ISSN 2229-5518

Novel Algorithm & methods of document security and verification system

Mukesh Kumar, M.Phil, MCA (Ph.D Scholar), Dr. R.K.Sharma, M.Tech, Ph.D

Abstract— Before discuss any more on the topic it is necessary to know about security in actual; it is the degree of protection against risk, hazard, threat, danger, damage, loss, misdeed and crime etc. We always compare Security with safety and reliability but the major difference between security and reliability is that security must take in to account of actions of people attempting to cause destruction. Securities as a form of protections which are structures and process that provide or improve security as a condition. In my opinion the conditions that prevent from hazards and provide guards from unauthorized or unauthenticated happenings or access to secrete and legal documents. We present the conceptualization, design and implementation of facts or transcript, an encrypting file system incorporates an advanced key management schemes, methods, algorithms and technique to provide a high grade of security while remaining will be fully transparent and easily usable.

Documents that contain sensitive or important information which are often subject to fraud, fake and tampering. One attempt toward security of documents has been through the application of watermarks on the printed documents. The watermarks are physical designs imprinted / embossed or stamped into the document that can be seen when the document is held up to a light. Visible watermarks are typically faded background images superimposed on the document. The conventional use of watermarks is typically insufficient to prevent the documents from succumbing to fraud, imitation, simulation, reproduction, counterfeiting or other tampering. It would therefore be beneficial to securely protect documents from further deters fraud, counterfeiting, or other tampering etc. The Invisible watermarks are typically pattern less arrangements of hidden bits in the documents.

Use of Invisible ink for document Security: Invisible ink is a substance used for writing, which is invisible either on application or soon thereafter, and which later on can be made visible by some means. Invisible ink is one form of steganography, and it has been used in intelligence uses include anti-counterfeiting, property marking, hand stamping for readmission, children's games, and marking for the purpose of identification in manufacturing. An Ultraviolet (UV) marker is a pen marks are fluorescent (Fluorescence is the emission of light by a substance that has absorbed light) but transparent: the marks can be seen only under an ultraviolet light. They are commonly used in security situations to identify belongings or to protect from the copying of banknotes. UV pens can now be bought at some stationery shops to security mark items of high value in case of theft

Increasing thefts of sensitive data is a world wide issue for all individuals as well as organizations, an enterprise encrypting file system must take a cohesive approach towards solving the matters associated with data security in organizations. These include flexibility for multi-user scenarios, transparent remote access of shared file system and protect against threats including insider attacks while trusting the fewest number of entities. We also formalize a general method and algorithms to secure documents by SDES (Simplified Data Encryption Standards), AES (Advance Encryption Standards) technique and by using fool proof methods in the current technological and advance scenario to defend data from attacks, intruders, crackers and unauthorized and unauthentic users.

Digital Watermarking is a technique to insert watermark signal into image in order to authenticate it. A binary watermark pattern was constructed from the information content of the image by selecting the minimum value from every block of size 2”x2”. As a solution to this issue, an innovative watermarking scheme is proposed. According to this, the low frequency sub-band of wavelet is utilized in the watermark construction process. The operation of embedding and extraction of watermark is done in high frequency domain of Discrete Wavelet Transform since small modifications in this domain are not supposed to see by human eyes. This watermarking scheme deals with the extraction of the watermark information in the absence of original image, hence the blind scheme was obtained. Peak Signal to Noise Ratio (PSNR) and Similarity Ratio (SR) are computed to measure image quality.

The ability to secure the protected data is differentiating by the algorithms efficiency to protect data against duplicity, hazards or attacks. This synopsis provides a performance comparison between most common encryption algorithms DES, 3DES and AES.

Index Terms— Security of document, Check authenticity of document, currency verification, document security, methods to detect fackeness of document, document verification, Algorithem to check document genuineness

1 INTRODUCTION

In recent years, the transmission of digital images over the Internet or personal digital mobile phones has been highly developed. Secure storage and transmission of digital images are becoming critically important. Most traditional ciphers, such as DES, and AES are not suitable to conduct the digital image encryption in real time due to large data volume involved. By using the Data Encryption Algorithm (DEA), SDES (Simplified data encryption standards) algorithm, Steganography methods of Hiding Data with in Data, and simplified advance encryption standard and Neuro fuzzy to explicit handwriting reorganization by Biometrics technology by using non-traditional methods / theory.

