Research Journal of Computer and Information Technology Sciences _____ISSN 2320 – 6527
Vol. **5(2),** 1-5, March **(2017)** *Res. J. Computer and IT Sci.*

Anti-phishing approach using TEA, steganography and MD5 algorithm

Rubeena Jabi, Deepty Dubey* and Punyaban Patel
Department of Computer Science & Engineering, Chhatrapati Shivaji Institute of Technology, Durg, CG, India
deeptydubey@csitdurg.in

Available online at: www.isca.in
Received 24th February 2016, revised 12th March 2017, accepted 18th March 2017

Abstract

Phishing is an endeavor by an individual, to theft the individual data of user, for example: secret word, Visa number and client id and so on. Phisher sends email to their victim and that email contain connection of fake site that resembles as original site and client experience that connection and enter their own data. It is used by phisher for profit, relates with cash and other false movement. This exploration work is to solve the issue of phishing by combining as to tiny encryption algorithm (TEA), Steganography and MD5 calculation. Before send their own data client will give security to their information. To begin with TEA, it encode the data , second level security is to conceal that encrypted data behind the photo using LSB steganography method furthermore utilizing MD5 estimation to give acceptance from client side and to check the honesty of data. Server will have the mechanism to decrypt the personal information of the client. By joining these method phishing assault can prevent. It will give confidentiality and also validation to the information of clients.

Keywords: Anti-Phishing, Security, Phishing, Steganography, MD5, TEA.

Introduction

The specialty of remodelling the information into misty organization alluded to as Cipher Text. The Encrypted message will decode (or Decrypt) into Plain Text singularly by individuals who have a mystery key, this methodology known as cryptography. Sender encodes the Plain Text by using mystery key and encoding recipe and it'll decode into Cipher Text. This Cipher Text send to the recipient through the system then, CT (Cipher Text) can transmit back to original Plain Text using decryption algorithm and secret key. In cryptanalysis, assault (attack) is the endeavor to "break" an Encrypted message created by framework. Security assault is any activity that compromises the security of data, possessed by an association and to identify, keep and recuperate from a security assault, a procedure is composed called security mechanism. There are two sorts of assault one is the Passive assault and another is Active assault[2].

In online ambush, phishing is understood now a day, there is e-saving cash and e-business is used and today, finance trades are done online. Yet for this, security is vital and to give the security to these online exchange different procedures are utilized. Phishing could be a system for assault, which is finished by a distinctive individual known as phisher; they assault in individual information of a client. For instance: Phisher send the email to the victim as, because of some specialized deficiency SBI database is crashed and urgently require your record related data, then client experience the connection that contain in email and visit fake site that precisely resemble a unique site. At the point, when clients enter the information (client ID, secret key) and submit these information then, phisher robbery this information, to use this for unlawful trickery[3].

There are different sorts of phishing, for example: Deceptive phishing, Malware based totally phishing, Hosts report harming, Domain Name System-based totally phishing, man-in-center phishing and search motor phishing. This proposed work will avert the data from Deceptive Phishing in which, with a message, electronic mails are sending to the casualty by a phisher. Clients unit of estimation affected to tap on a connection and visit the online site and victim enter their information. It used by wrongdoer for the fraud related to money transaction[1].

Hostile to Phishing methodology is identifying and keeping the phishing PC. A hostile to phishing (Anti-Phishing) project tries to confirm phishing substance contained in registering machine or email. In some cases fused with web programs. Hostile to phishing common sense may like manner be encased as an inbuilt limit of some application program. Against Phishing administration is a mechanical administration that helps to prevent unapproved access to secure data. Hostile to phishing administrations protect different sorts of information[4].

Tiny encryption algorithm (TEA)

“The Tiny Encryption Algorithm", is a shortened form of TEA. In 1994, it's made by David Wheeler and Roger Needham, additionally called as Wheeler and Needham, in the Computer exploration offices of University of Cambridge. The Tiny Encryption Algorithm or minor encryption count uses operations from mixed logarithmic social occasions (gatherings)

XOR, SHIFT and ADD. The two fold moves are used to blend all bits of the data and keys more than once. The key schedule estimation is essential; the key K of 128-piece (bit) is separated into four squares of 32-bit i.e. K = (K [0], K [1], K [2], K [3]). TEA is sufficiently short to compose into any project on any PC. The Tiny Encryption Algorithm (TEA) implement is very easy; it takes quick execution time and consumes insignificant storage space (memory).

