Volume 3, Issue 10, October 2013 ISSN: 2277 128X

International Journal of Advanced Research in Computer Science and Software Engineering

Research Paper

Available online at: www.ijarcsse.com

Anti Phishing Approach Using Probabilistic (t,∞) VC Scheme

Mangala S. Wale, Gayatri M. Bhandari

Computer Department & University

Pune, India

Abstract— With the advent of internet, various online attacks have been increased and among them the most popular attack is phishing. phishing is an attempt by an individual or a group to threats/hackers seeking to retrieve an individual's personal private information such as passwords, usernames/confidential information, credit card information etc. Fake websites which appear very similar to the original ones are being hosted to achieve this. In this paper, we have proposing a new approach named as "Anti Phishing approach using probabilistic (t,∞) VC Scheme" to solve the problem of phishing. Here an image based authentication using (t,n) Visual Cryptography is implemented. The (t,n) visual cryptography (VC) is a secret sharing scheme where a secret image is encoded into n transparencies(shares),and the stacking of any t out of n transparencies reveals the secret image. This paper proposes a Anti Phishing approach using (t,∞) VC scheme with unlimited based on the probabilistic model in which orignal image captcha will decomposing the into two shares that are stored in separate database servers (one with user and one with server) such that the original image captcha can be made only when both are made available; the individual part of images do not make the identity of the original image captcha. Once the original image captcha is revealed to the user it can be used as the password. Using this website cross verifies its identity and proves that it is a genuine website before the end users. For phishing detection and prevention, we are proposing a new approach to detect the phishing website.

Keywords— Phishing, Visual Cryptography, Image Captcha, Shares, Security,Secret sharing.

I. INTRODUCTION

Online transactions are nowadays become very common and there are various attacks present behind this. In these types of various attacks, phishing is identified as a major security threat and new innovative ideas are arising with this in each second so preventive mechanisms should also be so effective. Thus the security in these cases be very high and should not be easily tractable with implementation easiness.Today, most applications are only as secure as their underlying system. Since the design and technology of middleware has improved steadily, their detection is a difficult problem. As a result, it is nearly impossible to be sure whether a computer that is connected to the internet can be considered trustworthy and secure or not. Phishing scams are also becoming a problem for online banking and e-commerce users. The question is how to handle applications that require a high level of security. Phishing is a form of online identity theft that aims to steal sensitive information such as online banking passwords and credit card information from users. One definition of phishing is given as "it is a criminal activity using social engineering techniques. Phishers attempt to fraudulently acquire sensitive information, such as passwords and credit card details, by masquerading as a trustworthy person or business in an electronic communication". Another comprehensive definition of phishing states that it is "the act of sending an email to a user falsely claiming to be an established legitimate enterprise into an attempt to scam the user into surrendering private information that will be used for identity theft". The conduct of identity theft with this acquired sensitive information has also become easier with the use of technology and identity theft can be described as "a crime in which the impostor obtains key pieces of information such as Social Security and driver's license numbers and uses them for his or her own gain". Phishing [1] attacks rely upon a mix of technical deceit and social engineering practices. In the majority of cases the phisher must persuade the victim to intentionally perform a series of actions that will provide access to confidential information.

Communication channels such as email, web-pages, IRC and instant messaging services are popular. In all cases the phisher must impersonate a trusted source(e.g. the helpdesk of their bank, automated support response from their favourite online retailer, etc.) for the victim to believe. To date, the most successful phishing attacks have been initiated by email – where the phisher impersonates the sending authority (e.g. spoofing the source email address and embedding appropriate corporate logos). For example, the victim receives an email supposedly from support@mybank.com (address is spoofed) with the subject line 'security update', requesting them to follow the URL www.mybank-validate.info (a domain name that belongs to the attacker – not the bank) and provide their banking PIN number. So here introduces a new method which can be used as a safe way against phishing which is named as "A novel approach against Anti-phishing using visual cryptography". As the name describes, in this approach website cross verifies its own identity and proves that it is a genuine website (to use bank transaction, E-commerce and online booking system etc.) before the end users and make the both the sides of the system secure as well as an authenticated one. The concept of image processing and an improved visual cryptography is used. Image processing is a technique of processing an input image and to get the

output as either improved form of the same image and/or characteristics of the input image. In Visual Cryptography (VC) an image is decomposed into shares and in order to reveal the original image appropriate number of shares should be combined [9].

