International Journal for Rapid Research in Engineering Technology & Applied Science
ISSN (Online): 2455-4723
Vol 3 Issue 3 March 2017

A Novel Approach using both Encryption and Graphical Passwords for Artificial Intelligence Security Threats

Rohit Vibhandik*, Nisha Bhati**
M. Tech Student, B.M. College Of Technology, RGTU, Indore, Madhya Pradesh, India*
Assistant Professor, B.M. College Of Technology, RGTU, Indore, Madhya Pradesh, India**
vibhandik.rohit@gmail.com, nisha.bhatiujn@gmail.com***

Abstract: Most security based systems of primitives rely on the use of complex mathematical problems to be solved by attackers using either brute force of back engineering to decipher confidential information or credentials. The success of such algorithms lies in the fact that how quickly the attacking computation outruns the rate at which the complexity of the algorithm grows.[1] In this paper, a novel multiple encryption security primitive is proposed which is based on captcha for securing confidential paradigms which uses a dynamic modelling of the number of captcha images as well as the number of co-ordinates for securing entries or log-ins. It has been shown that such a mechanism has a much higher level of security compared to previously used mechanisms.

Keywords: Security primitive, Graphical Password Mechanisms, Guessing Attack, Dictionary Attack, hotspots, CaRP, Hard AI problems, Teacher

I Introduction

The most common design mechanism of security mechanisms is employing complex mathematical problems whose computation outruns the rate at which the security mechanisms grow. The growth of Artificial Intelligence has become a challenge in which bots are programmed to detect subtle weaknesses or apparently unnoticeable patterns in the user's security log-ins. In general passwords have remained the norm for securing the confidentiality of various avenues. The major challenge though remains the fact that passwords are vulnerable to online guessing attacks,, brute force attacks and reverse engineering mechanisms thus leading to a possibility of failure of security. Such a human machine interface (HMI) has culminated in the need for a security system which would need human intervention in accessing systems which would be competent enough to thwart back engineering or exhaustive searching algorithms thus making it a hard AI problem.

II Limitations of Previous Mechanisms

Conventional cryptography methods use cryptographic algorithms such as Data Encryption Standard (DES), Advanced Encryption Standard (AES), Rivest Cipher (RS) versions, The RSA algorithms just to name a few. The major challenge that this mechanism faces is the possibility of dictionary attacks or reverse engineering or brute force to find weaknesses in the algorithm used. With computational platforms becoming more advance by the day, the digital hardware employing vary large scale integration becomes capable to compute all possible combinations needed to break an employed algorithm without outnumbering the rate at which the algorithm grows. The following has been depicted by the following graph.

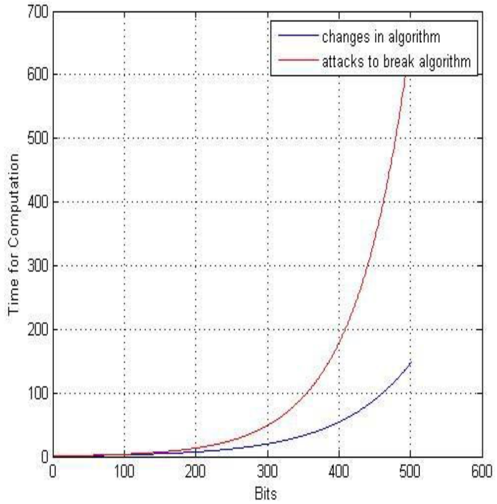


Fig.1 Infeasibility in breaking the cryptographic algorithm

After analyzing the basic characteristics of a cryptographic system, we ca formulate the basic problem strategies pertaining to any cryptographic network. The basic requirements of any cryptosystem can be defined as:

1. Randomness in the encrypted plaintext i.e. the cipher text should exhibit sufficient randomness which makes it difficult for the adversary to attack the plain text.
2. Each bit in the cipher text must change independently of each other so that the attackers cannot relate the cipher bits mathematically
3. It is often observed that the algorithms rendering high level of security also need maximum resources of hardware and time.
4. The above point indicates the necessity of space and time complexity of the algorithm
5. The cipher text should show large variations even for small variations in the plain text.
6. The above point makes the number of trials to break the encryption algorithm enormously large.

