quite best rate for 3-pirate with 0.064 while Anthapadamanabhan, Barg, and Dumer [31] have constructed a t-secure fingerprinting scheme that rate *t=2* and *3* are more higher to the previous rates realised, but deteriote exponentially with t. They used individual uniform randomization of codeword for two to three, using typicality extensive analysis and they were able to achieve the rate of 0.003.

Dumer [33] worked on simplified phase of fingerprinting under marking assumption that all codewords are selected uniformly at random from the string of hamming weight *P*, this prevent the pirates to generate compromised copies of hamming weight *P*. It [21] has related $\geq 0.9/t^2$, code construction is biased while attaining capacity using bias code generation, it requires not less than $O=\frac{\sqrt{t}}{4\ln(2)\log(t)}$ different bias values when pirates users are allowed to used any strategy ordain to marking assumption.

The most prominent hindrances to the construction of [28,39] is a slow accusation algorithm, which is exponential in t, as such more effort have been made for improvement in rating the construction [37] and some papers lately keep the running time linear. Huang and Moulin [40], building on [39] in construction of a family of subgroup fingerprinting scheme which has linear time accusation algorithm, their rate is higher than [37] and game theory nature.

2.3.5 Identifiable parent property code

The previous work of binary fingerprint code would not be deterministic, which proof to be general to fingerprint code over an alphabetic size, the important point for the proof is the that in a fingerprint code a group of pirates that sight the symbols that differs in a column can use any element symbol of the alphabet, in such column of the compromised copy. If the restriction in the set of eligible pirate strategies, then we achieved a version of restricted fingerprinted code which the deterministic exist. This group of code with restriction are identifiable parent property, in such code any column pirate can yield an output in one of the element symbol that they can identify in their codeword.

This group of code in [24] such can be contrast with 2-fingerprint code; Let *C* code of length *n*, two codeword $a = a_1 ..... a_n$ and $b=b_1.....b_n$. Let $D(a,b) = \{C|\forall\ 1 \leq i \leq n, Ci \in \{a_i, b_i\}\}$.

D (a,b) = set of descendant of a and b, while *a,b* are the parent to each member of *D(a, b)*. $C^* = U_{a,b} \in_c D(a,b)$, then we say *C* has identifiable parent property (IPP), if each $c \in C^*$ pair *(a,b)* such that C $\in$ D(a,b) have common member.

In another [28], the researchers determine the value in precision of *f(3,q)* which in more specifically *let* $(r+1)-1 \leq q\ (r+2)^2 - 1$, then this can be written $q= r^2 + r + k$, where $0 \leq k \leq 2r+2$.

The shows that for $q \geq 24$, maximum size of IPP q-ray code length three is equal to $3r^2 + m$, where m may be zero or $3k-6 \leq m \leq 3k-2$, in which the exact value of mm depend on the value of r and *k.* The maximum size of IPP-code for $q \leq 48$ was found, alot have considered decoding algorithm for IPP-code and general IPP-code to arbitrary number of pirates.

Unbinding Problem

Determination for fingerprinting protocol to solve the issue of unbinding so that the relation between a certain transaction conducted by a buyer and seller, and fingerprinting technique information, Memom and Wong [25], the fingerprint information which is going to be embedded is not well considered, but merely relative in respect to user's information such as gender, name, address etc. When in any situation the selling authority finds the illegal copy of the content and detect the matching to buyer by identifying (extracting) the fingerprint, he can provide the collected proof to the court. However a malicious seller can frame the detected buyer with embedding the detected fingerprint in to the other content that he sold, so it is possible for the seller to obtain the detected fingerprint and re-embed it in to another more expensive content data, as such he can get more compensation.

Lei et al. [24], provided fingerprinting binded with a common agreement by providing signature of certification authority and digital contents associated with unique form with a log file. The anonymous of buyers, digital certification authority is produce in the fingerprinted protocol. A buyer *B* is randomly selected with a key of (PK_B, SK_B) where PK_B and SK_B are public and secret keys respectively. He send PK_B which is a pseudonym associated with *B*, to *CA* for retrieving anonymous certification ($Cert_{CA}(PK_B)$); when *B* madean order to a seller *S*, he check the authenticity of $Cert_{CA}(PK_B)$, then *S* ask *WCA* to generate a uniquely watermark *W* for the specific transaction between Buyer and Seller

Picking a Fingerprinting Code

In the selected choices of fingerprinting code is quite flexible, whereby could be possible to select a different code in every phase (stage) of subset that has been identified as traitor, and in due wishes to keep same fingerprint code throughout. Additionally choice will show when every deciphering is taking place in the contents, but the fingerprinting code is independent from the key stored in the device. The code to be used is re-arrange in a careful manner by re-assigning in a subset to a new codeword, the most important selection that differs in fingerprint code and previous mechanism that are employing some tracing ability [5].

This security called Watermarking has important dual portions of stage, One is the embedding stages and the other is detection stage. From these two stages, the former is the stage when the watermarking is integrated in to the content of the data, to prevent changes in the content of that particular data. While the detection stage is the one that, the embedded watermarked content can be identified using the secrete key, used at embedding stage, only known to the authorized user. The expected qualities derive and exhibit by the good watermarking technique, includes;

Security- The watermark should be able to secure the content of digital data, by not allowing the attacker to make any form of attack on the digitally multimedia content, and he could not possibly know/find the secrete key, even if he knew the algorithm used in the technical processes.

Robustness- The watermark is going to robust, if the content of the multimedia data is been modified and yet, the quality is identified to scale through, for the instance in conversion and compressions of data.

Capacity- When the content of the data is capable of allowing the several integration of watermark to same content, so as to enforce the detection of reliable content.

All in the process of watermarking, it is quite important to note with differencelessness in the content in terms of qualities of any forms such as visual, Audio etc.