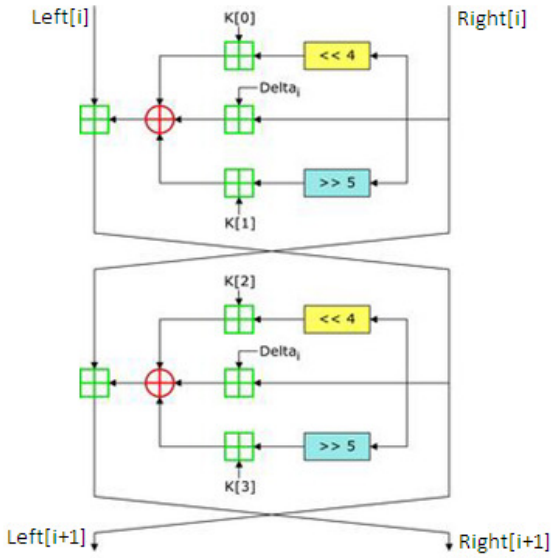

Figure-1: i[th] Cycle of Tiny Encryption Algorithm.

Encryption Function:

$$L[i+1] = L[i] + f(R[i], K[0, 1], d[i]) \quad (1)$$

$$R[i+1] = R[i] + f(L[i+1], K[2, 3], d[i]) \quad (2)$$

Where: Left= L, Right= R, Delta=d, d = 0x9e3779b9, d [i] = (i+1/2)* d.

Let,
R[i] =M1 and L [i+1] =M2

$$f(M1, K[0,1], d[i]) = ((M1<<4) + K[0]) \wedge (M1+d[i]) \wedge ((M1>>5) + K[1]) \quad (3)$$

$$f(M2, K[2,3], d[i]) = ((M2<<4)+K[2]) \wedge (M2+d[i]) \wedge ((M2>>5)+K[3]) \quad (4)$$

Decryption Function:

$$L[i+1] = L[i] - f(R[i], K[0, 1], d[i]) \quad (5)$$

$$R[i+1] = R[i] - f(L[i+1], K[2, 3], d[i] \quad (6)$$

Where,
d = 0x9e3779b9
d [i] = (i+1/2)* d

Let,
R[i] =M1 and L [i+1] = M2

$$f(M1, K[0,1], d[i]) = ((M1<<4)+K[0]) \wedge (M1+d[i]) \wedge ((M1>>5) + K[1]) \quad (7)$$

$$f(M2, K[2,3], d[i]) = ((M2<<4)+K[2]) \wedge (M2+d[i]) \wedge ((M2>>5) + K[3]). [5] \quad (8)$$

TEA is fiestel sort figure. In this the substance being encoded is partitioned into 2 sections. What's more, round function connected to one half and X-OR key with the shifted estimation of the last phase of each piece of message content. Function's output is in a matter of seconds X-OR with another half. For each round of TEA, two segments are swapped. No swap for the last round. The steady delta is gotten from the brilliant number extent (golden number proportion) to ensure that the 4 keys are unmistakable and keys qualities have no cryptographic noteworthiness. Decryption is inverse process of encryption procedure There are the TEA encode just 64 bit piece of info content at once so, it takes 64 bit of block information and divide it into two parts i.e. two 32 bit block known as Left[i] and Right[i]. It utilize 128 piece key and separation it into 4 sub keys of 32 bits. It contains 32 cycles and 64 rounds. Every cycle incorporate 2 rounds. There are four keys as: K0, K1, K2, and K3 thusly, K0 and K1 used for the odd rounds and K2 and K3 used for the even conforms[5].