Infeasibility in breaking any cryptographic algorithm stems from the fact that the number of trials outnumbers the rate at which the system parameters change. Such a system may be possible to break but infeasible to work upon practically. But such a security primitive may either involve enormous amount of space or time complexity to be practically utilized or be prone to attacks employing Artificial Intelligence.

III Captcha as Graphical Password

As discussed previously, the sole purpose of the cryptographic mechanisms a to make the computation infeasible or design a mechanism which shows the following traits:

1. $Y(i) = f(X(i)) \quad \forall \quad X(i);$
2. But Y (i) varies randomly for $X(i+\Delta);$
3. here Δ is a change in X.

Here, X is the plain text,
Y is the corresponding Cipher Text.
Such a condition may be able to evade dictionary or brute force attacks but it cannot be ascertained.
A captcha on the other hand uses the human capability to extract meaning from random data. The important aspect remains the fact that different humans perceive data differently and the train of thoughts cannot be trained into a machine to show exact similarity. The underlying principle can be understood using the following figure. Although the computer algorithm using character reorganization in conjugation with Teacher Learner Based Optimization strategies or machine learning mechanisms, yet it needs human intervention to extract the correct meaning.

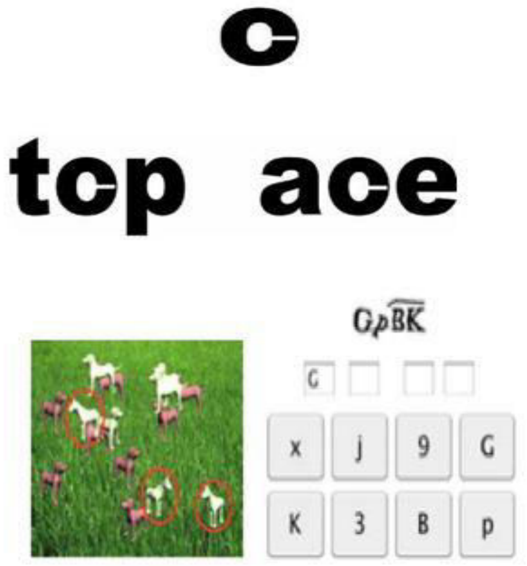


Fig.2 Example Citing Need of Human Intervention in Extraction of Information

The information to be retrieved depends upon the user's perception of the given captcha data. Different users may infer data differently.

IV Proposed Algorithm

The algorithm used for the proposed technique can be explained as:

1) Display the home screen where user login can be done using username.
2) After user logs in, user is activated after admin login.
3) User enters password and captcha code.
4) User selects image coordinates from image.
5) Image can be uploaded after encryption using either AES or Blowfish algorithms.
6) End user can download image (still not decrypted) and after that enter image coordinates.

7) If entered coordinates are correct, image is decrypted.
8) Else, user is blocked and an e-mail is sent to the registered e-mail id of the registered user informing about a security attack.

V Advantages of Proposed Method

The methodology behind the proposed method is based on the fact that the sequential stages of data inference in the case of multiple captcha needs human intervention and cannot be accomplished by programmed bots. Also the critical points of the images need to be specified by the user in terms of coordinates. The critical points of a captcha can be understood using the example of the following figure:

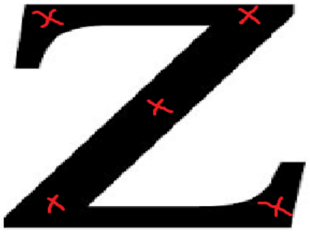


Fig.3 Critical Points of a Captcha Image

If the coordinates specified by the end user are incorrect, then the user is blocked and a mail is sent to the authorized user. Thus the security is increased twofold.