By the fully technique for the establishment of watermarks, a copy right enforcement measures has been ordained, in digital contents and tracing capabilities. This provide a comprehensive defence mechanism for digitally embedded watermarks in to content (image), as such detection and readability is addressed, largely linking to the detail information about the owner/distributor for right to license. [13], [14].

2.3.6 Various Forms of Attack

1.Cryptographic Attack; this is an attack purposely intended to break the security methods provided by the watermarking technique, by illegally removing the hidden information embedded, through a cryptographic processes such as

brute-force search of oracle attack (production of non watermarked signal). It is practically very complex computation, but theoretically possible.

2. Re-modulation Attacks: This type of attack is aim at modifying the prepared watermark by using modulation, in an opposition to watermark embedding. For instance, where the original watermark (estimated) had been correlated with the original watermark, it yields a good estimation, and at same time, the estimated watermark can be subtracted from the original (watermarked) content. However, subtracting close less may decrease the quality of the content, not necessarily the watermarking. And also be defeated, when subtracting an amplified part of the estimated watermarks.

3.Removing attack; Is the phenomenal attack of removing the watermark embedded in a content, without breaking the security of the watermark algorithm.

4. Geometric attack; this type of attack that does not completely intended for the removal of embedded watermarking, but provide to cause distortion to watermark detector synchronization with embedded information. It is only when proper synchronization is attaining, then the detector can be able to recover the embedded watermarks.

5. Protocol Attack; This attack can be defined as a concept of invertible watermark, that is when the attacker extract his particular watermarked content and subsequently claiming the right to ownership of the data content. So to counter this attack, the watermark embedded shall never be extracted from non watermarked content. And the non inevitability to prevent the creation of ambiguity on the claims in respect of right to ownership.

2.3.7 Traitor Tracing

The case scenario; the owner of the content broadcast encrypted data that should be accessible to only central privileged users, each user has unique decrypted key that can be used to decrypt the broadcasted data. But some user may decide to collude and create a new key which is totally different from any of the authorized users and still be able to decrypt and access the content. In a traitor tracing schemes if the number of users colluding is less

than some given threshold, it should be possible with low probability of error to trace atleast one of the creators of the new compromised key. The traitor tracing was first introduce by Chor, Fiat and Noar [1]. Boneh and later updated in [15]. [16] [17] [18] [19] [20] [21].

2.4 Understanding some terms

Understanding these parameters will help in the better significances on the subject matters;

Partition; is a user set in the collection of subdivision, equivalent to disjoint set of codeword's. Partitions in every subset contain one finite block cell.

Attributes (disjoint); is an index subset features of singleton in the user set of a block partition.

an Object can be best understood as a collection of digitally stored data such as texts, images, videos, and Audio. When an object is referred as same working purpose for usage, this could be having different meaning for different people and depend on what the object is used for.

Illegal Distributions; all forms of distributions to other then by a rightful distributors is considered as unwanted and should not be practice but due a certain users, they leak the contents to authorized personnel's, these users are called traitors.

Pirates are users performing illegal distribution. Number of pirate *c* and the set of pirate is denoted $P = P_1, P_2 ..... P_C$, if P contain more than one pirate is called collusion cooperation The pirates will distribute to a person other than the user who want a copy, this is prohibited and to discourage the distributor should use fingerprinting.

A collusion is the coming together (cooperation) to work in a purpose of producing an illegal copy of a fingerprinted object.

The set of marks are what consisted in the binary form for the fingerprints code, which is denoted by 0, 1. In coming together of the pirates, they can detect their difference in the means in the marks in between of their copies, else the

marks are undetectable. A detectable marks can be change to any of the state, for example 0, 1 ,? in which the marks' ?' is invisible and can never be understood which is it, among '0's and '1's. It is noticeable assumptions in the marks that satisfying this sort of property exist for the object being fingerprinted. Note that a situation where by no collusion fingerprinting is trivial by the marking assumption, the fingerprinting allocated to a pirate will be visible (readable) in any illegal copy he distributed.

2.4.1 Distributions

A mark is a bit encoded in an object, in a systematic way such that '0' is encoded as no changes is made from the original content of the data and '1' is encoded as some specific changes made from the original. Every object contains *n* means the located positions of the marks are kept secret by the distributor. A detectable mark is that which differs between compared objects and undetectable marks is a mark that is the same in all compared objects, the marks are assumed to;

In the collusion strategy, the pirate can find a specific marks if only if it is detectable, they cannot changes an undetectable marks without destroying the fingerprinted object, for any undetectable mark the pirates can create a new object in which that mark contain '0' and '1' or 'e' where is the erasure (wiped-out) denote anything but '0' and'1', sometime restriction and allow a detected mark's state to be chosen only from 0, 1. The systematic way described the fingerprinting system works;

The publish content own by distributor distributes to number of marks. Distributor chooses random code model for the performance measure.

The distributor finds marks and embeds the codeword's of the embedded code in to original, creating copies. The distributor distributes the copies to the user. While the illegal distributions reflect the unlawful act by the pirates in distributing the illegal copy of object contents after creating a new copy. An illegal copy is the copy that has been produced (created) by the pirates and illegally distributed.

CHAPTER 3

3.1 Rethinking the Fingerprinting Code as Partitioning

Given a user set U, we define partitioning $P = \{S_1, S_2, \ldots, S_m\}$ of that user set to be collection of disjoint subsets such that their union consists of the user set, i.e. $\cup_{j=1} S_j = U$.

Given an (l, n, q)-code $C = \{c^1, \ldots, c^n\}$ over an alphabet $Q = \{e_1, e_2, \ldots e_q\}$ it is possible to create l many portioning since each codeword C^i is an turple $(C^i_1 \ldots\ldots C^i_l)$ where

$C^i_j \in Q$ for $1 \leq j \leq l$. More specifically a portioning $P_j = \{S^j_1, S^j_2 \ldots\ldots S^j_q\}$ of the user set

$U = \{1, 2, \ldots\ldots, n\}$ for the j^{th} column is define to be as follows:

$$Sji = \{k \in U : C^k_j = e_i\}$$

Fingerprinting code in classical approach is that it provides ` many partitioning where ` is the number of columns in a fingerprinting code. In this process, traitors compromises number of user's codeword's through a forging algorithm, that enable him produces a pirated codeword as a valid codeword over a number of user's columns.