II. RELATED WORK

Automated Challenge Response Method [1] is one such authentication mechanisms, includes challenge generation module from server which in turn interacts with Challenge-Response interface in client and request for response from user. Challenge-Response module in turn will call the get response application which is installed in the client machine. Method ensures two way authentications and simplicity. It prevents man-in-the middle attacks.

DNS-based anti-phishing approach technique which mainly includes blacklists, heuristic detection, the page similarity assessment. But they do have some shortcomings. Blacklist is a DNS based anti-phishing approach technique now most commonly used by the browser [2]. Heuristic-based anti-phishing technique is to estimate whether a page has some phishing heuristics characteristics. For example, some heuristics characteristics used by the SpoofGuard [3] toolbar include checking the host name, checking the URL for common spoofing techniques, and checking against previously seen images.The following technologies used, but they have several drawbacks:

1. Blacklist-based technique with low false alarm probability, but it cannot detect the websites that are not in the blacklist database. Because the life cycle of phishing websites is too short and the establishment of blacklist has a long lag time, the accuracy of blacklist is not too high.
2. Heuristic-based anti-phishing technique, with a high probability of false and failed alarm, and it is easy for the attacker to use technical means to avoid the heuristic characteristics detection.
3. Similarity assessment based technique is time-consuming. It needs too long time to calculate a pair of pages, so using the method to detect phishing websites on the client terminal is not suitable. And there is low accuracy rate for this method depends on many factors, such as the text, images, and similarity measurement technique. However, this technique (in particular, image similarity identification technique) is not perfect enough yet.
4. An offline phishing detection system [4] named LARX, acronym for Large-scale Anti-phishing by Retrospective dataeXploration to counter phishing attacks has been proposed but ends out to be offline and not dynamic to real world events.
5. An anti phishing mechanism [5] using Bayesian approach happens to be better. This framework synthesizes multiple cues, i.e., textual content and visual content, from the given web page and automatically reports a phishing web page. But it also happens to be computationally expensive process.

III. METHODOLOGY

A.(2, 2) VISUAL CRYPTOGRAPHY SCHEME ALGORITHM

One of the best known techniques to protect data is cryptography. It is the art of sending and receiving encrypted messages that can be decrypted only by the sender or the receiver. Encryption and decryption are accomplished by using mathematical algorithms in such a way that no one but the intended recipient can decrypt and read the message.

VCS is a cryptographic technique that allows for the encryption of visual information such that decryption can be performed using the human visual system[8]. We can achieve this by one of the following access structure schemes.

1. (2, 2) - Threshold VCS scheme- This is a simplest threshold scheme that takes a secret message and encrypts it in two different shares that reveal the secret image when they are overlaid.
2. (n , n) -Threshold VCS scheme-This scheme encrypts the secret image to n shares such that when all n of the shares are combined will the secret image be revealed.
3.(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 overlaid the secret image will be revealed.

In the case of (2, 2) VCS, each pixel P in the original image is encrypted into two sub pixels called shares. Figure.1 denotes the shares of a white pixel and a black pixel. Note that the choice of shares for a white and black pixel is randomly determined (there are two choices available for each pixel). Neither share provides any clue about the original pixel since different pixels in the secret image will be encrypted using independent random choices. When the two shares are superimposed, the value of the original pixel P can be determined. If P is a black pixel, we get two black sub pixels; if it is a white pixel, we get one black sub pixel and one white sub pixel [9].

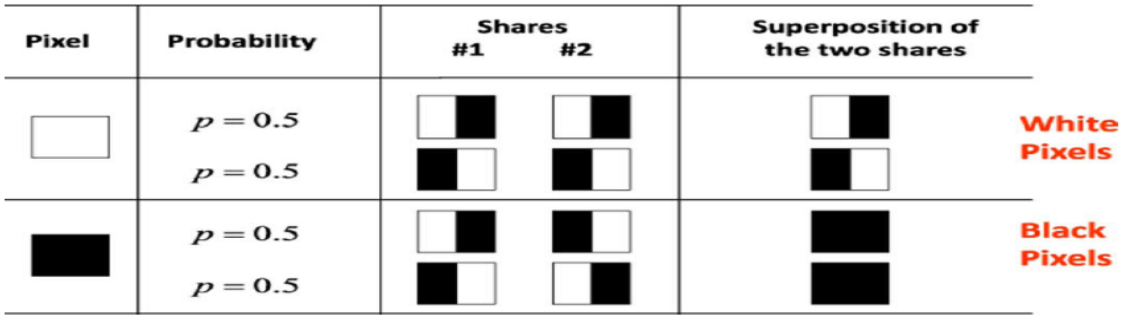

Fig 1: Illustration of a 2-out-of-2 VCS scheme with 2 sub pixel construction.