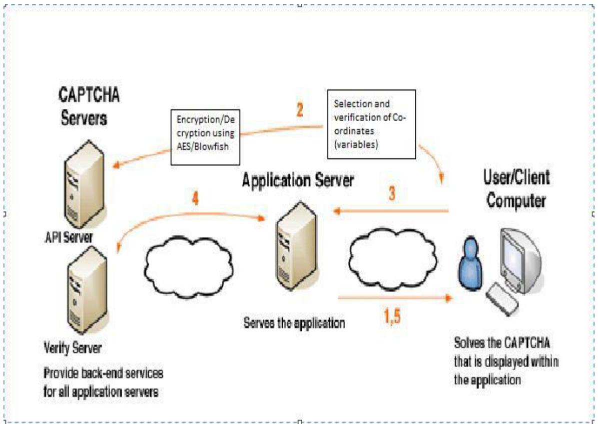

CAPTCHA Servers
Encryption/Decryption using AES/Blowfish
2
Selection and verification of Co-ordinates (variables)
Application Server
User/Client Computer
4
3
API Server
Serves the application
1,5
Verify Server
Provide back-end services for all application servers
Solves the CAPTCHA that is displayed within the application

Fig.4 Architecture Diagram of Proposed Technique

The images uploaded are done through AES/Blowfish whose choice is provided. Thus tampering with the image or its coordinates needs the breaking of AES/Blowfish which are strong encryption algorithms.

Also as the number of captcha images needed for the sequential data inference may be varied, hence the level of difficulty in case of brute force or guessing attacks becomes enormously large. Moreover the number of coordinates can also be varied according to needs thus rendering more randomness in the algorithm.

VI. Results and Performance Evaluation

The results of the proposed algorithm as compared to the previously referred technique can be summarised using the following mathematical formulations.

Mathematical Modelling Of The Proposed Technique.

The effect of adding extra images in the captcha list can be mathematically understood using the following expression. If one among n Images are to be selected by the user, them the total number of combinations can be $\mathbf{{}^{n}C_{1.}}$

Comparing the values of such combinations for 3 Images **(Bin. B. Zhu et al, IEEE Transactions, 2015)** with the proposed technique which has an adaptive image adding feature, we obtain the total probable cases as:

$\mathbf{{}^{3}C_{1}, {}^{6}C_{1}, {}^{7}C_{1}, {}^{8}C_{1,}....}$

Here it has been assumed that the order or the sequencing of the images is immaterial. Has it not been the case, the sequence would have boiled down to:

$\mathbf{{}^{3}P_{1}, {}^{6}P_{1}, {}^{7}P_{1}, {}^{8}{}_{1,}....}$

The progressive randomness increases the security of the proposed mechanism.

The probability of breaking the algorithm by brute force or guessing attacks thus becomes,

$\mathbf{\{1/{}^{3}P_{1}\},\{1/\ {}^{6}P_{1}\},\{1/\ {}^{7}P_{1}\},\ \{1/{}^{8}P_{1}\},....}$

Similarly the increase in the number of coordinates follows the trend of:

2*3, 5*6, 5*7, 5*8....

It can be clearly observed that the decryption of the proposed algorithm by brute force keeps increasing as the number of co-ordinates increases. The benefit of keeping the number of coordinates variable gives us the liberty of deciding the complexity of the system.

The analysis of the previously mentioned concepts can be understood clearly using the following graphical representations. The above results clearly illustrate the fact that the proposed algorithm serves better in terms of security as compared to previously implemented systems

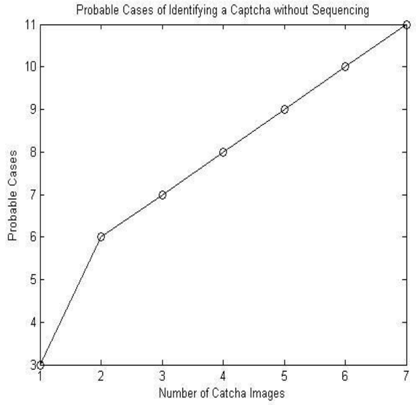


Fig.5 Effect of Adding Additional Captcha Images

It can be seen that the probable cases keeps increasing with the addition of the number of captcha images. It has been assumed t\here that the sequence in which the captcha images are to be selected is immaterial and that only one of the n images are to be selected. The probability of breaking the algorithm by guessing attacks or brute force can be seen from the following graph

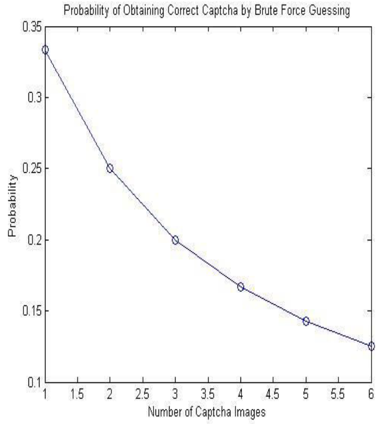


Fig.6 Probability of Obtaining the Correct Captcha by Brute Force or Guessing

The increase in the number of coordinates of the captcha follows a linear curve depicted by:

