Sensitive Credit Card Data Security

Sagar Tawde, Aditi Uchil, Anjali Sharma
Department of Computer Engineering,
K.C. College of Engineering & Management studies
Kopri, Thane (E)-400 603, India

***Abstract*— The domain of e-commerce has grown into a billion dollar industry today. As a result of which different security techniques have been developed. Here we are now proposing a system to secure credit/debit card details. Data such as credit/debit card number, cardholder's name, CVV number, etc are embedded into any wanton image. For this we use different security methods such as arithmetic encoding, image processing techniques, etc. The paper also provides an additional layer for enhancing the security via the use of Hamming Code.**

***Keywords*— *Data compression, Huffman coding, Bit Plane Architecture, Error detection and correction, Image embedding techniques*.**

I. INTRODUCTION

E-commerce has grown considerably and along with it the use of credit cards. Ensuring the protection of such sensitive information over a network is therefore a must. So we now introduce a technique that whenever a user enters his/her credit card detail, all his sensitive information gets embedded in the background of the image. The customer name, card number, CVV number and its expiry date gets embedded into the image. For this we use arithmetic encoding which has a high embedding capacity. Also hamming code is applied. The image is then sent over the web network. The receiver decodes this image and gets the user's credit card information.

II. LITERATURE SURVEY

There exist many water marking techniques for information hiding which includes DCT, DWT [1] (DCT is used for low frequency images and DWT is used for high frequency images.). But our innovation uses lossless code for hiding a data (credit card information) as a watermark [1][2]. There is limited capacity in SVD-DWT. The flaw is such that the embedding and the recovery itself takes up a lot of time. This is because the zigzag scanning to map the coefficients into four quadrants based on the frequency takes a lot of computation [2].

We also have password based cryptography [3] which is vulnerable. Complexity of the system increases and also complex numerical code is involved. That is why the compile time and run time of system increases. By limiting the scale of the damage, ensuring (in the case of password via the mechanism of salting) that the computational effort to compromise all of m instances. Biometric Technology [4][5] is a relatively old technology, and the equipment is very expansive. Holographic Technology [6] is vastly more expensive and risky to counterfeit a new security device in a few days than in a few months. Then, if the copyright holder changes the security device every few weeks, the counterfeiter is unlikely to be able to make a profit; the copyright holder wins the race [14].

Light Weight Cryptography [7] is relatively expensive to be used. Also it is unsuitable for environment. We use arithmetic coding along with cryptography [8] to encrypt the entire credit card data and then embed it onto the image. Tokenization is an outdated method [9]. One of the problems in tokenization systems is that they themselves have quickly become attractive targets to fraudsters - since the middlemen companies no longer store valuable information, criminals are turning to the token holders. Additionally, there are few security standards and best-practices regarding tokenization at this point, which can lead to some unsafe implementations and potential fragmentation throughout the industry[15]. Perhaps one of the most unfortunate drawbacks of tokenization and encryption is the cost, which could either dissuade businesses from using it or encourage them to jack up prices to offset additional costs.

Hamming codes are error detection [10] and error correction codes [11], detecting to two simultaneous bit errors, and correcting single bit errors. Mathematically, they are linear block codes. Although primarily used in communication, Hamming codes [12] also serve the purpose of introducing Data Redundancy, extending their use into Steganography.

A bit plane of an image is a set of bits having the same position in the respective binary number. In bit plane substitution [13] we substitute least significant bits with our desired data (now in binary), since these LSB's are considered to be visually redundant, it does not degrade the image.

III. METHODOLOGY

A. The encoders site (For Customer Name)

The length of a name on any credit card is limited up to 20 alphabets (including spaces). The user is asked for entering the name on the card .The encryption technique converts it to binary format. Alphabet in the user name (customer name) will correspond to its alphabetical

position, in the code. The entered number or position is converted to the 6 digit binary format, so as to keep whole stream of bits always an even number .The process is iterative until whole string is scanned and this becomes our cipher text. The redundancy in the cipher text forms our Hamming code. A hamming code word is generated for every character in the 'Customer Name' string.

Hence by using a (10, 4) Hamming generator matrix, hamming codes for each of the character is generated. Now each character is a 10 bit data word. Therefore, the total cipher text length is now 10* (length of name). The length of the string is also converted into a (11, 4) hamming code before embedding it into the background of an image. Now the data to be embedded onto the image is the 'Customer name' string with each character with a 10 bit data word and the length of the string which is a 11 bit data word.