B. (t, n) SHARE GENERATION VCS ALGORITHM

In this project, we propose a probabilistic model of (t, n) VC scheme with unlimited n. The proposed scheme allows changes of users without regeneration and redistribution of VC transparencies, which reduce the computing and communication resources in accommodating user changes. The scheme is capable of generating an arbitrary number of transparencies and the explicit algorithms are proposed to generate the transparencies. for a group with n' initial users, the proposed Algorithm 1 explicitly generates the required n' transparencies [10]. For n'' newly joining participants, the n'' new transparencies can be explicitly generated through Algorithm 2, and the new transparencies can be distributed to the new participants without the need to update the original transparencies. The secondary contribution is that this paper designs an (t,∞) implementation of VC based on the probabilistic model, and the proposed scheme allows the unlimited number of users. For the conventional VC scheme, the mathematical manipulations of infinite size of basis matrices and variables are often required, which is computationally prohibitive. We also derive an optimization problem L(t) to solve the maximal contrast of the proposed VC scheme.

Notations

The threshold of a (t, n) VC scheme	t
The number of generated transparencies of a (t, n) VC scheme.	n
The number of sub pixels to encode a secret pixel; i.e., the width of the basis matrices.	m
Two vectors to record the nonzero terms in Δp	X,Y
An optimization problem to find the optimal contrast of(t, n→∞) VC scheme.	L(t)
A binary secret image.	S
A ready-to-process pixel taken from	s
The ith generated transparency.	T_i
A pixel at Ti, and the position corresponds to the position of s .	t_i
n' is the number of original participants in the user group initially, and n'' is the number of new participants to join the user group.	n & n''
An index table Z[w,h] where is the index of the used memoryless sequence E(xZ[w,h])to encode the secret pixel s[w,h] .	Z

For a given value of t , the transparencies can be continuously generated with the (t,∞) OptPrVC scheme. However, practical applications require the algorithm to terminate within finite steps. To meet the requirement, a finite number n' is used to specify the number of transparencies in the algorithm. The algorithm requires (X,Y) , obtained by solving(t). The outputs of Algorithm 1 are transparencies and an index table Z , where Z[w,h] is the index of the used memoryless sequence E(xZ[w,h]) to encode the secret pixel s[w,h].

Algorithm 1. The algorithm of (t,∞) OptPrVC scheme
Input: A binary secret image S, two positive integers t,n', and two vectors(X,Y) .
Output: n' transparencies T1,T2,… $T'n$ an index table Z.
1 **for** each pixel s[w,h] in S **do**
2 **if** s[w,h]=white **then**
3 Generate an integer z € {t-2k | k=0,1,…..[t/2]} and P(z=t-2k)= $y_t - 2k$
4 **else**
5 Generate an integer z € {t-1-2k | k=0,1,…..[t-1/2]} and P(z=t-1-2k)=$-y_t - 2k$
6 **end if**
7 Z[w,h]=z.
8 for k=1 **to** n' **do**
9 Assign randomly$T_k[w,h]$ to 0 or 1 where P($T_i[w,h] = 0$)= x_z
10 **end for**
11 **end for**

C. (t,∞) NEXT SHARE GENERATION VCS ALGORITHM

In the first round, we use Algorithm 1 to generate transparencies and Z . If we need not to generate more transparencies in the future, Z is not required and discarded. Otherwise, has to be stored in a safe place, and we can generate more transparencies T1',T2'…….T'n'' by utilizing Z .

Algorithm 2. The algorithm of (t,∞) OptPrVC scheme by the index table Z.
Input :An index table Z , a positive integer n'' , and a vector X .
Output: n'' transparencies $T'_1, T'_2, \ldots T'n''$.

