Anti-Phishing Structure Based On Visual Cryptography and RSA Algorithm

Sayali Vaidya[#1], Shreya Zarkar[#2], Prof. Achal N. Bharambe [#3] , Arifa Tadvi[#4] ,Tanashree Chavan[#5]

[1,2,3,4,5]*Department of Computer Engineering, Modern Education Society's College of Engineering, Savitribai Phule Pune University, Pune, India.*

***Abstract*— Now-a-days online attacks have increased to a great extent and the most popular attack among them is phishing. Phishing can be basically defined as one kind of attack in which various attackers acquire the confidential and sensitive information of the victims. Thus, security in such cases should be very high to avoid the online attacks. Phishing steals the confidential information such as password, credit card information, etc which is carried out by fraudsters. So it is very much important for the users to identify the fake website and avoid falling prey to it. In this paper we have proposed a new approach named as "Anti-phishing structure based on visual cryptography and RSA algorithm" to solve the problem of phishing. Here an image based authentication using Visual Cryptography (VC) and the encryption algorithm is used. Visual cryptography is mainly done by splitting the original image into two shares one with user database and one with the server database. And the original image can be obtained only by both the shares of the image.**

***Keywords*— Phishing, Visual cryptography, Encryption Algorithm, RSA, Decryption, Visual Cryptographic Scheme.**

I. INTRODUCTION

Now-a-days, online transactions are very commonly used and various online attacks are present behind this. So with the rising threats effective preventive mechanisms should be made. Phishing is one kind of attack in which confidential and sensitive information is gained by various attackers. Among all online attacks phishing is identified as the major attack. So therefore security in such cases should be very high. Now-a-days the applications are only as secure as their underlying system and as there is improvement in the technology of the middleware so the detection of the problem becomes difficult. Phishing attacks are becoming a threat for online transactions users and e- commerce user's. Phishing can be basically defined as a form of online identity theft that steals the confidential and sensitive information such as password, credit card information, etc of various victims [1] as per Divya James and Mintu Philip. Another definition can be given as "Phishing is a criminal activity using social engineering". Phishing is an attempt by fraudsters to steal victims account related information such as user id's, passwords by sending e-mails which appears to originate from trusted source like Banks, Tax authorities, etc. These emails states urgently updating the accounts information.

Phishing mainly aims to acquire the user's confidential information such as username, password and credit card information. In phishing attacks are mainly done on the websites where the attackers carry out fraudulent activities such as the financial transactions from the user's side by sending an email which contains a fake URL that redirects user to fake website. Attacker uses replica of original website that is send to the user, user fills and submits the sensitive and useful information into the website. The attacker extracts all the information of the victim and saves the data for its own illegal use. And there are different types of phishing attacks like deceptive phishing, malware based phishing, Web Trojans, System reconfiguration attack.

II. PHISHING AND ITS CLASSIFICATION

The most successful phishing attacks is when the victim receives an email from attacker regarding need to verify account information, re-enter users information because of system failure, undesirable account changes, new free service, and may other scams with the hope that victim will enter their information and caught in to attackers trap. So in the figure similar thing is shown that the phisher sends lots of mails to any random victims. And the victims getting trapped will enter all the confidential data and this is saved by the attackers for serving criminal purpose.

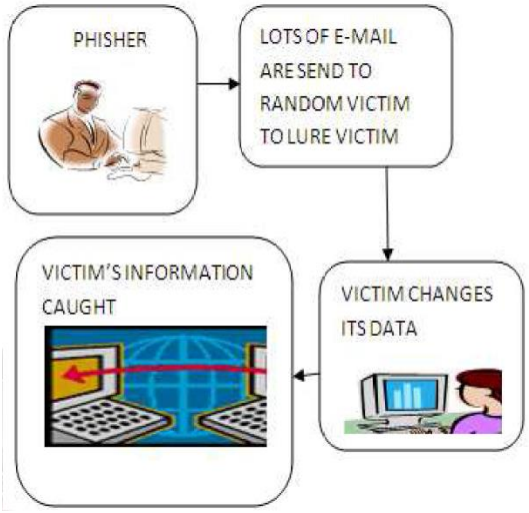

Fig. 1 Process of Phishing

Phishing attacks can be classified in to following types according to the way attack is done. The different types of the phishing attacks are as follows:

1) Deceptive phishing: In this type of phishing the attacker send or broadcasts an email which contains the following type of messages like need to verify account information, re-enter users information because of system failure, undesirable account changes, new free service, etc. The attacker thinks that the victim will enter their information and get caught into the attackers trap.