The need for data security emerges from the widespread deployment of shared file systems, greater mobility of computers and the rapid smallness of storage devices. It is increasingly obvious that the value of data is much more than the value of the underlying devices. The theft of a personal laptop or a USB thumb drive leaves the victim vulnerable to the risk of identify theft in addition to the loss of personal or financial data and intellectual property. Several recent incidents of data theft emphasize the need for a cohesive solution to the problem of storage security. Hence, it is fast becoming necessary to protect stored data from unauthorized access using strong cryptographic methods.

Encrypting File Systems

An encrypting file system employs secure and efficient mechanisms to encrypt or decrypt data on-the-fly as it is being written to or read from the underlying disk, to provide a level of data privacy that goes beyond simple access control. Also, issues such as trust models, backups and data recovery must be resolved. An encrypting file system must also be tightly integrated with the operating system for ease of use and flexibility. Other challenges faced when designing a storage security framework include immunity from attacks launched by privileged entities, enabling legitimate remote access to shared encrypted volumes and providing a scalable and transparent key management scheme suitable for enterprise deployment. Although the design of such systems is a well-researched problem, existing implementations still lack the security and usability features that must be present in a truly scalable system that can be successfully deployed in enterprises.

2 PROBLEM DESCRIPTION

With the advent of technique and technologies, it is easier for a person to create a fake copy of the legal document very easily. High resolution scanner and printer made people capable of preparing almost perfect copy of the document. So, we need a technique by which we can verify the truth ness of the document image. To propose a technique, this can verify the truth ness of informatory financial or legal documents. To verify the document, technique should be capable of digital, invisible / hidden watermark or encryption methods during transmission a document from one source to another destination and verify on the basis of methods, algorithms etc.

In the consumer goods industry, counterfeiting is a significant and growing problem. While fashion and luxury goods have long been targets of counterfeiters, nearly any branded product can be the subject of counterfeiting. For example, products such as cosmetics, automotive parts etc. etc. subjects of counterfeiting. Counterfeiting is difficult to detect, investigate, and quantify. Consequently, it is difficult to know the full extent of the problem. However, by some estimates, between five to seven percent of all world trade is in counterfeit goods, amounting to an annual value that exceeds 2500 crores. It is necessary to stop these vulnerabilities to detect and prevent counterfeiting.

Hand written signature is the most widely form of personal identification as well as for document verification, especially for cashing cheques, deeds etc However, for several reasons the task of verifying human signature can't be considered for pattern recognition because signature samples from the same person may be similar but not identical. A person's signature often changes radically during their life. We can see much variability in signature according to country, age, time, and psychological or mental state. So to remove all such vulnerabilities form the system to check the genuineness of the documents, we can use such methods that verify the legitimacy of the document that we use for social value purpose.

3 MOTIVATION

It is well known that the document fraud, such as credit card fraud, counterfeit, cracking in the fruitful information during transmission of e-data. Internet fraud and theft of crucial document and reproduction of fake document can be reduced in all.

3.1.1 Data Security Inspirations

An enterprise data protection system is vital in military organizations where classified and secret data need to be shared and secured simultaneously. Recent news reports of security breaches and data thefts from India's military and intelligence agencies accentuate the critical need for a cryptographic solution to this problem. According to reports relating to one case, important information was leaked through stolen USB thumb drives. Another case purportedly involved a computer administrator who was able to pass secret data illegitimately to a foreign country. The fact that both these cases involved insiders motivates the need for a secure data protection mechanism that thwarts theft attempts and ensures that undue power is not left in the hands of individual employees or administrators. Data protection systems are increasingly playing a crucial

role in commercial environments too. A recent study conducted by Symantec Corporation surveyed laptop users across Europe, the Middle East and Africa to estimate the value of the commercially sensitive contents of their laptops. The study estimated the average worth of a single laptop to be about a million dollars. Clearly, there is a pressing need to design and develop secure and usable data protection mechanisms that cater to the above application scenarios. Encrypting file systems fill this void to enable individuals and organizations to keep their storage systems highly available and protected from unauthorized access at the same time.