Steganography

Steganography is gotten from Greek word "Stegano". Stegano signifies "cover" and graphy signifies "writing". Steganography use to send data clandestinely and riddle picture is taking to disguise the information behind riddle picture and picture's each pixel has 3 shading (color) furthermore, picture incorporate a large number. The simplest and most standard kind of steganography is LSB (Least noteworthy piece). The one bit of byte is used to encode the information. Content, Video, Image and Audio are different sorts of steganogaphy. In this paper, utilize picture steganography to conceal information behind picture[6].

A computerized picture is portrayed utilizing a 2-D network of the color intestines at every matrix point (i.e. pixel). Normally gray pictures utilize 8 bits, where as hued (colour) uses 24 bits to portray the color model, for example, RGB model. The Steganography structure, which utilizes a photo as the cover, there is using some steps, to conceal information inside cover picture. The LSB systems control the cover picture's pixel bit values to implant the mystery data. The mystery bits are composed straightforwardly to the cover picture pixel bytes. LSB systems are basic and simple to actualize. The LSB i.e. Least Significant Bit is one of the primary procedures in spatial space picture Steganography.

Use picked pixel value to address character instead of a shading esteem. Each character in message changed over into its ASCII regard that, change over into bits and introduce each bit of message into LSB position of each pixel position.

Example1: Suppose unique pixels as bits:
(Rd7 Rd6 Rd5 Rd4 Rd3 Rd2 Rd1 Rd0 (Red), Gr7 Gr6 Gr5 Gr4 Gr3 Gr2 Gr1 Gr0 (Green), Be7 Be6 Be5 Be4 Be3 Be2 Be1 Be0 (Blue))

Bits of Character as take:
(Ch7 Ch6 Ch5 Ch4 Ch3 Ch2 Ch1 Ch0)

Spot the character piece at LSB of pixel bits as:
(Rd7 Rd6 Rd5 Rd4 Rd3 Ch7 Ch6 Ch5, Gr7 Gr6 Gr5 Gr4 Gr3 Ch4 Ch3 Ch2, Be7 Be6 Be5 Be4 Be3 Be2 Ch1 Ch0)

Example2: Pixel value (225,100,100) with character A. Pixel in bits (11100001, 01100100, and 01100100). And ASCII value of A=65 and A in bits (01000001).

New pixel bit values = (11100010, 01100000, 01100101)

New pixel integer values = (226, 96, 101).

In example 2, gives the amount of pixels i.e. red, green and blue. It will changed over into the bits and character of message substance's ASCII qualities are changed over into the byte organize then, in LSB position of the pixel values message byte is set. Through this procedure a message is hiding behind the picture.

The pixel estimation of picture before stowing away and pixel estimations of picture after prevent form being seen is not very diverse. In this way, there are both pictures appear to be identical. There is no an excess of distinction between before picture and after picture for steganography[7].

In this structure the puzzle (mystery) picture pixel is partitioned into sub pixel or 2x2 piece of sub pixel. Security of mystery picture is relying upon the color composition of the mystery picture. There are to cover the mystery picture segment, cover picture is required and mystery picture recuperation required the cover picture, ought to be deciding the shape and pattern of mystery picture and decide the limit between 2 diverse color districts in picture[9].

md5 Algorithm

MD5 is message digest calculation created by the Ron Rivest. Message digest is a synopsis (summary) of message and it is utilized to check the integrity of information. MD5 is very straightforward and quick, it produces 128 bit message digest. Over years scientists have created potential shortcoming in these. However so, for MD5 has possessed the capacity to effectively guard itself against collision[10].

There is case taken that appears, how a number message will change over into digest form: Message = 4672389. Perform the operations to convert message into message digest. It is impossible to create reverse of the message digest.

Table-1: Message digest example.

Operations	Result
Multiply 4x5	24
Discard first digit	4
Multiply 4x7	28
Discard 2	8
Multiply 8x2	16
Discard 1	6
Multiply 6x3	18
Discard 1	8
Multiply 8x8	64
Discard 6	4
Multiply 4x9	36
Discard 3	6

Message digest = 6.