It is not possible for the forging algorithm to be effective and produces a pirated code- word without the involvement of valid codeword from a traitor/pirate. That signifies the marking assumptions. Since the code is not changeable, this corresponds to several fixed partitioning. The phenomena of fingerprinting code

necessitate the assignment of unique codeword over the set of users in a form of partition. Codeword is a set of marks. The user obtains a content attached with some marks that are set in accordance with his assigned codeword through a defined attributes. Attribute are singled fixed partition. The codeword is kept secret in order to prove for the benefit to identify a traitor. A traitor is a privilege user who involve in an illegal activity but coalition has less information in constructing the pirate codeword, in which in many cases the actual deployment of fingerprinting code, When users colludes, they can find the difference of the marks in between their object codeword, but the aimed property of the marks is that the user of the object cannot make any alteration on the object content, if they cannot find the marks in the object in which subsequently makes it useless. But in the case where no, then it will be the user's serial number easily to be identified. After the codeword's are randomly generated, and the pirate user to produce the valid codeword, through a forging algorithm. Forging Algorithm; is the set of codeword's that can be produced by a pirate using the codeword owned by authorized user (Traitor). So the forging would be corresponding to producing a valid pirate codeword out of original codeword. the tracer will run identification algorithm hoping to identify a codeword that was given to one activities of piracy (piracy) who took part/role in the forgery. This algorithm runs over several partitions obtained in a fully collusion restricted code. This is what is happening in the traditional way of traitor tracing/tracking, but in this work (thesis) the new way is provided which is expected to sufficiently work effectively by making the fingerprinting code shorter and robust in tracing the traitor and will not allow a whole detrimental to be made by the pirate user before they should be traced and identified. The strategic process in an adaptive formation in which fingerprinting code is a dynamic phenomenon for the assignment of codeword generation and tracing schemes. The concept is codeword is generated in a partitions proportion embedded with the content of the data to be distributed to users, and the authority is observing the behaviours of the first released partitioned content object data and when they (authority) find copy of the illegal duplicate content data other than the one legitimately assigned to, then the responsibility of the tracer is to generate the identification algorithm for tracing. Notably the numbers of users N are divided in to subset with respect to the size and

implementation through ”Walking Procedure”. The user set will be updated to the new current partition and generate and assigned the new codeword in the fingerprinting code, from the walking argument that is usually involve to every conventional tracing scheme, to this also was involves by taking the pattern of the content in a random fashion until coming to a certain point where the pirate decoder could not decrypt the content. This process is repeatedly continuous with various different portions of the partitions. After all the observations of the partitions behaviour of the resulting output will infer a corrupted user identity. This type of strategic pattern is linear in formation and can utilize the number of corrupted users using the fingerprinting code

3.2 Shortening Code with Selective Partitioning

After detail descriptions; instead of selection in a conventional way, we may choose another way of selection and still be able to trace one of the traitor who took part in the illegal revelation of content in a much shorter time. The question is how do we do it; is we select length l of them (set of code partitions) as a subset and then apply a setting of fingerprinting code base on classical fingerprinting code. In a binary fingerprinting, we make the partition in a form of subset level hierarchy structure, forming a phase of levels up to the final singletons. In this new strategy, the traitor can be found in any phase of partition identified and that partitions can be further split in to two phase partition then we apply the binary fingerprint code. This process is repeatedly continuous until it enables the tracer to categorically identify a traitor. In this process; the subset cover strategy is adapted, whenever the com- promised content is realized, where an individual singleton of the subset at all level is assigned a different marks in conjunction with the fingerprinting code to be chooses. The response acquired after a significant casting in transmission (which is same length with the code) yielding a compromising in the course of feedback with a “pirated code word”. A tracing algorithm is employed afterward to reveal the identity of a traitor, in the subset level. Immediately after the identify a traitor, the subset level split to another two subset and update the set of users (new singleton), to be reassigning new codeword in fingerprinting code. It is strictly optional, when selecting fingerprinting code for the subsequent levels, in either new different code to each phase level. One of the most important part

particularly when selecting the fingerprinting code, in this work (thesis) comparing with the prior techniques, is that; this code required with the respect to the number of codeword is corresponding to the number of traitors. While in the prior techniques; it provides code in respect with number of codeword equivalents to the size population yielding a fixed minor number of traitors. Because of that, introducing fingerprinting code will allow (random) traitors in collusion to be traced (unlimited).

3.3 Applications of fingerprinting code

The (CodeGen and Identify) are two algorithm defined as a fingerprinting code. CodeGen; is an algorithm in which input 1^n and illustrate a couple (C, tk), where C= (l,n,q)-Code over an alphabetic Q with l as a function of q, n; F (n, q). And *tk* is an additional information apply for identifying that could be empty (useless). Identify algorithm applied on input (C, tk), illustrated from CodeGen (1^n), with a codeword $C \in Q^l$), will link to a traitor $t \in E[n]$ else fails. When the codeword is constructed out of the traitors collusion ω, after they average their codeword and identify algorithm yields a traitor with a failure probability of at most α, then we say the fingerprinting is an (α, ω)-identifier.

3.4 Encryption for Multiple Users

The encryption system operation has some state; that includes (key distribution, Transmission, Receives) with some parameters such as number of receivers that are related with (set of keys, plaintex, and ciphertext).

Key Distribution; is a probabilistic algorithm that input 1^n, and yield (tk, ek, S_k $i \ldots S_{K\ n}$). S_k i is a keys that is used for decryption that will be assign to i^{th} user, and ek is the encryption key, tk is the key referred as tracing key.