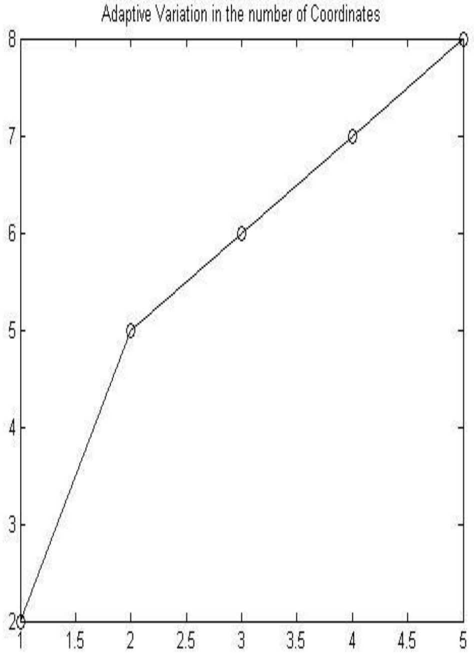


Fig.7 Adaptive Variation in the Number of Coordinates

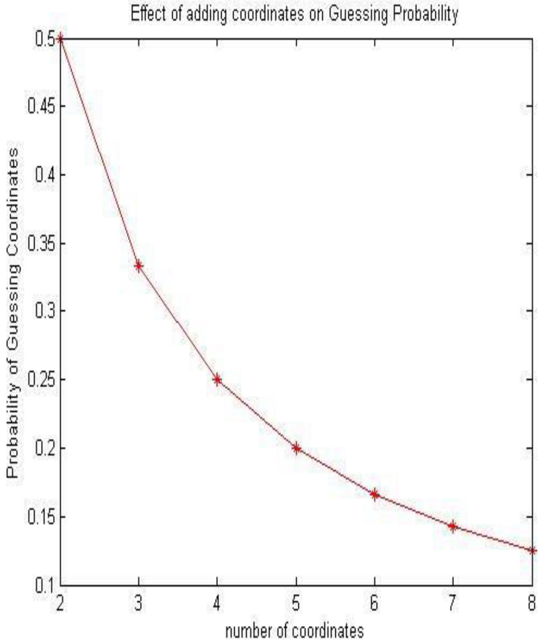


Fig.8 Effect of Adding Coordinates on Guessing Probability

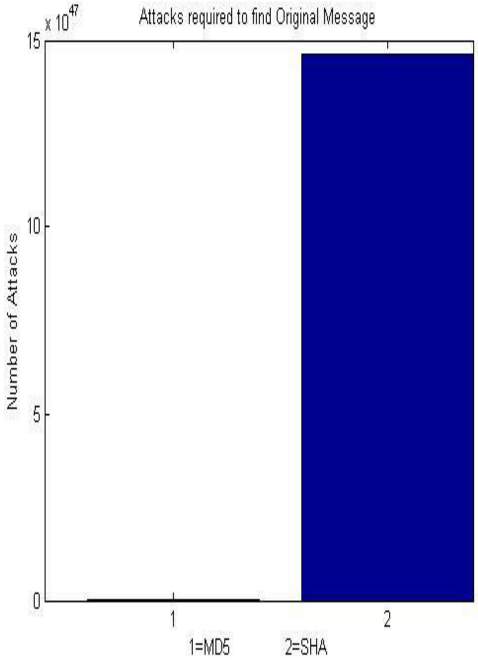


Fig.9 Attacks needed to decipher Cipher Text

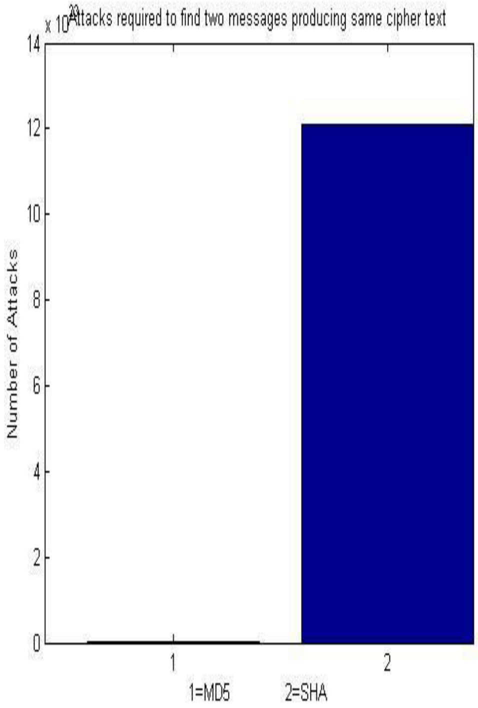
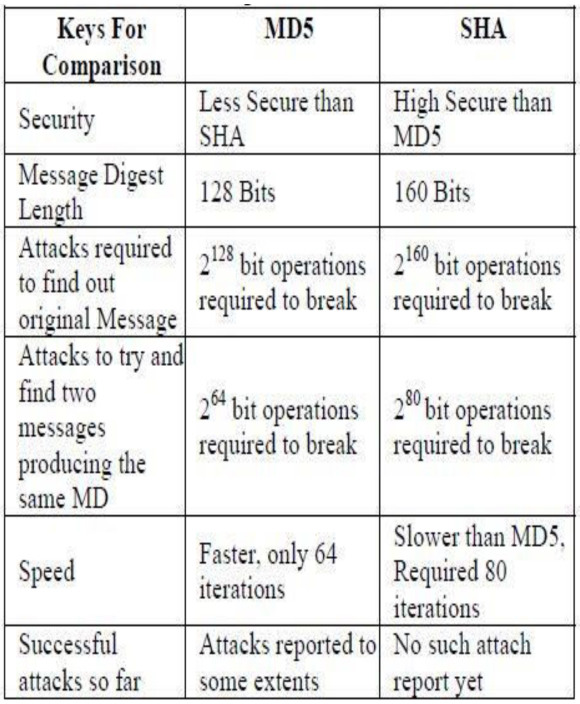


Fig.10 Attacks needed to find two different Messages producing same cipher text

Apart from the above mathematical augmentations, the proposed technique uses the **SHA-1** algorithm instead of the weaker version of **MD-5** used in **(Bin. B. Zhu et al, IEEE Transactions, 2015).**

A ready reference of the differences among MD5 and SHA are given in the figure below:

Table.1
Comparison Table depicting the relevant parameters of MD5 and SHA algorithms.



Keys For Comparison	MD5	SHA
Security	Less Secure than SHA	High Secure than MD5
Message Digest Length	128 Bits	160 Bits
Attacks required to find out original Message	2^{128} bit operations required to break	2^{160} bit operations required to break
Attacks to try and find two messages producing the same MD	2^{64} bit operations required to break	2^{80} bit operations required to break
Speed	Faster, only 64 iterations	Slower than MD5, Required 80 iterations
Successful attacks so far	Attacks reported to some extents	No such attach report yet

Also the image is encrypted and uploaded and decrypted while downloaded by the end user using **AES/Blowfish** hence enhancing security of the algorithm.

Table.2
Prominent Attack Mechanisms on Image Based Captcha Based Systems

S. No	Type of Attack on Captcha
1.	Neural Network Based Attacks: In these type of attacks, neural networks can be trained to attack captchas

2.	Heuristic attacks depend on the findings in the process of reaching a particular result. It relies on monitoring the amount of data requested and monitored, IP addresses visited, pages visited and data entry techniques.
3.	Online Captcha breaking services enable captcha hijacking.

Apart from the above comparison, its also important to look into the common types of attacks on CAPTCHA based systems. As a thumb rule, 20 bots should not have a success rate of greater than 0.01%. It can be checked using IP monitoring.

VII Comparative Time Complexity Analysis of MD5 and SHA

The algorithms used as the backbone of the proposed technique and the previous technique are MD5 and SHA respectively. Therefore a natural question about the time complexity analysis of both the algorithms comes into picture. Generally there are 3 guiding paradigms to evaluate the time complexity of algorithms. They are:

1) **The big O notation**: It describes the Worst Case running time of an algorithm.