3.1.2 The main aim of my project is

High Level of Security

Completely specified and easy to understand

Cryptographic security do not depend on algorithm secrecy

Adaptable to the diversified applications

Economical implementation on hardware / software

Efficient on high data transfer rate

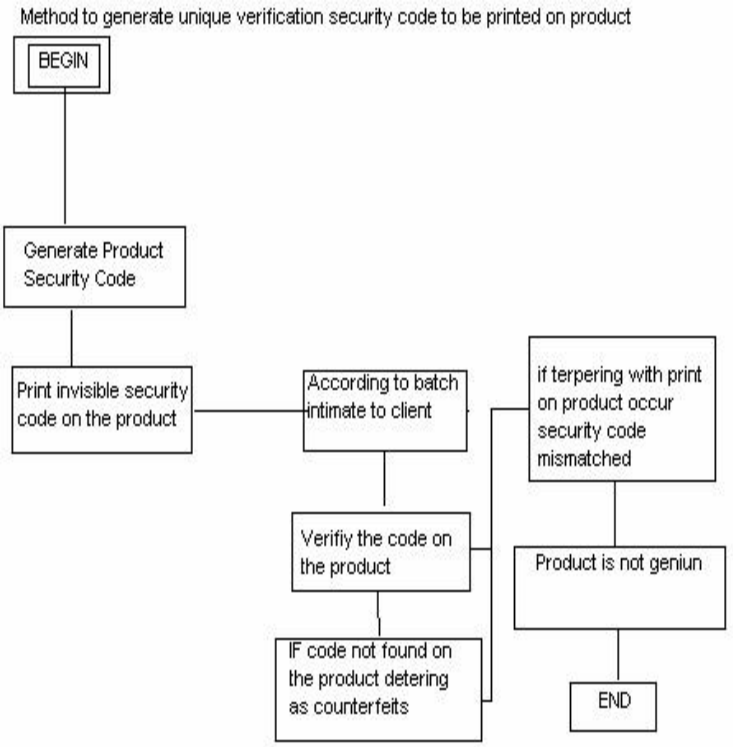
4 METHOD OF VERIFICATION

4.1 Method and system for deterring product counterfeiting

A method and system for authenticating goods and thereby detecting and deterring counterfeits are disclosed. According to one aspect of the invention, a client utilizes data received from a host to generate a verification of security codes and to direct a printing device to print the majority of security codes on a verification of products, without retaining the verification of security codes after the printing device has printed the security codes on the security of products. After the security codes have been printed, a person can communicate the security code to the host, which can verify its authenticity.

In one personification of the invention, the security codes may be printed on a tamper-evident seal. Accordingly, the tamper-evident seal may be positioned on the product in such a way that the tamper-evident seal is destroyed when the product is opened, or otherwise used. Consequently, once destroyed, the security codes cannot be reused.

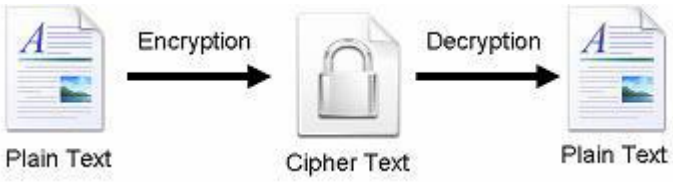
Illustrate a method, according to one personification of the invention, for generating a verification of unique security codes to be printed on products;



1. Steganography: Hiding Data With in Data

4.2 Cryptography

The science of writing in secret codes — addresses all of the elements necessary for secure communication over an insecure channel, namely privacy, confidentiality, key exchange, authentication, and non-repudiation. But cryptography does not always provide safe communication.



Cryptography Goals:-

Every security system must provide security functions that can assure the secrecy of the system. These functions are usually referred to as the goals of the security system.

Authenticity / Legitimacy: - This means that before sending and receiving data using the system, the receiver and sender identity should be verified.

Secrecy /Confidentiality: - Usually this function (feature) is how most people identify a secure system. It means that only the authenticated people are able to interpret the message (date) content and no one else.