1 **for** each Z[w,h] in Z **do**
2 **for** k=1 **to** n'' **do**
3Assign randomly $T_k[w,h]$ to 0 or 1 where P($T_k[w,h] = 0$)= x_z.
4 **end for**
5 **end for**

IV. DESIGN PROCESS

For phishing detection and prevention, we are proposing a new methodology to detect the phishing website. Our methodology is based on the Anti-Phishing Image Captcha validation scheme using visual cryptography[10]. It prevents password and other confidential information from the phishing websites.It can be used when user lost his share,In this project 2 algorithm is used.1st algorithm is use for creating new share,2nd algorithm is used to for creating next share.
The current approach is divided into three phases:

A. REGISTRATION PHASE

In the registration phase, a key string (password) is asked from the user at the time of registration for the secure website. The key string can be a combination of alphabets and numbers to provide more secure environment. This string is concatenated with randomly generated string in the server and an image captcha[7] is generated. The image captcha is divided into two shares such that one of the shares is kept with the user and the other share is kept in the server. The user's share and the original image captcha is sent to the user for later verification during login phase. The image captcha is also stored in the actual database of any confidential website as confidential data. After the registration, the user can change the key string when it is needed. Registration process is depicted in Figure 2.

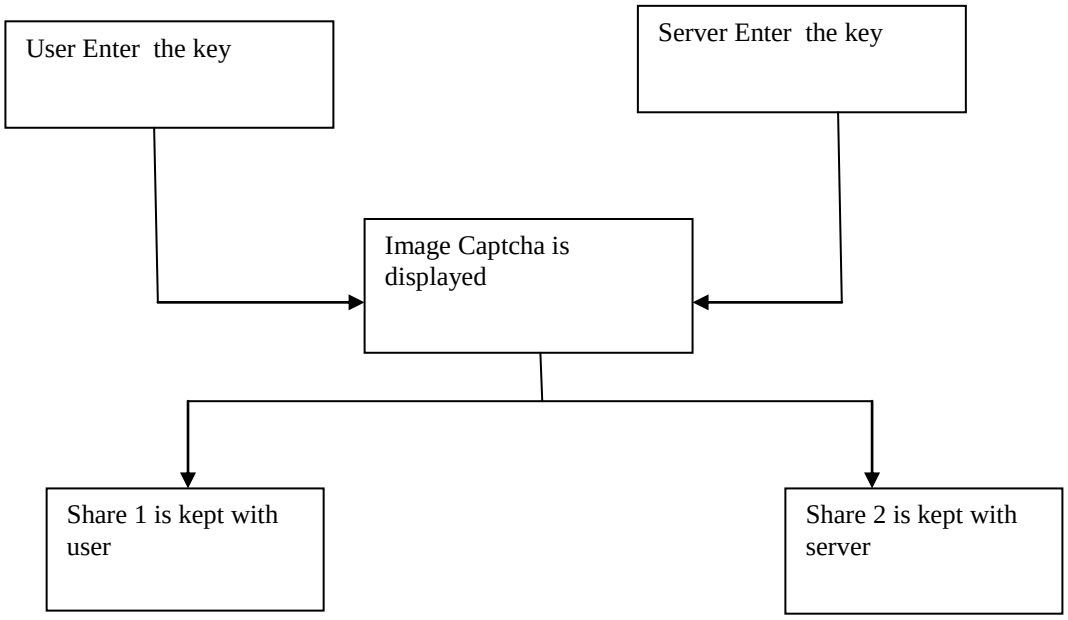

Fig.2. Registration Phase

B. LOGIN PHASE

In the Login phase first the user is prompted for the username (user id).Then the user is asked to enter his share which is kept with him. This share is sent to the server where the user's share and share which is stored in the database of the website, for each user, is stacked together to produce the image captcha[7]. The image captcha is displayed to the user. Here the end user can check whether the displayed image captcha matches with the captcha created at the time of registration. The end user is required to enter the text displayed in the image captcha and this can serve the purpose of password and using this, the user can log in into the website. Using the username and image captcha generated by stacking two shares one can verify whether the website is genuine/secure website or a phishing website and can also verify whether the user is a human user or not. Figure.3 can be used to illustrate the login phase.