2) Malware based phishing: This type of attack basically involves running malicious software on victim's machine or pc. Malwares can be introduced as an email attachment,

downloadable file from website or by exploiting security vulnerabilities.

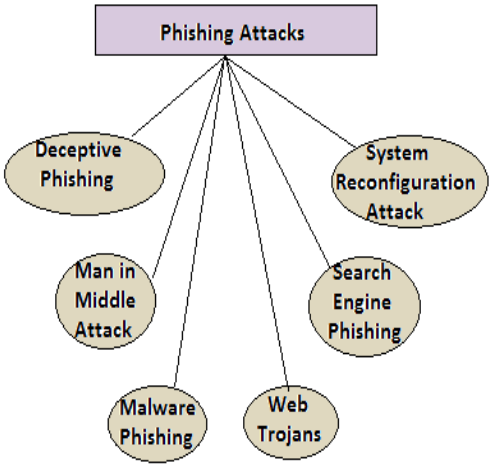

Fig. 2 Types of phishing attacks

3) Web Trojans: In this kind of attacks pops up are coming while using the websites. And they are invisible when users attempt to log in. They collect user's information by storing in the database and then transmit the confidential data to the phisher.

4) System reconfiguration attack: In this type of attack users pc configuration is changed for performing illegal activities. And it redirect users to the URL look alike, for example the Banks URL may be changed from www.gmail.com to www.gmai1.com, here l is replaced by 1 which confuses the victim.

5) Man in middle phishing: In this type of phishing attacker acts as middleware between the user and legal website. The attacker records the user's information and continues to work on the legal website so that user cannot identify that he is being phished. And also the user's transactions are also not affected. And then the attackers may sell or use the user's information when user is not accessing the account.

6) Search engine phishing: In this type of phishing attacker creates very much attractive website by using sound effects and animations. So when users do normal search they fall prey to such kind of websites and end giving up their confidential information.

III. VISUAL CRYPTOGRAPHY

Visual Cryptography is the best technique to protect the data. In this type of technique the original data is converted into some other form to serve the purpose of security. It can be defined as the process of sending and receiving the encrypted data in the form of the messages and this data can only be decrypted by the sender and the receiver. Encryption and decryption are performed by using mathematical algorithms in such a way that only the targeted recipient can read the message by decrypting it. Naor and Shamir [2] introduced the visual cryptography scheme (VCS) and it is a simple and secure way which allows the secret sharing of images without performing any type of cryptographic manipulations and computations. A survey of the related work in the area of visual cryptography is presented in brief in this paper. Visual cryptography schemes were independently introduced by Shamir [3] and Blakley [4]. Their main aim was to secure the cryptographic keys. These schemes are mainly used in the construction of different types of cryptographic protocols [5] and there are many applications based on cryptography in different areas like opening a bank account, opening a safety deposit box, launching of missiles, etc. A segment based visual cryptography suggested by Borchert [6] can be used only to encrypt the messages containing symbols, especially numbers like bank account number, amount etc.

Visual Cryptography Scheme is a cryptographic technique that allows for the encryption of visual information such that decryption can be performed using the human visual system. We can achieve this by one of the following access structure schemes.

The access schemes are as follows:

1. (2, 2) Threshold VCS scheme- This is one of the simplest threshold scheme. It takes a secret message and performs encryption. And two different shares are formed that reveal the secret image only when they are overlapped. Additional information is not required to create such kind of access structure.
2. (2, n) Threshold VCS scheme-This scheme encrypts the secret image into n shares in such a way that when any two (or more) of the shares are overlaid or overlapped the secret image is obtained. The user will be prompted for n, where n is the number of participants.
3. (n, n) Threshold VCS scheme-This scheme encrypts the secret image to n shares such that when all n of the shares are combined or overlapped will the secret image be formed. The user will be prompted for n, the number of participants.
4. (k, n) Threshold VCS scheme- This scheme encrypts the secret image to n shares such that when any group of at least k shares are overlapped the secret image will be obtained. The user will be prompted for k, and n is the number of participants.