Existing methodology

There are RSA principle is used for the key composition and mystery composing of client id and secret word. For the validation client pick an image captcha and by visual cryptography, it separates that picture into shares of 2, one share kept with client and another share kept with the trusted server; these all procedure is done in registration stage. In login stage, client utilizes that picture share as a secret key. In login stage enter client id, select picture share and enter public key. The client id, picture share is encoded by utilizing the general population key and sent to the server, where it's decrypted by utilizing public key or private key. Presently server's share and client's share stacked along, a unique picture is demonstrated which is then sent to the client's program. A client uses that picture captcha for login to check whether the site is phishing site or not. In this methodology RSA calculation is utilized which has an issue of factoring of substantial entire number. RSA is an Asymmetric key cryptography and it has complex numerical count[1].

The progressions of visual cryptography during which, two stages are conveyed. Initially is registration stage and second is login stage. In registration stage the mystery key is approached from the client for the protected site. The server enters the key and client enters the key, then by the every key string composed in a rundown and picture captcha is produced and separated it into two shares utilizing visual cryptography. One key kept with

client and another kept with the server. In login stage client enter that share of picture captcha rather than mystery. When, the client's share stake with the server's share then, unique picture captcha is uncover. Through the uncovered picture captcha client identifies its phishing site or not. In the event that uncover picture captcha is same with the created captcha of registration part, in this manner it isn't phishing site however in the event that it isn't coordinated then its phishing site. Coordinated captcha is used to sign in into site. It gives validation in the middle of client and server. Be that as it may, it just gives the validation in the middle of customer and server, not gives secrecy to client's information[8].

The technique for information exchange between the server of phishing site and phisher's victim is justified in which the phishing site server requests that the victim present the information and on submission the phisher will snatch the data. When client or victim inquiries the data to the phishing site server for recovering data then, no information is given by the phishing site server. In this strategy phishing is prevented utilizing code word system. It contains two stages in particular sign-in stage and sign-up stage. In sign-in stage client can enlist by going into unique site, then subsequent to filling of registration form a client ID, secret word is made. By the code generation strategy a novel code is produced by site (association) that is saved with the details of client. The created code is likewise sent to client. This code should be reviewed by customer**.** In sign - up stage through the email client will get the connection of site then client experience that connection, fills their client id and any 2 digits as an interesting code. On the off chance that the digit is right then complete code is shown by the server on the screen of client, if the code coordinates the code that was created at the season of registration then, the client can be sure that the page is not a phishing page. The code is a mix of client ID character, secret word character number (it must be of 5 characters) and date and month (it ought to be more noteworthy than or equivalent to ten, however in the event that it's not exactly in date and month's entirety include 10). In the sign-up stage client will enter any random code in page and if client code is right then the server will give complete right code on client's screen. This confirms the site is legitimate[11].

Proposed methodology

For the phishing prevention, this paper proposing a new methodology. It based on 2 levels security; it will give verification and additionally secrecy to client's information.

Sender Side Process: In Figure-2 (a) the customer enters his/her sensitive information and before sending this sensitive data (delicate information) that is in plain content will experience the 2 levels of the security. To start with, plain content will encode through the TEA and key of 128 bit then, it will make the CT1. In second level of security, encoded information is hide behind the picture utilizing LSB steganography and produce CT and utilizing MD5 calculation for give the verification plain text's message digest will make and it is joined with the CT that send to the server or collector.