II- Transmit; is a probabilistic algorithm that is given $m \in M$, and assemble an element $c \in C$ and output C: transmit (em, m) i.e ($E_{k\ 1}(m)$) $\ldots E_{k\ n}(m)$.

III- Receives; A deterministic algorithm that input C, given from the state (transmit); (ek, m) and user key ($S_{k\,i}$) access the number of user $i \in [n]$

CHAPTER 4

4.1 Asymmetric scheme

In addition to the information about fingerprinting code, our strategies and symmetric, but understanding the other component of fingerprinting as asymmetric fingerprinting notion in B.pftizman and M.schunter []. In this types of fingerprinting code (Asymmetric), where the content document is distributed or sold, the merchant will like to make the document/content does not redistributed illegally o a large scale, especially dealing with world wide web. And in curtailing this some approach has to be involved such as temper-resistant modules in a buyer's machine that prevent copying of at least internal description of the content data, despite the fact that this has some limitation. The other approach is indeed the fingerprinting code which does not rely on certain special hardware, to deter copying and redistribution by the costumers.

Certainly in some schemes, the two entities will have knew the content data with the fingerprinting such as; the buyer (receiver) of the original and merchant, but contrary to the only one person can make the digital signature, the essence of this scheme is after the receiver(buyer), he is the only personnel who knows the fingerprint, however subsequently if the merchant found such copy else where he can identify the buyer and prove to third parties that the buyer bought (received) this certain copy.

4.2 Asymmetric property

As introduced by pfitzmann and sadeghi [26] for the purpose of attaining an asymmetric property of public key cryptosystem has to be homomorphic; where in the privilege provide in a homomorphic, allow the selling authority to acquire a ciphertext of fingerprinted content by operating encrypted fingerprinting added to an encrypted content of data (original). As the content data to be broadcasted (ciphertext) is computed using receiver's (buyer) encryption key, so that only the buyer will be able to decrypt it, as such will only be the one to obtain a copy of data

with fingerprint. The property of homomorphic to public-key cryptosystems is quite deliverarate to be applified to protocol of cryptography as an operation that can be conducted without exposing the plain value [27].

Let E(M)= encrypted message (ciphertext). The homomorphic property of the below equation; g (E(M_1), E(M_2)) = E (f(M_1, M_2)), where g, f can be one of the authentication operation; XOR, multiplication, summation etc, that is relative to the applied cryptosystem and attaining algorithm. In more often, the public-key cryptosystem chooses product for *g(),* for instance, if M_1 is a copy of content and M_2 as a fingerprinted, the fingerprint can be attached on to the copy of content without decryption by multiplication of these ciphertext. Especially there is manipulated using user's public key encryption, as such the fingerprinted content is decrypt. Only by the buyer, hence the asymmetric property is attain and satisfied. The embedding technique base on the property of homomorphic is basically conducted to each element of fingerprint detail that is exposed in a bit-sequence or spread-spectrum, hence each element is individually embedded on to its corresponding position, for example M_1 could be just part not entire copy of content, but some components such as (frequency) element to be fingerprinted by watermarking, explained and provided by katzenbeisser and petitcolas [] in a frequency components; E(M_1) ={E($M_{1,1}$), E($M_{1,2}$),.....E($M_{n,n-1}$)}

Asymmetric Protocol base on bit fingerprint; the scheme of asymmetric fingerprinting with the transaction party (Seller and Buyer) involve in the embing of fingerprinting, where the buyer encrypt a fingerprint and send it to seller, then the seller verify the copy sent by the buyer from real fingerprint and embeds it to his encrypted copy by taking the product of the ciphertext and lastly the buyer got the fingerprinted copy without exposing his identity. In the context of an anonymous fingerprint protocol presented by pftizman and sadeghi [26], the parties involved (Buyer B, seller S and Registration centre RC) and the registration centre generate a pair of keys secret and public keys and distribute to numerous receivers of the system. If *B* begin a transaction to a seller *S*, firstly *B* must register at *RC*, then *B* withdrew a digital coin which include an identify proof (*id*) of his identity (fingerprint), id, and signature that will be verify by registration centre, public key and validated the legitimacy of the buyer.

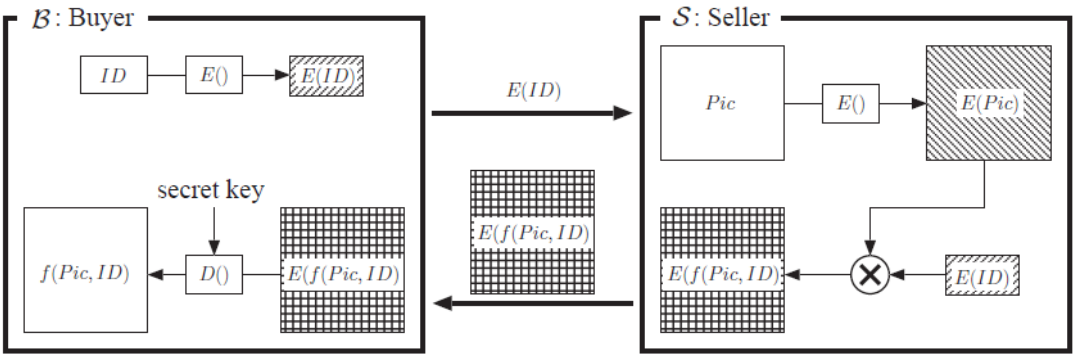

Figure 4.1 Flow of asymmetric fingerprint protocol

In the protocol of fingerprint, B encrypt his fingerprint and send to S and by using zero-knowledge proof, B proves that the copy of the content is equivalent to W. After S verified and convinced with evidence of validity of the content (ciphertxt) then the he encrypt his image and multiply with the received ciphertext and be embedded with fingerprint on his image base on a homomorphic properties. For proving the validity of ciphertext really convey with fingerprint without revealing the plain value, two bit-commitment scheme are possibly applicable, one can could be discrete algorithm assumption and quadratic residue brassard et al.[40], which the security issues is depend on P-subgroup assumption and residue respectively.