Integrity / Reliability: - Integrity means that the content of the

communicated data is assured to be free from any type of modification between the end points (sender and receiver). The basic form of integrity is packet check sum in IPv4 packets.

Agreement / Non-denial of Services:- This function implies that neither the sender nor the receiver can falsely deny that they have sent a certain message.

Easy to use services: Since secure systems usually get attacked by intruders, which may affect their availability and type of service to their users. Such systems should provide a way to grant their users the quality of service they expect.
Now consider an environment where the very use of encrypted messages causes suspicion. If a Internet service provider (ISP) is looking for encrypted messages, they can easily find them. Consider the following text file; what else is it likely to be if not encrypted?

qANQ R1DBwU 4D/TlT6 8XXuiUQC ADfj2o fdklklglf 4aFY BcWu mA7hR1Wvz9rbv2BR6WbEUsy ZBIEFtjyqCd9 6qF38s p9IQiJIK lNaZfx2GLRWikPZw chUXxB+AA5+lqsG/ELBvRa c9XefaYpbbAZ6z6LkOQ+eE0XASe7aEEPfdxvZZT37dVyiy xu BBRYNLN8Bphdr2zvz/9Ak4/OLnLiJR k05/2UNE5Z0a+ 3lcv ITMmfGajvRh kXqocavPOKiin3hv7+Vx88uLLem2/=

The message above is a sentence in English that is encrypted using Pretty Good Privacy (PGP), probably the most commonly used e-mail encryption software today.

4.3 Steganography

Steganography is the science of hiding information. Whereas the goal of cryptography is to make data unreadable by a third party, the goal of steganography is to hide the data from a third party. There are a large number of steganographic methods that most of us are familiar with invisible ink and microdots to secreting a hidden message in the second letter of each word of a large body of text and spread spectrum radio communication. With computers and networks, there are many other ways of hiding information, such as:
Covert channels (e.g. some distributed denial-of-service tools use the ICMP (Internet Control Message Protocol) as the communications channel and a compromised system)
Hidden text within Web pages.

Steganography today, however, is significantly more sophisticated than the examples above suggest, allowing a user to hide large amounts of information within image and audio files. These forms of steganography often are used in conjunction with cryptography so that the information is doubly protected; first it is encrypted and then hidden so that an adversary has to first find the information and then decrypt it.

Here are a number of uses for steganography besides the simple innovations. One of the most widely used applications is for so-called digital watermarking. A watermark, historically, is the replication of an image, logo, or text on paper stock so that the source of the document can be at least partially authenticated. A digital watermark can accomplish the same function; a graphic artist, for example, might post sample images on her Web site complete with an embedded signature so that ownership can be proved in all.

4.3.1 Steganography method

The following formula provides a very generic description of the pieces of the steganographic process:

cover_medium + hidden_data + stego_key = stego_medium

4.4 Neuro fuzzy to explicit handwriting recognition

The neuro-fuzzy system that exploits explicit knowledge on character's shape and execution plans was presented for online handwriting recognition. The first stage extracts prototypes using the Fuzzy ARTMAP based extraction method. These prototypes provide the explicit knowledge about shapes and execution plans and are used to initialize the second stage of the recognizer consisting of a series of LVQ (Learning Vector Quantization) codebooks. It has been shown that the explicit knowledge extracted by the first stage improves the rates of the handwriting recognizer. The comparison of our system's performance with other relevant recognizers showed interesting results that may foster further improvements.

4.5 Detection Techniques for Identification of Duplicate and Near Duplicate Documents

An array of 400 million web pages obtained from search engines were employed to apply the technique. These pages were grouped into clusters of incredibly similar documents. They identified that in their dataset almost one third of the pages were near duplicates of other pages. The grainy hash vectors (GHV) have the ability to detect near-duplicates and exact duplicates. They have mathematical properties that are well-defined. They conducted experiments and demonstrated that GHVs identify duplicate and near-duplicate document pairs at merge time efficiently and effectively. The experiments illustrated that the proposed methods detected clones among static Web pages and the efficiency of the method was proved by a manual verification. The method produced comparable results, but different computational costs were involved. In spite of the fact that the overall process is satisfactorily efficient with computer resources, practically, human attention to consider the many results is the bottleneck. In conclusion, a special handling for exact duplicates and a way to reduce a frequent source of false alarms-template similarity is provided.