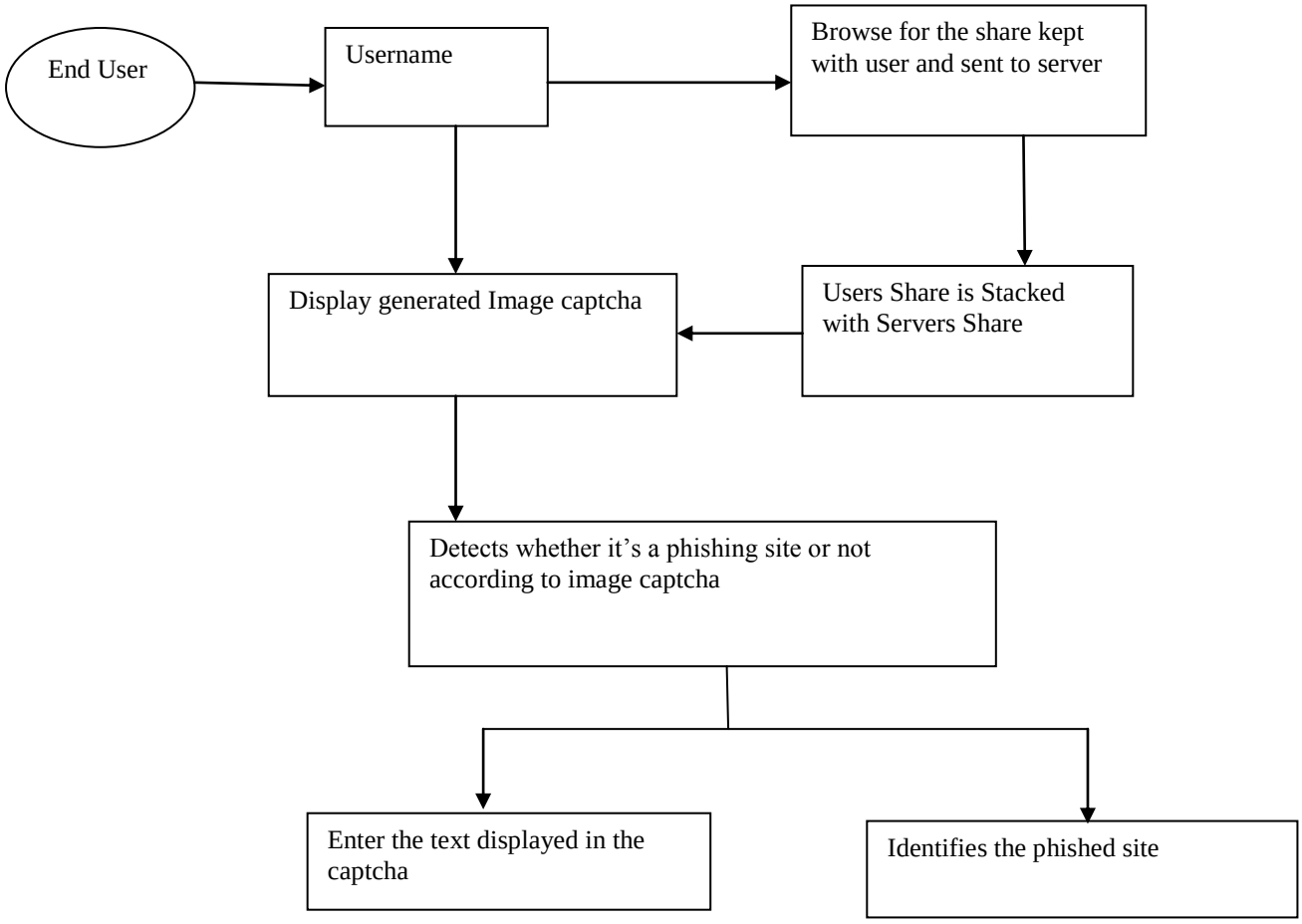

Fig 3.Login Phase

C. SHARE RECOVERY PHASE

The Share recovery Phase is used when user lost or corrupt his share. In the registration Phase when User enter username and try upload his share from server. if she/he lost or corrupt his/her share then he request for new share at that time server crosscheck whether user is authorized or not. server uses next share algoritham (t,∞) for generating new share which is compatible with users share. server generate new share for user.user download new share and process continue with login page is as shown in Figure.4