4.3 Algorithm and parameters

The fingerprinting code scheme works for a specific space content of content data to be sold out, we referred it as *Pic_Space*, in consideration of illustrated pictures. The original content merchant sold has to be secret and the receiver might have initial understanding about it, we have to assume that *Pic_Space* is equipped with an arbitrary probability distribution. Receiver identified with a public key that is constant over some time, may buy/receives several data from same merchant. We use an input parameter text to the fingerprinting code to uniquely separate these sales, the text would be identifying the contract which referred as sale (This is important if different data/contents are sold with different licensed terms).

Definition 1: This scheme of asymmetric consists of four components protocols, *Key_gen, Fing*, identify and dispute. And they have a certain security parameter *K* as common input and probabilistic polynomial-time in K. So also the scheme contain *Pic_Space*, the space of content of data to be sold, with given probability distribution and *text_space*; both are the subset of {0,1}*, In order to have much simpler we have remove in the notion that *Pic_Space* and *text_space* may depend *on K. N* is denoting

the maximum number of times the merchant can sell same content of data. The parameters as follows;

1 Key_gen: In the key generation protocol, the simplest way of understanding this is when the buyer generate pair of values (SK_B, PK_B) that serve as keys secret and public respectively, while transmitting PK_B to several merchant and others such as third party (certification authorities)

2 Fing: The evolvement of two parties between a merchant, M and buyer, B. The merchant inputs the content of the data to be broadcasted *(pic)*, the identity of the buyer, represented by the buyer's key PK_B and string text E *text_space*. Expectedly the algorithm may depend on the history described by $Record\_list_{pic}$ which serves as a record from the previous transaction of the same content of data; The buyer input text with her secret key, SK_B while the output to the buyer are the content *data* Pic_{bought} with small error. The buyer may also obtain the records ($Record_B$) to be presented against future disputes or instead she may acquire a certain specified output failed, that means the protocol fail; The output from the merchant is a record of the sales, $Record_M$ or fail.

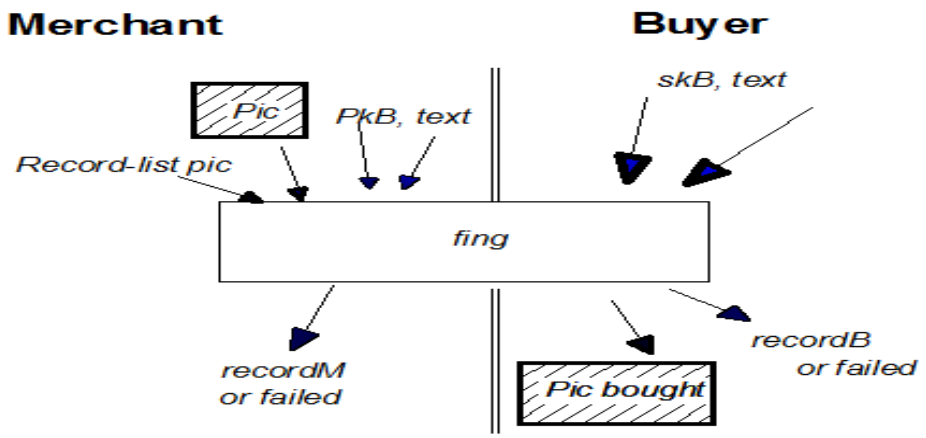

Figure 4.2 Parameters of Asymmetric fingerprinting

3 Identity: Is the algorithm; where the merchant process to uniquely identify the actual buyer to a certain specific content of data copy. Actually the inputs are attached with small error, Pic_{found}, the content merchant sold, *pic* and list of records, $Record\_list_{pic}$ of all records. E expectedly the output is able to identify the identity of the buyer by PK_B , text and a string Proof.

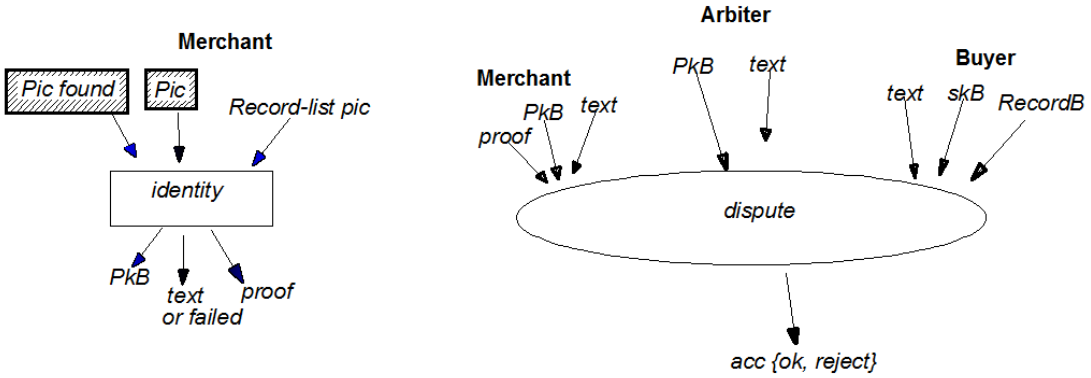

Figure 4.3 original owner identification and dispute

4 Disputes: The protocols between the merchant, an arbitrary (third party), which is referred as arbiter, and the culprit buyer. The merchant and the third party (arbiter) input PK_B and text. The merchant add input Proof in any circumstance the accused buyer involved, she input text, her secret key, SK_B and Record, the previous record from the transaction (purchase) with text subsequently the outputting is a Boolean value acc for the arbiter, that acc =ok, which refers that the arbiter finds the buyer guilty or that the arbiter accept that the merchant has found the content of the data bought by the accused buyer with text.

Security

In the asymmetric fingerprinting scheme the precautions with the regards to the asymmetric fingerprinting is the effectiveness in the sense that buyers acquired significant content of the data so long no body compromise it. Secondly, the merchant should be protected from cheating buyers, in as much the turning up of compromised copy of the content realised, the authenticating authority will have some buyer to be identified and hold guilty responsible. And also the buyers should be protected from cheating merchant and other buyers, while if they are innocent in terms of redistribution of their acquired copy, they should be falsely guilty.