4.6 Simplified Data Encryption Standards

S-DES encryption (decryption) algorithm takes 8-bit block of plaintext (cipher text) and a 10-bit key, and produces 8-bit cipher-text (plaintext) block. Encryption algorithm involves 5 functions: an initial permutation (IP); a complex function f K, which involves both permutation and substitution and depends on a key input; a simple permutation function that switches (SW) have the 2 bisect of the data; the function f K again; and finally, a permutation function that is the inverse of the initial permutation (IP-1). Decryption process is similar.

The function f K takes 8-bit key which is obtained from the 10-bit initial one two times. The key is first subjected to a permutation P10. Then a shift operation is performed. The output of the shift operation then passes through a permutation function that produces an 8-bit output (P8) for the first sub key (K1). The output of the shift operation also feeds into another shift and another instance of P8 to produce the 2nd sub key K2. We can express encryption algorithm as:

Cipher-text= IP-1 (((((int))))) 2 1 f SW f IP plain text K K

Where 8((10( ))) 1 K Shift P key
8(((10( )))) 2 K Shift Shift P key

Decryption is the reverse of encryption:
Plain text= IP-1 ((((( ))))) 1 2 f SW f IP cipher-text K K

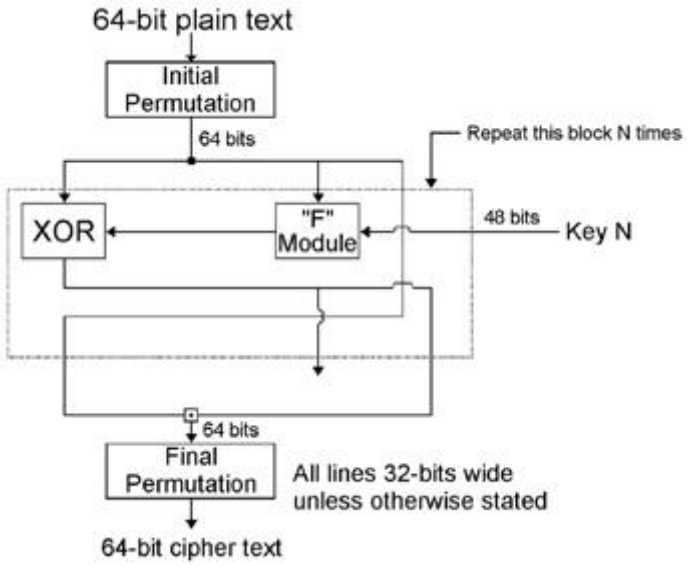
4.7 Data Encryption Algorithm

DES was designed with an effective key length of 56 bits, which is vulnerable to comprehensive search. It also has some weaknesses against differential and linear cryptanalysis: these allow recovering the key using, respectively, 247 chosen plaintexts, or 243 known plaintexts. A known plaintext is an encrypted block (an 8-byte block, for DES) for which the attacker knows the corresponding decrypted block.

A chosen plaintext is a kind of known plaintext where the attacker gets to choose himself the decrypted block. In practical attack conditions, such huge amounts of known or chosen plaintexts cannot really be obtained, hence differential and linear cryptanalysis do not really impact the actual security of DES; the weakest point is the short key. Still, the existence of those attacks, which, from an academic point of view, have less complexity than the exhaustive key search (which uses 255 invocations on average), is perceived as a lack in security.

The structural weaknesses of DES are thus its key size, and its short block size: with n-bit blocks, some encryption modes begin to have trouble when 2n/2 blocks are encrypted with the same key. For the 64-bit DES blocks, this occurs after encrypting 32 gigabytes worth of data, a big but not huge number.