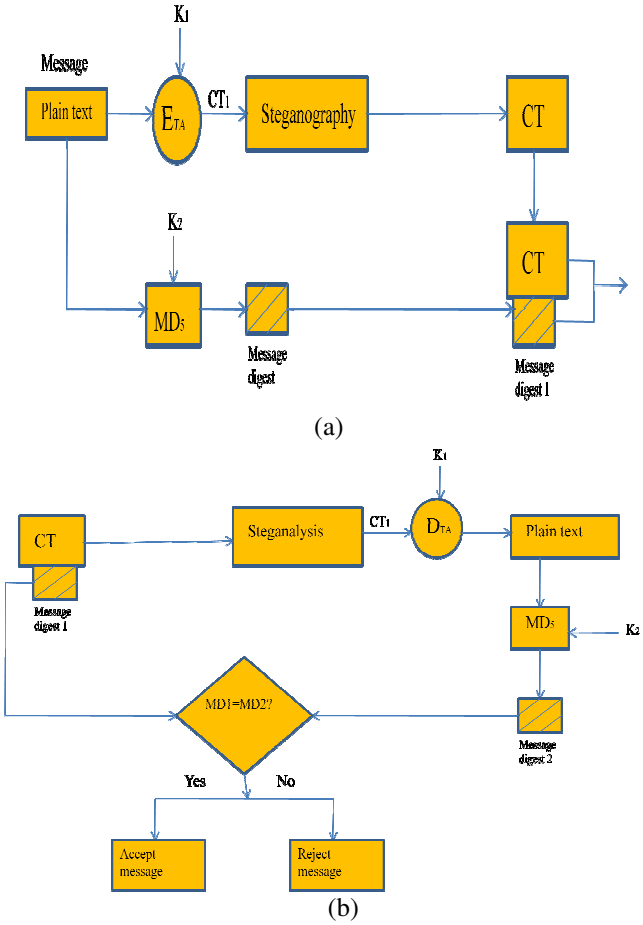

Figure-2: (a) Sender side process (b) Receiver side process for Confidentiality and Authentication.

Receiver Side Process: At the receiver side of Figure-2(b) the server or collector will get, the consolidated cipher text and message digest. It performs the decoding of CT in 2 levels. To begin with, second level, it will extract the encoded information from the picture that is known as steganalysis and in first level, it will decode the encrypted information utilizing tiny encryption algorithm and a mystery key. At that point, collector will get the plain substance. The identity of the sender is confirm by using the MD5 algorithm and it will make the message summary of the decoded plain substance by MD5 calculation utilization, For the verification and it will look at the MD1 and MD2, if both are comparative then, it will accepted by the server or recipient else it will dismisses by the collector. After receiving data of client, server will send acknowledge to the client.

Conclusion

To prevent the phishing, there are various methodology have been developed. Phishing is an issue that extending regulated along these lines, this necessity the high state of security. This

venture will give the 2 levels security to the delicate data of the customer before sending its information over the web. If, any phishing site will visit by the customer thusly, before sending its information, customer will give the 2 levels of security to its information. In the first place use TEA to encodes data then, using steganography to cover up encrypted data behind the photo. This technique gives the confidentiality to the client and the server. Here, in like manner use MD5 to check the integrity of data and besides give the confirmation.

References

1. Vaidya S., Zarkar S., Bharambe Achal N., Tadvi A. and Chavan T. (2015). Anti-Phishing Structure Based on Visual Cryptography and RSA Algorithm. *International Journal of Engineering Trends and Technology*, 20(4), 209-213.

2. Stallings William (2006). Cryptography and Network Security principles and practices. Pearson education, South Asia, 1-680. ISBN: 978-81-7758-774-6.

3. Gaurav Mishra M. and Jain A. (2012). Anti-Phishing technique: A Review. *International Journal of Engineering Research and application*, 2(2), 350-355.

4. Kumar N., Kumar S. (2015). Anti-Phishing software at DMOZ. https://en.wikipedia.org/wiki/Anti-phishing_ software. 15/11/2015.

5. Shoeb M. and Gupta V.K. (2013). A cryptanalysis of the tiny encryption algorithm in key generation. *International journal of communication and computer technologies*, 1(5), 123-128.

6. Tyagi V. (2012). Data hiding in image using least significant bit with Cryptography. *International journal of Advanced Research in computer science and software engineering*, 2(4), 120-123.

7. Sing S. and Agarwal G. (2010). Use of image to secure text message with the help of LSB replacement. *International Journal of applied engineering research*, 1(2), 200-205.

8. James D. and Philip M. (2012). A Novel Anti phishing framework based on visual cryptography. *International Journal of Distributed and Parallel System,* 3(1), 207-218.

9. Hou Y.C. (2003). Visual Cryptography for color images. *The Journal of the pattern recognition society*, 36(7), 1619-1629.

10. Mishra D.P. (2007). Notes on Cryptography. Department of Computer Science & Engg. B.I.T., Durg.

11. Gaurav Mishra M. and Jain A. (2012). A Preventive Anti-Phishing Technique using Code Word. *International Journal of Computer Science and Informational Technologies*, 3(3), 4248-4250.