In the case of (2, 2) VCS, each pixel P in the original image is encrypted into two sub pixels. Sub pixels are also called as shares. Fig.1 shows the shares of a white pixel and a black pixel. And choosing of shares for a white and black pixel is done randomly determined. And for each pixel there are two choices available.

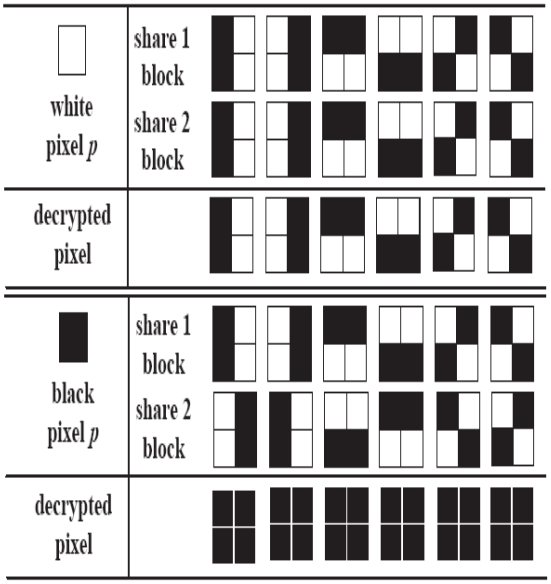

Fig. 3 2-out-of-2 VCS schemes with 2 sub pixel Construction.

Neither of shares provides any clue about the original pixel as different pixels in the secret image will be encrypted using independent random choices. The value of the original pixel P can be determined only when the two shares are overlapped. If we get two black sub pixels then P is a black pixel. And if we get one black sub pixel and one white sub pixel then is a white pixel.

IV. EXISTING METHODOLOGY

The figure of the current scenario is show below. Here the end user tries to access his confidential information online while accessing his account. The end user logs in into his bank account and enters his confidential information like username, password, credit card number, etc. on the login page thinking it as original login page. But if it is a phishing website then the confidential information is saved by the attacker in the databases. A phishing website can gathers the confidential login information of the user when the user enters it and redirect him to the original site. So user doesn't come to know about this. And later on when the user becomes inactive then this information is used by the attackers for illegal purposes like money transfer. Confidential and all the necessary information can be directly obtained from the user at the time of his login.

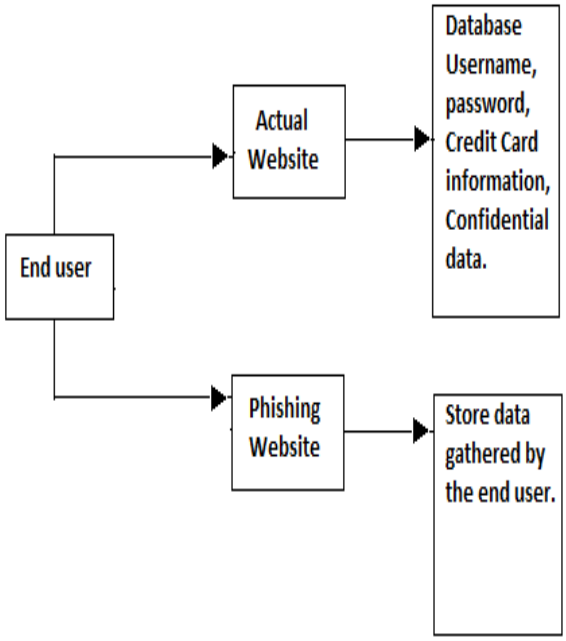

Fig. 4 Current scenario

In the current scenario there are a lot of chances of end users falling prey to the phishing websites. So to avoid the drawbacks of the current methodology some changes need to be made in it. Thus the changes prove to be beneficial and very much helpful.

V. PROPOSED METHODOLOGY

In this paper we have proposed a new methodology to detect the phishing website and preventing it.
The proposed architecture can be divided into two main phase:
A. Registration Phase
B. Login Phase

A. Registration Phase

The registration phase mainly consists of the most important part that is the creation of shares from the image. And here one share is kept with the user and other share is kept with the server. If server under test sends some different share then the stacking or overlapping of shares will create unrecognizable form of image.