A variant on DES is called 3DES: that's, more or less, three DES instances in a row. This solves the key size issue: a 3DES key consists in 168 bits (nominally 192 bits, out of which 24 bits are supposed to serve as parity check, but are in practice wholly ignored), and exhaustive search on a 168-bit key is wholly out of reach of human technology. From (again) an academic point of view, there is an attack with cost 2112 on 3DES, which is not feasible either. Differential and linear cryptanalysis are defeated by 3DES (their complexity rises quite a bit with the number of rounds, and 3DES represents 48 rounds vs 16 rounds for the plain DES). Yet 3DES still suffers from the block size issues of DES. Also, it is quite slow (DES was meant for hardware implementations, not software, and 3DES is even three times slower than DES).



Advanced Encryption Standard (AES) algorithm

The resistance of AES towards differential and linear cryptanalysis comes from a better "flooding effect" (a bit flip at some point quickly propagates to the complete internal state) and specially crafted, bigger S-boxes (a S-box is a small lookup table used within the algorithm, and is an easy way to add non-linearity; in AES, S-boxes have 8-bit inputs and 8-bit outputs).

On a theoretical point of view, we could say that what makes a cryptographic primitive secure is the amount of effort invested in its design. At least, that effort is what creates the perception of security: when I use a cryptosystem, I want it to be secure, but I also want to be certain that it is secure (I want to sleep at night). The public design and analysis process helps quite a lot in building that trust.

The validation tests are designed to assist in the detection of accidental implementation errors, and are not designed to detect intentional/ deliberate attempts to misrepresent conformance. Thus, validation should not be interpreted as an

evaluation or approval of overall product security. The design philosophy of AES communicates data via REQUEST and RESPONSE files.

For each mode implemented, selections are available for the key sizes (i.e., 128-bit, 192- bit, and 256-bit) supported as well as the ciphering direction (i.e., encryption and decryption). It is not necessary for every mode implemented to support the same key sizes and ciphering directions. Once configuration information has been provided, appropriate REQUEST files will be generated. REQUEST files are the means by which test data is communicated to the Implementation Under Test (IUT). To initiate the validation process of the AES, a vendor submits an application to an accredited laboratory requesting the validation of their implementation of the algorithm. More specifically, the request for validation should include:

1. Vendor Name
2: Product Name
3: Product Version;
4. Implementation in software, firmware, or hardware
5. Processor and O.S

The following is a sample data set for Multi block message test is:

KEY = 00000000000000000000000000000000
IV = 00000000000000000000000000000000
PLAINTEXT = 6a84867cd77e12ad07ea1be895c53fa3

The RESPONSE file for the tests contains the same data as the REQUEST file with the addition of the ciphertext for encryption (or plaintext for decryption). The following is a sample data set:

KEY = 00000000000000000000000000000000
IV = 00000000000000000000000000000000
PLAINTEXT = 6a84867cd77e12ad07ea1be895c53fa3
CIPHERTEXT = 732281c0a0aab8f7a54a0c67a0c45ecf

4.8 Signature verification / handwriting reorganization by Biometrics

Biometrics is an emerging technology that is quickly garnering wide attention for its promises to improve security violation and offer solutions to key challenges likes financial theft and fraud etc. Biometrics is used to verify identity by way of physiological or behavioral characteristics that are unique to each individual and cannot be forgotten, lost or stolen. It can take the form of several different techniques such as hand geometry, iris or retinal scans, dynamic signature verification, face and voice recognition, fingerprinting and others.

Biometric handwriting characteristics are absolutely unique to an individual and virtually impossible to duplicate. Therefore, handwriting still remains one of the most powerful human identifiers today. In dynamic signature verification, multiple biometric characteristics of a signature in question are analyze / scrutinized and compared against a reference signature kept on file to make a conclusion about measure the confidence value of the signature's genuineness. If several genuine reference signatures are available, the measure of the stability of the particular feature is developed and used to estimate the probability of deviations observed in the questionable signature.

The most advanced signature verification systems employ a powerful combination of engines using different approaches for comprehensive signature verification. Each engine analyzes biometric characteristics such as speed, acceleration, deceleration, stroke sequencing and length, pen pressure and timing information received directly during the act of signing together with other innovative technology that scrutinizes signature shape. Finally, the results received from different methods of analysis are combined to provide a reliable measure of the likelihood of coincidence between the signature in question and genuine reference signature(s). The success of dynamic signature verification in such systems relies on analysis of graphical representation of a signature and biometric characteristics received during the process of signing. Usage of several independent methods of analysis leads to a dramatic performance improvement and adds substantial robustness to the signature verification software.