It is noticeably, the buyer (receiver) is obtaining the copy, depending on the type of data content copy fingerprinted. We assumed that an arbitrary given relation similar1; similar1 (Pic, Pic_{bought}) which *means* Pic_{bought} is significant enough as a replacement for Pic which will be understood as pictures, similar1 referred to $Pic_{bought} = Pic + E_{bought}$ with error vector of minor small hamming weight, similarly from the other part we assumed arbitrary relation similar2 which means similar (Pic,

Pic_{found}) means compromised copy of content *Pic*, Pic_{found} is nearly to *pic* that the merchant supposed to rectify the identity of the original receiver (buyer).

Definition 2: The effectiveness of fingerprinting code scheme to a certain binary relation such that similar1 on *Pic_space*, if it holds; when the *fing* is executed honestly by both parties, then for *Pic E Pic_space*, with key generated, and history then the protocol succeeds while the buyer acquires a new enough variant Pic_{bought} of *pic* which is similar1 (*Pic*, Pic_{bought}) evolved.

Definition 3: In the secure for the merchant under *coll_size* collusion and binary relation similar2 such that the *coll_size* is a function $N \longrightarrow N$, whenever it holds for all the probabilistic polynomial-time interactive algorithm B. Whenever the merchant select content data *Pic* from *Pic_space* and execute *fing* on the data with B at most *coll_size* (K) times, and B will select the merchant's input text arbitrary and PK_b among the key generated and dispute with B and interaction with others (buyers)might exist in between;

- If B then output data Pic_{found} similar to pic which in the sense that similar2 (*Pic*, Pic_{found}) holds.
- Then identify o input Pic_{found} , *Pic* and the list current list of records, output (PK_B, text, Proof) where PK_B is the input to *fing* from the previous ones.
- If the dispute is executed expectedly evolving these parameters input by the merchant and arbiter and 3-party with B, the arbiter's output will be ok except with negligible probability.

Definition 4: In the scheme for the buyer on the fingerprinting scheme, if the following holds; no probability polynomial time algorithm M;

- Conveying out *fing* with B arbitrary pair of key generated by B and transmitted PK_B , where *M* can select the buyers input text arbitrary for the identical data or non identical and likely conveying out some disputes across.
- Acquiring output Pic_{input} in an executions, that M may select in an adaptive phases by inputting matching text.
- Can compute parameters text and proof, where in which the text is not in the mist of those for which acquired Pic_{bought} in the previous stages.

- Then the execution of disputes with an arbiter have input PK_B and text of the arbiter is OK with the exception of negligibility in probability.

In these *fing* execution, there could be a parallel or sequential approach and may have several flavour of requirement to its applications, like securing the merchant from making a false accusation and allowing merchant to deliver significant information in disputes.

4.4 General Constructions

Considering in a general symmetric fingerprinting schemes and other cryptographic primitives are secure participation in disputes by the buyer is not needed or store records. The algorithm in this general scheme of symmetric (fing, identify), a content data space *Pic_space* with a given probability distribution and identity space, *id_space*. And these algorithm are assumed to be carryout by the merchant.

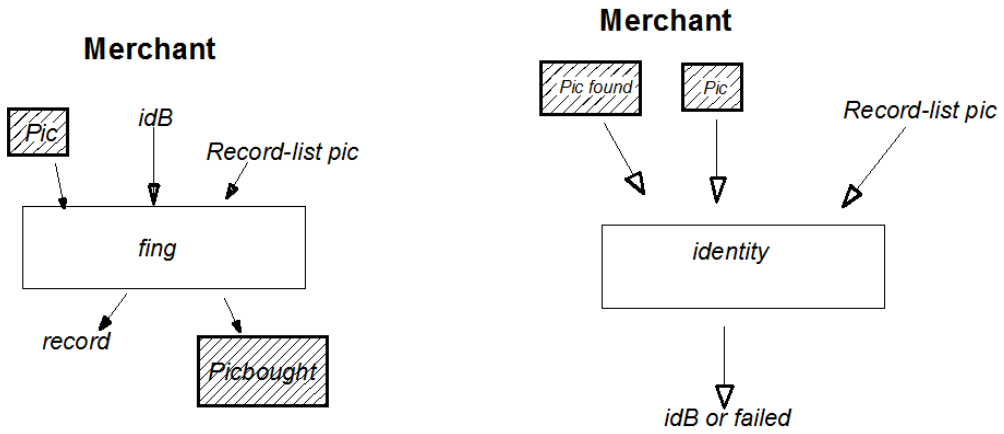

Figure 4.4 Model symmetric Fingerprinting

Definition 5: In construct of symmetric fingerprinting code scheme consist;

- Memory: less the *fing* algorithm does not provide record depending on *the* id_b ,the data content of *Record_list* will yield from *pic* and the random bits string by the merchant uniquely, it is denoted by its $Record_{pic}$ which is the secret turple of mark position in the data content (pic).

- Adequate substantial identity space *id_space* is $\{0, 1\}^{id\_len\ (k)}$ that is the function *id_len* of the security parameter *K*, $id\_len(k) \geq K^{\varepsilon}$ where K > 0 serve.

Construction

Let the symmetric fingerprinting scheme has 2-party protocol, scheme of signature with algorithm sign and verify, one-way function (*f*), *sym* to serve as symmetric, *asym* as asymmetric scheme. Key_gen_{asym} is the generated key signature scheme, $fing_{asym}$ refers to the first buyer, selection of $Record_{pic\ ,asym}$ by the merchant base on content (pic) and reserve (store) to serve as part of $Record\_list_{pic,\ asym}$. For any firstly buy (buyer) choose an identity $id_{sym} \in id\_space_{sym}$ randomly, so also the transaction entities (buyer and the merchant) execute a secure protocol evolution of two party scheme protocols.