B. The encoder site (For Customer Card Number)

For encryption of the credit card number we have assumed a credit card number has a 16 digit unique number and additional 3 significant digits CVV number which is necessary for all online transactions. It also encrypts the expiry date in (mm/yy format- 4 digits), which is important for internet transactions. We then take these 19+4 digits separately as 4+4+4+4+3+4 blocks of data and then encode them separately. The user inputs are given in a sequential manner.

The symbols that are set for this experiment is (0-9) and the corresponding symbol probabilities are found from large data sample space. The arithmetic encoder then takes these probabilities and the symbol is set as well as the input sequence and outputs the lower range of the final subinterval.

An example of arithmetic coding is shown below in Table I. After the arithmetic coding process in the above example we obtain six double precision values corresponding to the six input sequences. We round these to six significant digits and then convert it to binary format using our encoding algorithm. Table II illustrates the results of the encoded values for a sample input. This code word, along with the cipher text obtained can be now embedded into the image (refer Fig.1) the image is compressed and then sent to the receiver over the web network.

TABLE I. AN EXAMPLE OF ARITHMETIC CODING

Source Symbol	Probability	Initial Sub-Interval
1	0.08	[0.0,0.08)
2	0.2	[0.08,0.28)
3	0.2	[0.28,0.48)
4	0.12	[0.48,0.6)
5	0.16	[0.6,0.76)
6	0.08	[0.76,0.84)
0	0.04	[0.84,0.88)
.	0.04	[0.88,0.92)
8	0.04	[0.92,0.96)
7	0.04	[0.96,1.0)

TABLE II. STEPS TO ENCODE A CREDIT CARD NUMBER

Input data sequence	5296
Corresponding code word	0.451252245474587
Code word in six significant digit	451252
Equivalent 20-bit binary	11011100010101101 00

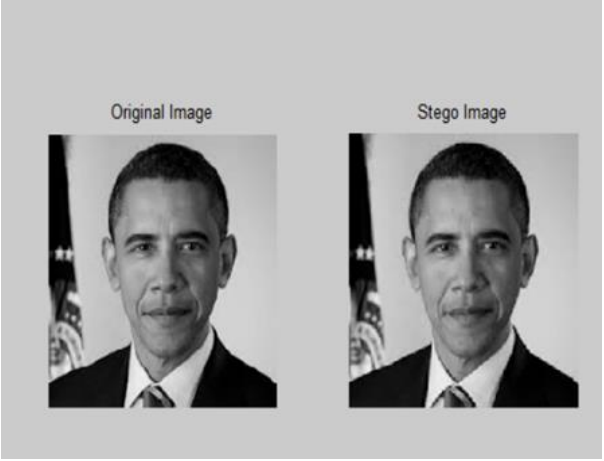

Fig 1. Original and Embedded Image

The entire credit card information is embedded in the pixels of the image (shown in fig 1). Various level of security is integrated and applied on the customer information in order to maintain security services like data confidentiality and data integrity. Thus a scheme of three level security is provided.

The flow chart of entire encoding process is show below (shown in Fig 2)

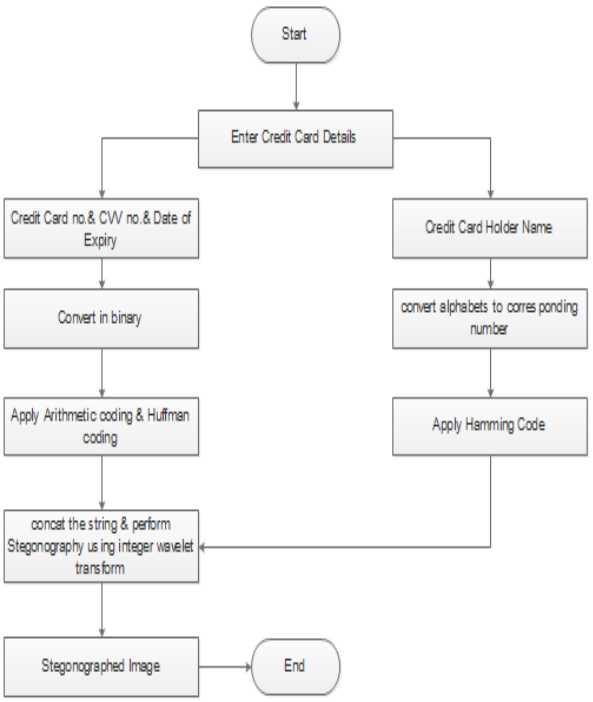

Fig 2. Flowchart for the complete encoding process

C. *At the decoding side/receiver side (For Customer Card Number)*

After receiving the embedded image, the receiver carries out the decryption process using decoding software which employs the exact reverse process. Firstly, the length of the 'Customer name' string is extracted. On the basis of that, remaining data bits are extracted from the image following the reverse algorithm.