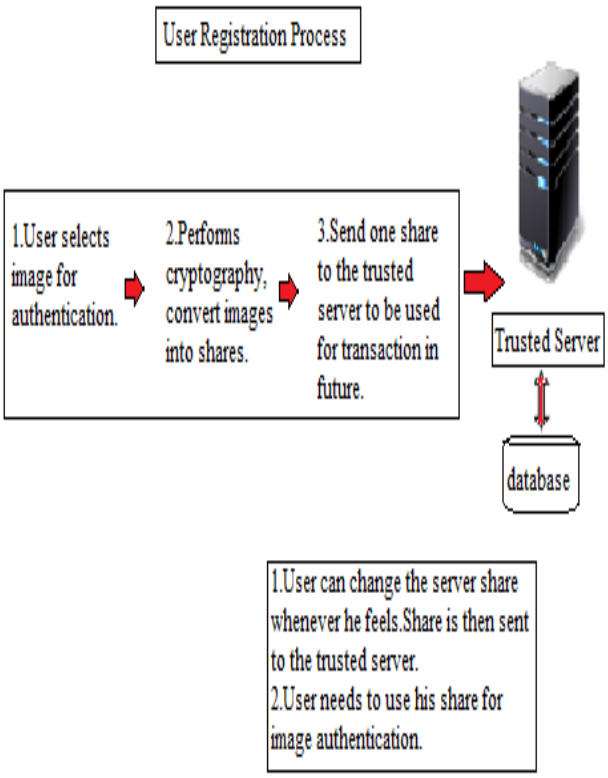

Fig. 5 Registration Phase

And the user can change the share of the images according to his/her wish. Basically the user needs his share of image for the authentication purpose. Then the public key and private keys for the user will be obtained. These keys are very important and will be used by the user during the login phase. All the information regarding the user such as username, password and public and private keys will be stored into database keys will be stored into the database of the server. The user goes through the registration phase only once for registering. And if the user wishes to change his information then he has to create new registration for the user. The stepwise procedure is shown in the figure given above. And the registration phase is very important as login depends on this phase. The figure helps to understand the registration phase more clearly.

B. Login Phase

In the login phase the user needs to enter the user id and his share of image with the public key. The user id, share of image and public key is then sent to the server. Then decryption of the user using the public and private key is done.

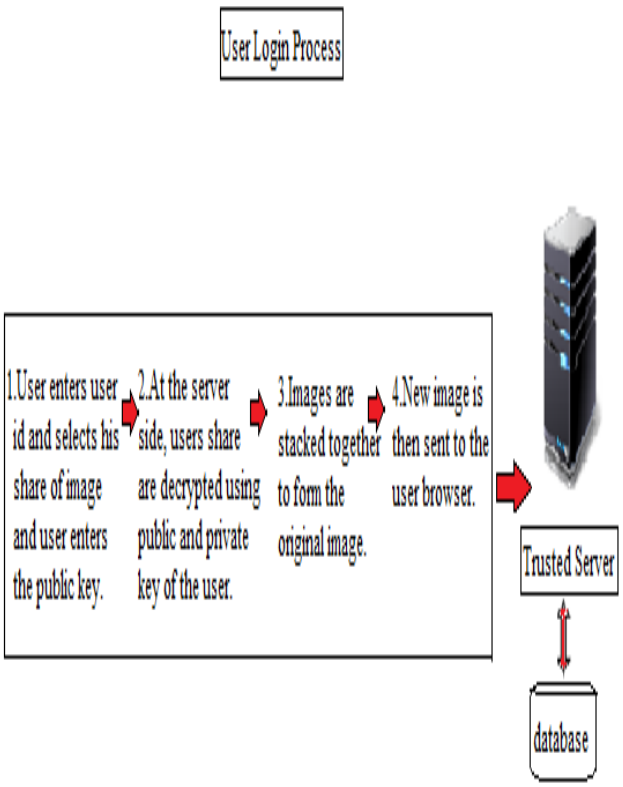

Fig. 6 Login Phase

While at the server side both the share of images that is the server share and the user share are stacked together to form the original image. Then this original image is sent to the user's browser window. Now the user will be able to identify that this is the trusted server and user can enter his further credentials details. So the confidential information entered will be secure and not at any risk. And there won't be any risk as the site is not phishing site. Login phase is represented diagrammatically in the above figure. Here only when the server is found to be trusted then the confidential details are given. The figure helps to understand the Login phase more easily.