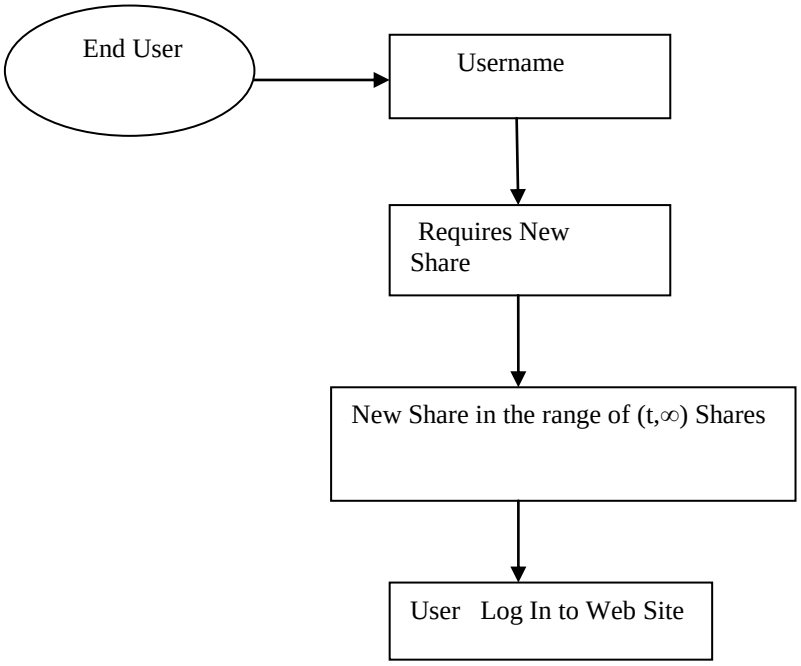

Fig.4 Share Recovery Phase

V. CONCLUSION

The proposed methodology verifies whether the website is a genuine/secure website or a phishing website. It validates image Captcha and ensure that the site as well as the user is permitted one or not. So, using image Captcha technique, no machine based user can crack the password or other confidential information of the users. It also prevents intruders' attacks on the user's account. This method provides additional security in terms of not letting the intruder log in into the account even when the user knows the username of a particular user. It isuseful to prevent the attacks of phishing websites on financial web portal, banking portal, online shopping market. The proposed methodology is used when user lost his share by creating new share.

REFERENCES

[1] Thiyagarajan, P.; Venkatesan, V.P.; Aghila, G.; "Anti-Phishing Technique using Automated Challenge Response Method"', in Proceedings of IEEE- International Conference on Communications and Computational Intelligience, 2010.

[2] Sun Bin.; Wen Qiaoyan.; Liang Xiaoying.; "A DNS based Anti-Phishing Approach," in Proceedings of IEEE-Second International Conference on Networks Security, Wireless Communications and Trusted Computing, 2010

[3] Nourian, A.; Ishtiaq, S.; Maheswaran, M.;" CASTLE: A social framework for collaborative antiphishing databases", in Proceedings of IEEE- 5th International Conference on Collaborative Computing:Networking, Applications and Worksharing, 2009.

[4] Sid Stamm, Zulfikar Ramzan, "Drive-By Pharming", v4861 LNCS,p495-506, 2007, Information and Communications Security - 9th International Conference, ICICS 2007, Proceedings.

[5] Anthony Y. Fu, Liu Wenyin, "Detecting Phishing Web Pages with Visual Similarity Assessment Based on Earth Mover's Distance (EMD)",IEEE Transactions on Dependable and Secure Computing, v 3, n 4, p301-311, October/December 2006

[6] A Novel Anti Phishing Framework Based on Visual Cryptography, 2012, Divya James, Mintu Philip. 2012

[7] A Text-Graphics Character CAPTCHA for Password Authentication Matthew Dailey Chanathip Namprempre

[8] M. Naor and A. Shamir, "Visual cryptography," in Proc. Advances in Cryptography (EUROCRYPT'94), 1995, vol. 950, LNCS, pp. 1–12.

[9] R. Ito, H. Kuwakado, and H. Tanaka, "Image size invariant visual cryptography,"*IEICE Trans. Fundam. Electron., Commun., Comput. Sci.*,vol. 82, pp. 2172–2177, Oct. 1999.

[10] Sian-Jheng Lin and Wei-Ho Chung A Probabilistic Model of (t,n) Visual Cryptography Scheme With Dynamic Group,2012