Next, the redundant bits are removed with the help of the hamming code matrix. The extracted code word will be sent to the decoder. The decoder will go on extracting 6 continuous bits from the received stream of bits (which will always be a multiple of 6) and convert it to decimal format whose range will be from 1 to 26. Then the decimal data will be mapped to the position of alphabet and corresponding alphabet will be added in the output string. This recursive process will get us the original customer information (the printed name) to the receiver. The decoder now proceeds as follows.

The receiver should have our decoding algorithm. In addition to these the symbol probability, symbol is set and the length of the sequence must be stored. If a password was set at the encoding end then the user must enter the same. The bits are obtained using our reverse algorithm.

Table III illustrates the decoding procedure for the output data in Table II.

TABLE III. STEPS TO DECODE A CREDIT CARD NUMBER

Binary data from image after concatenation	1101110001010110100
Corresponding decimal Number	451252
Decimal number after division by 100	0.451252
Obtained credit card Number	5296

CONCLUSION

We have actually conceptualized a reliable system which will change the way sensitive data is being transferred via internet. We will try to achieve all our criteria which were set for this system and also overcome challenges posed in the way successfully.

The software will be inexpensive and can be efficiently used within a group of systems. Through the literature survey we have also tried to compare our lossless technique with the lossy watermarking techniques like DWT, DCT to show that this technique can be effectively used in future applications of credit card information security. Thus we believe that this paper in its own way will make a significant contribution in the field of e-retail.

ACKNOWLEDGMENT

No work is ever complete without the guidance of those experts who have already walked on this path before and hence become master of it and as a result, our leader. So we would like to take this opportunity to take all those individuals how have helped us in whichever way possible.

We express our deep gratitude to our project guide Mrs. Sonal Balpande for providing timely assistant to our query and guiding us in her own way. She has indeed been a lighthouse for us in this journey.

We extend our sincerity appreciation to all our Professors from K.C.COLLEGE OF ENGINEERING, for their valuable insight and tips. Their contributions have been valuable in so many ways that we find it difficult to acknowledge of them individually.

We also great full to our HOD Mrs. Amarja Adgaonkar, Principal Dr.Hansraj Guhilot for extending their help directly and indirectly through various channel in our work.

REFERENCES

[1] Mei Jiansheng, Li Sukang and Tan Xiaomei "A Digital Watermarking Algorithm Based On DCT and DWT" .
[2] Sumit Kumar Prajapati, Amit Naik, Anjulata Yadav "Robust Digital Watermarking using DWT-DCT-SVD"
[3] Mihir Bellare, Thomas Ristenpart, Stefano Tessaro "Multi-Instance Security and its Application to Password-Based Cryptography"
[4] A Report By Ingenia Technology "Identity, Secure Documents & AntiCounterfeiting" May 2012

[5] Prithika M., P Rajalaksmi "Credit Card Duplication and Crime revention Using Biometrics"
[6] Stephen P. McGrew, President New Light Industries, Ltd. "Holographic Technology for Anti-Counterfeit Security: Present and Future"
[7] A. Poschmann, M. Robshaw , F. Vater , C.Paar "Lightweight Cryptography and RFID: Tackling the Hidden Overheads"
[8] Vishwa gupta,Gajendra Singh,Ravindra Gupta "Advanced cryptography algorithm for improving data security"
[9] A First Data White Paper September 2009 "Data Encryption and Tokeniation:an innovative One-Two punch to increase Data security and reduce the challenges of PCI DSS Complaince"
[10] Error Detection and Correction
[11] Allen D Holliday "Hamming error correction code"
[12] Eltayeb S. Abuelyaman and Abdul-Aziz S. Al-Sehibani "Optimization of the Hamming Code for Error Prone Media"
[13] Mazdak Zamani,Azizah Bt Abdul Manap,Shahidan M Abdullah,Saman Shojae Chaeikar "Correlation between PSNR and Bit Per Sample Rate in Audio Steganography"
[14]Stephen P. McGrew, President, New Light Industries Ltd , July 1996"Holographic Technology for Anti-Counterfeit Security: Present and Future"
[15] Datacap systems inc "The advantages and disadvantages of tokenization"