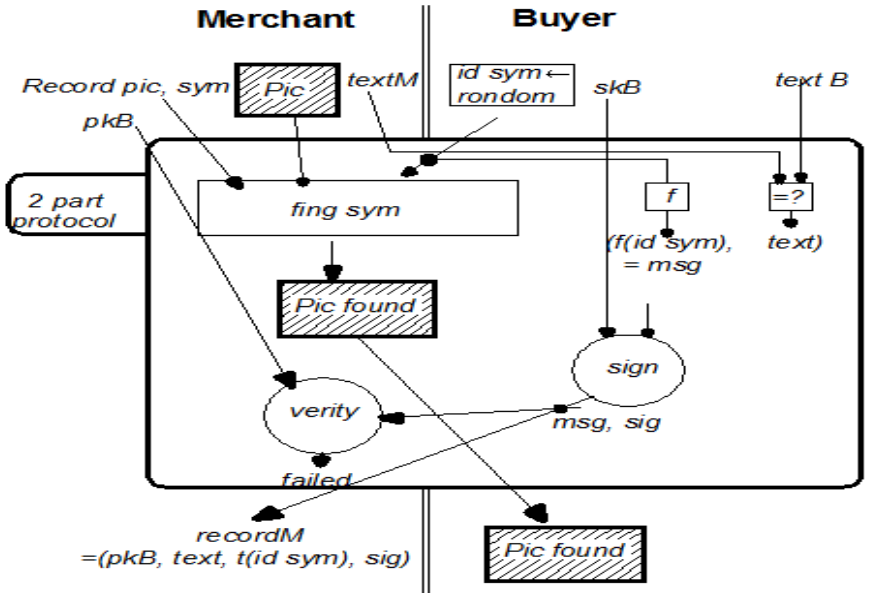

Figure 4.5 asymmetric construction out of symmetric scheme

The identification; $identify_{asym}$ input pic_{found} ,*pic*, and $Record\_list_{pic,asym}$;

1- Execute $identify_{sym}$ $(pic_{found}, pic, Record_{pic,sym})$, when the outputting failed, it stop, else obtain a value id_{sym}

2- It process f(id_{sym}) and look to a record with such function identity $f(id_{sym}) = f(id_{sym})$
3- If the record rectified with outputting yielding the values PK_B and text from the record proof(id_{sym} sig)

In the disputes, the input (PK_B $text$) which the arbiter has and proof (id_{sym}, sig), the sig is valid in regards to PK_B on $Msg = (f(id_{sym}), text)$.

CHAPTER 5

5.1 Conclusion

Application of fingerprinting code to the contents will facilitate the deterrence of involving the users from participating in the illegal activities of copying and redistributions as we have seeing. And we came to notice this type of scheme (fingerprinting code) provide a solution to numerous problem especially right to ownership, which allow the publisher to have a sole proprietorship of the content data published, and at same time do away with fear of losing a lot of resources (capital) that may cause, if their production fall in the handwork of traitors or even pirate users.

We will noticed that some of these techniques; Watermarking, Fingerprinting, stenography are all stipulated on just inserting the hidden information, mainly for intended purposes. Because there are number of issues specifically covered with production and communication of digital contents, through a media. The watermark embedded the information that identify the ownership/copyright details in the content (object), while fingerprinting code hide identification, it recognizes the user. These terms are going to be used interchangeably (fingerprint and watermarking). The motive behind improved fingerprint code sacrifices to any modification to the object (content) that may jeopardize the fingerprint shall result to the object useless. And the robust nature of fingerprint code technique needs its attached [23]

And we have seeing how the checkmate items help us released a good property of fingerprinting code, for the purposes intended, such as in the process of key distributions, watermarking, pirate rebroadcast, improving round complexity. All these enjoy the numerous capability and scalability of acquired by the fingerprinting code technique (scheme) in different settings. In the contained elements of fingerprinting form all mentioned above areas, information is invisible to the information and these hidden information should be robust against any form of

attacks and when collusion of various user to yield a new form of product(remember the traitor tracing scenario). And ultimate result will be possible to find atleast one of the users in the collusion team. And capability of asymmetric scheme in making hindrance to bad transaction activities realised and subsequently the guilty is found.

REFERENCES

[1] Chor, B., Fiat, A., Naor, M., and Pinkas, B. (2000). Tracing traitors. IEEE Transactions on Information Theory, 46(3):893–910.

[2] Kiayias, A., and Pehlivanoglu S. (2009). Tracing and revoking pirate rebroadcasts. In International Conference on Applied Cryptography and Network Security, pages 253–271. Springer.

[3] Kiayias, A., and Pehlivanoglu S. (2010). Improving the round complexity of traitor tracing schemes. In International Conference on Applied Cryptography and Network Security, pages 273–290. Springer.

[4] Schaathun, H.,G. The boneh-shaw (2006). Fingerprinting scheme is better than we thought. IEEE Transactions on Information Forensics and Security, 1(2):248–255.

[5] Petitocolas, F., A. (1999). rj,"evaluation of copyright marking systems". In Proceedings of the IEEE International conference 'in Multimedia Computing and Systems(ICMCS'99). Florence, Italy, **volume 1**, pages 574–579.

[6] Ettinger. J., M. (1998). Steganalysis and game equilibria. In International Workshop Information Hiding, pages 319–328. Springer.

[7] Duffany, J. L and Velez M., D. (2012). Steganography and steganalysis in digital images. Technical report, DTIC Document.

[8] Cachin, C. (1998). An information-theoretic model for steganography. In International Workshop on Information Hiding, pages 306–318. Springer.

[9] Hopper, N. J., Langford, J. and Ahn, L. V. (2002). Provably secure steganography. In Annual International Cryptology Conference, pages 77–92. Springer.

[10] Lai, J. J. (2008). Image watermarking techniques. Digital Signal Processing 18.5, 762-776.