Since there are many different characteristics involved in the analysis, the biometric signature verification technology is able to ensure a high efficiency of verification even if certain characteristics of signing (i.e. pressure) are not tracked. This important aspect of signature authentication reduces dependence on the type, specifics and quality of pen-enabled or pointing devices. For example, pressure characteristics are very important if the signature is obtained on pressure sensitive devices and less important if captured on pointing devices built into many laptop computers that are not sensitive to exerted pressure.

Technological advancements have increased the accuracy of biometrics systems while making them more widely affordable as a viable method of verification. Due to the level of reliability, convenience and high security that biometrics provides, it is already being used extensively in a number of applications to provide a competitive edge to alternative technologies. At the same time, not all biometric methods are equally acceptable in all industries and all applications. One of the greatest challenges of biometrics are privacy concerns due to its relatively intrusive nature. It is for this reason, for example, that fingerprinting and iris scans are not accepted in much retail, banking and financial services applications. Dynamic signature verification is a much more easily digested biometric method of identification. The act of signing one's name is socially accepted and commonplace in our legal and commercial lives. As such, individuals are less likely to object to their sig-

nature being confirmed as compared to other possible biometric analyses. This allows dynamic signature verification to be seamlessly integrated into existing working processes.

Not only is dynamic signature verification the least controversial of current biometric methods on the market, it is also one of the most accurate, intuitive, fast and cost effective and operates with compact data. All these factors make it an ideal solution for document authentication and enterprise workflow. Nowadays a wide range of equipment is available for digitizing signatures: palmtop or PDA-type devices, digitizer tablets, pointing devices, and smart phones. Biometric signature verification software universally supports any form of pen-enabled input device on which a signature is written.

The solution can supports any signature authentication application, from homeland security to banking and retail applications, providing organizations and individuals with enhanced security and control over the documents and transactions that are originated, transacted and stored in today's business environments. Based on state-of–the-art technology, signature verification software extracts maximum data concealed in a biometric signature, captured with a digitizer, and converts the data to information that allows a more reliable detection of forgery than any other solution available on the market, including manual verification.

5 FUTURE ASPECTS OF TECHNOLOGY

5.1 Potential of digital watermarking technology

The future of this technology is to detect any alterations and modifications in an image. Especially it is related to encourage copy protection of printed media. Examples here include the protection of bills, legal documents, currency with digital watermarks.

Digital watermarks can also be adapted to mark white paper with the goal of authenticating the originator, verify the authenticity of the document contents etc. For example, the digital watermark can be used to embed the name of the originator or important information such as key with monetary amounts. In the event of a dispute, the digital watermark is the key to read and certify authentication of key information in the contract. I proposed to process an invisible mark with normal and visible ink.

5.2 Cryptographic Technique to provide high level protection for digital document

The Classical cryptography offers a high level of protection for digital documents and is essential in ensuring an efficient and secure electronic communication. However, a big challenge has been securing the path from the human to the cryptographic module. This path is currently the weakest link, which limits the security of the whole cryptographic chain. An approach for bidirectional document authentication based on visual cryptography and watermarking is proposed. The proposed challenge-response mechanism prevents the man-in-the-middle attacker from obtaining a signed document without author's approval. Assuming that the author would not approve a forged document, the attacker is prevented from obtaining a valid signature on a forged document. On the user's side, the method requires a list of watermarks and transparencies. For each document page, a watermark and a transparency are needed. The watermarks, which appear only faded over the document, can in the list be printed reduced in size, so to fit on a sheet of paper. The method is not intended for to be used among arbitrary number of users and trusted devices. It essentially relies on symmetric cryptography, so the number of key sets increases linearly with the number of user per trusted device.

6 CONCLUSION

1. The occurrence of financial fraud has become a shocking trend in today's world. The significance of these crimes has resulted in a renewed interest in advanced security means. Institutions are increasingly demanding more reliable, less costly authentication and authorization for everyday activities, such as performing financial transactions, secure entry points etc.