VI. IMPLEMENTATION

For public-key cryptography, RSA algorithm is used. And this is based on the presumed difficulty of factoring of the larger integers which is the factoring problem. RSA basically stands for Ron Rivest, Adi Shamir and Leonard Adleman, who first publically described the algorithm in the year 1977.

The RSA algorithm works as follows. Initially the user of RSA algorithm creates and publishes the product of two large prime numbers with an auxiliary value. And the auxiliary value acts as its public key. Here the prime factors must be kept secret. The public key is used to encrypt a message which can be used by anyone. Considering the currently published methods, if the public key is large, then the decoding of message should be done by someone who is having knowledge of the prime factors. Whether breaking RSA encryption is as hard as factoring is a big question known as the RSA problem. The RSA algorithm is as follows:

The RSA algorithm involves three steps:

A. Key generation
B. Encryption
C. Decryption

A. Key generation

The RSA contains a public key and a private key. The public key can be known to anybody and is mainly used for encrypting messages. The messages can be encrypted with the public key and can be decrypted using the private key only. The keys can be generated as follows for the RSA algorithm.

a. Choose two distinct prime numbers p and q. For security purposes, the integer p and integer q should be chosen at random, and should be of similar bit-length.

b. Compute n = pq. n is used as the modulus for both the public and private keys.

c. Compute φ (n) = (p – 1) (q – 1), where φ is Euler's totient function.

d. Choose an integer e such that 1 < e < φ (n) and greatest common divisor of (e, φ (n)) = 1; i.e., e and φ (n) are co-prime. e is released as the public key exponent. e having a short bit-length and small. Hamming weight results in more efficient encryption.

e. Determine d as: D=e-1 mod (ϕ(n)). This is more clearly stated as solve for d given (de) = 1 mod φ(n) d is kept as the private key exponent. By construction, d*e= 1 mod φ(n). The public key consists of the modulus n and the public (or encryption) exponent e. The private key consists of the modulus n and the private (or decryption) exponent d which must be kept secret. (p, q, and φ(n) must also be kept secret because they can be used to calculate d.)

B. Encryption

Sender A does the following.

a. Obtains the recipient B's public key (n, e). 0<=m<n

b. Represents the plain text message as a positive integer m such that computes the cipher text. Where C can be given as, C= me (mod n)

c. Sends the cipher text c to B.

C. Decryption

Recipient B does the following:

d. Alice can recover by using her private key exponent via computing. Then it extracts the plaintext from the integer representative m.

VII. CONCLUSION

Now-a-days internet is used on a large scale so the phishing attacks are becoming very common. The phishing attacks can globally acquire the user's confidential information like username, credit card number, and password, etc. And this data may be stored on to the database and may be used for the illegal purposes. Phishing is basically the attack mainly done to gain the access to confidential information of the victims. By using the proposed method of "Anti-phishing Structure Based on Visual Cryptography and RSA Algorithm" phishing websites can be identified. Thus with help of the techniques used in the paper we can successfully helped the users to identify the fake and genuine website so that he doesn't fall prey to the phishing attacks. Thus the security purpose gets served here.

ACKNOWLEDGMENT

The authors gratefully acknowledge the contributions of M. Naor and A. Shamir for their work in the field of visual cryptography as a simple and secure way to allow the secret sharing of images without performing any cryptographic computations. So the sharing of the images becomes simple.

REFERENCES

[1] Divya James and Mintu Philip, A Novel Anti-phishing framework based on visual Cryptography, 2012.

[2] M. Naor and A. Shamir, "Visual cryptography," in Proc. EUROCRYPT, 1994, pp. 1–12.

[3] A. Shamir, .How to Share a Secret, Communication ACM, vol. 22, 1979, pp. 612-613.

[4] G. R. Blakley, Safeguarding Cryptographic Keys, Proceedings of AFIPS Conference, vol. 48, 1970, pp. 313-317.

[5] A. Menezes, P. Van Oorschot and S. Vanstone, Handbook of Applied Cryptography, CRC Press, Boca Raton, FL, 1997.

[6] B. Borchert, .Segment Based Visual Cryptography, WSI Press, Germany, 2007.