[11] Bruekers, A. A. Geert, Depovere, F. G., Nuijten, P. A. and Oomen, A. W. (2000). Embedding supplemental data in an encoded signal, such as audio/video watermarks, US Patent 6,157,330.

[12] Kuribayashi, M. and Tanaka, H. (2005). Fingerprinting protocol for images based on additive homomorphic property. IEEE Transactions on Image Processing, 14(12):2129–2139.

[13] Cox, I. J., Miller, M. L., Bloom, J. A. and Honsinger, C. (2002). Digital watermarking, volume 1558607145. Springer.

[14] Katzenbeisser, S. and Petitcolas, F. (2000). Information hiding techniques for steganography and digital watermarking. Artech house.

[15] Boneh, D. and Shaw, J. (1995). Collusion-secure fingerprinting for digital data. In Annual International Cryptology Conference, pages 452–465. Springer.

[16] Kurosawa, K. and Yoshida, T. (2002). Linear code implies public-key traitor tracing. In International Workshop on Public Key Cryptography, pages 172–187. Springer.

[17] Praun, E., Hoppe, H. And Finkelstein, A. (1999). Robust mesh watermarking. In Proceedings of the 26th annual conference on Computer graphics and interactive techniques, pages 49–56. ACM Press/Addison-Wesley Publishing Co.

[18] Delerablée, C. (2007). Identity-based broadcast encryption with constant size cipher- texts and private keys. In International Conference on the Theory and Application of Cryptology and Information Security, pages 200–215. Springer.

[19] Ding, Y. and Fan, L. (2011). Traitor tracing and revocation mechanisms with privacy- preserving. In Computational Intelligence and Security (CIS), 2011 Seventh Inter- national Conference on, pages 842–846. IEEE, 2011.

[20] Yacobi, Y. (2001). Improved boneh-shaw content fingerprinting. In Cryptographers' Track at the RSA Conference, pages 378–391. Springer.

[21] Naor, M. and Pinkas, B. (1998). Threshold traitor tracing. In Annual International Cryptology Conference, pages 502–517. Springer.

[22] Sharifnia, S. (2013). Analysis total and better the boneh–shaw fingerprinting codes. Journal of Computations & Modelling, 3(3):57–73.

[23] Kiayias, A. and Pehlivanoglu, S. (2010). Encryption for digital content, **volume 52**. Springer Science & Business Media.

[24] Lei, C., Yu, P., Tsai, P. & Chan, M. (2004). An efficient and anonymous buyer-seller watermarking protocol, IEEE Trans. Image Process. **13**(12): 1618–1626.

[25] Memon, N. & Wong, P. W. (2001). A buyer-seller watermarking protocol, IEEE Trans. Image Process. **10**(4): 643–649.

[26] Pfitzmann, B. & Sadeghi, A. (1999). Coin-based anonymous fingerprinting, Advances in Cryptology– EUROCRYPT'99, Vol. 1592 of LNCS, Springer-Verlag, pp. 150–164.

[27] Pfitzmann, B. & Sadeghi, A. (2000). Anonymous fingerprinting with direct non-repudiation, Advances in Cryptology – ASIACRYPT'2000, Vol. 1976 of LNCS, SpringerVerlag, pp. 401–414.

[28] Amiri, E., Tardos, G. (2009). High rate fingerprinting and fingerprinting capacity, In Proceedings of the Nineteenth Annual ACM -SIAM Symposium on Discrete Algorithms, pages 336-345.

[29] Anderson, R. J. ed. (1996). Information Hiding: First International Workshop, Lecture Notes in Computer Science, **volume. 1174**.

[30] Anthapadmanabhan, N. P., Barg, A. (2006). Random Binary Fingerprinting Codes for Arbitrarily Sized Coalitions, Proceedings of IEEE International Symposium on Information Theory, ISIT 2006, pages 351-355.

[31] Anthapadmanabhan, N. P., Barg, A., Dumer, I. (2007). Fingerprinting capacity under the marking assumption. Submitted to IEEE Transaction on Information Theory - Special Issue on Information-theoretic Security. Available from

arXiv:cs/0612073v2. Preliminary version appeared in the Proceedings of the 2007 IEEE International Symposium on Information Theory, (ISIT 2007), 2007.

[32] Blayer, O. Tassa, T. (2008). Improved versions of Tardos' fingerprinting scheme.

[33] Dumer, I. (2009). Equal-Weight Fingerprinting Codes, Second International workshop on codeing and cryptology, Lecture Notes in Computer Science, **Volume 5557**, pages 43-51.

[34] Johnson, N. F., Katzenbeisser, S. (2000). A Survey of Steganographic Techniques, Chapter 3

[35] Katzenbeisser, S. (ed.), Petitcolas, F.A.P. (ed.) (2000). Information Hiding Techniques for Steganography and Digital Watermarking, Artech House Books.

[36] Skoric, B. Katzenbeisser, B. Celik, M. U. (2808). Symmetric Tardos fingerprinting codes for arbitrary alphabet sizes. To appear in Designs, Codes and Cryptography.

[37] Tardos, G. (2003). Optimal probabilistic fingerprint codes, Journal of the ACM, to appear. Preliminary version appeared in Proceedings of the 35th Annual ACM Symposium on Theory of Computing, (STOC 2003), 116125.

[38] Skoric, B.,Vladimirova, T.U., Celik, M., Talstra, J. C (2006). Tardos fingerprinting is better than we thought. Available from arXiv:cs/0607131

[39] Huang, Y.W., Moulin, P. (2009). Saddle Point Solution of the Fingerprinting Capacity Game Under the Marking Assumption, in 2009 IEEE International Symposium on Information Theory (ISIT 2009), Seoul, Korea.

[40] Schaathun, H. G. (2003). Fighting Two Pirates, In International Symposium on Applied Algebraic Algorithms and Error Correcting Codes, Also in Lecture Notes in Computer Science **Volume 2643**, pages 71-78.