2. AES encryption by using dynamic matrices based on m-bit additional secret key. We can enhance m bit security level to default AES with respect to exhaustive key search. We briefly discuss the effectiveness of current naive and enhanced methods of cryptanalysis against AES by introducing DMCT and present possible security enrichment for different methods of attack. We also achieve same level of processing cost as default AES encryption algorithm using current processors. AES become more secure with no effect on throughput. To enhance security of ciphers against developing cryptanalytic techniques and this area has research potential to a great extent.

3. The explosive growth of information sources available on the World Wide Web has necessitated the users to make use of automated tools to locate desired information resources and to follow and asses their usage patterns. Web contains duplicate pages and mirrored web pages in abundance. The efficient identification of duplicate and near duplicates is a vital issue that has arose from the escalating amount of data and the necessity to integrate data from diverse sources and needs to be addressed. In this synopsis, I have presented a comprehensive survey of up-to-date researches of Duplicate / near duplicate document detection both in general and web crawling. To understand the available methods and help to perform their research in further direction.

4. A two stage neuro-fuzzy system that exploits explicit knowledge on character's shape and execution plans was presented for on-line handwriting recognition. The first stage extracts prototypes using the Fuzzy ARTMAP based extraction method. These prototypes provide the explicit knowledge about shapes and execution plans found in training data and are used to initialize the second stage of the recognizer consisting of a series of LVQ (Learning Vector Quantization) codebooks.

We can use such method on which, we are able to check the truth ness of document whether it is clone copy or true copy by high resolution printing or scanning technique. So improvement can be made by making it resolution independent.

7 REFERENCES

[1] http://www.itl.nist.gov/fipspubs/fip180-1.htm.Invisible ink / digital markers as document markers

[2] Lin Zhang, Jianhua Wu, "Image Encryption with Discrete Fractional Cosine Transform and Chaos", the Fifth International Conference on Information Assurance and Security, China, pp. 61- 64, 2009.

[3] Ling jie, Li jizhong, "An Improved Security Technique for the Terminal Sensitive Documents",IJIPM, Vol. 2, No. 2, pp. 108-113, 2011.

[4] Biryukov, A., Khovratovich, D., Nikolic, I.: Distinguisher and related-key attack on the full AES-256. In CRYPTO'09, LNCS. Springer Verlag (2009).

[5] Murtaza, G., Ikram, N.: New Methods of Generating MDS Matrices. In: Proceedings of ICWC 2008, pp 129-133, ISBN: 978-983-44069, Kuala Lumpur, (2008).

[6] ZHAO Rui, WANG Qingsheng, WEN Huiping, "Design of AES algorithm Based on TwoDimensional Logistic and Chebyshev Chaotic Mapping", Microcomputer Information, vol.24,No.33, pp. 43-45, 2008.

[7] QIAO Zhiwei, Han Yan, WEI Xueye, "Design and Implementation of the File EncryptionAlgorithm Based on One Dimensional Chaos Method", Journal of Norh University ofChina(Natural Science Edition), Vol.28, No.6, pp. 517-520, 2007.

[8] Biham, E., Dunkelman, O., Keller, N.: Related-key impossible differential attacks on AES-192, In CT-RSA'06, LNCS, vol. 3860, pp. 2–-31, Springer Verlag (2006).

[9] Qijun Zhao, Hongtao Lu, ``A PCA-based WatermarkingScheme for Tamper-proof of Web Pages", Pattern Recognition, vol. 38, pp. 1321-1323, 2005.

[10] Wang Huafeng, Yao Weihong, Li Hong, "Application Design and Analysis of ChaoticCharacteristics in 1-Demesion Logistic Mapping", National Software Application Conference, pp.486-490, 2004.

[11] Ching-Te Wang, Tung-Shou Chen, and Zhen-Ming Xu, ``ARobust Watermarking System Based on the Properties of Low Frequency in Perceptual Audio Coding", IEICE Trans.

[12] on Fundamentals of Electronics, Communications and Computer Sciences, vol. E87-A, no. 8, pp. 2152-2159, August 2004.

[13] Steganography and Watermarking - Attacks and Countermeasures by N.F. Johnson, (Kluwer Academic Publishers, 2000)