2) **The big Ω notation**: It describes the Best Case running time of an algorithm.

3) **Big θ notation**: It represents the Asymptotic efficiency of an algorithm. Big-O is an upper bound for a function and Big-Omega is a lower-bound. If the two bounds agree, then the function is sandwiched and is Big-Theta of the common function.

It is important to mention here that its not always feasible to evaluate the time complexity of algorithms using the Big O or Big Ω or Big θ notations since breaking down an algorithm into the **Tight Bounds** of dependent known evaluated mathematical functions is infeasible or extremely complex. Therefore the way out for such cases is evaluating the overall mean time required for the execution of the algorithm rather than trying to decompose it into interlinked dependencies of known mathematical functions which have a well defined time complexity analysis.

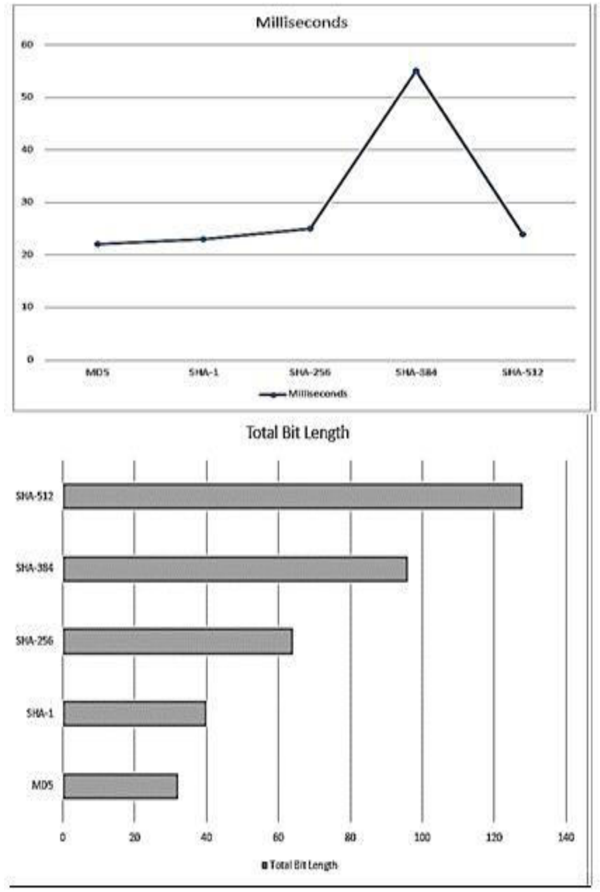


Fig.11 A Comparative Execution Time Complexity Analysis of MD5 and SHA

It is worth mentioning that there have been devastating collision attacks on **MD5.** To cite just one, in 2013, an attack by Xie Tao, Fanbao Liu, and Dengguo Feng broke MD5 collision resistance in 2^{18} time. This attack runs in less than a second on a regular computer. No such attack has thus been found on SHA. Therefore conclusively it can be said that SHA is much more secure compared to MD5.

Limitations:

The limitations of the proposed technique can be summarized as follows:

1) With increasing number of coordinates in the image, the complexity and time of execution would also increase.

2) Although it is assumed that machine learning algorithms do not possess cognitive capabilities at all, still it's a misconception since complex machine learning techniques working in synchronism with deep learning can extract features accurate enough to break the Captcha security.

VIII. Conclusion

It can be seen from the previous discussions that the proposed technique achieves much higher level of randomness compared to previous techniques. The effect is the plummeting probability of breaking the algorithm by brute force or guessing attacks. It is evident that such a mechanism needs human intervention and hence pre programmed bots cannot be used to break the algorithm. Moreover employing SHA in place of MD5 enhances the security of the algorithm. Using AES/Blowfish for encryption of the data leads to much lesser vulnerability in eavesdropping or modifications in the coordinates of the captcha images. Thus the proposed technique can be used as an effective tool while dealing with hard Artificial Intelligence problems.

References

[1] Captcha as graphical passwords-A new security primitive based on hard AI problems.
IEEE Transactions on Information forensics and security, vol.9 June 2014.

[2] M. Motoyama, K. Levchenko, C. Kanich, D. McCoy, G. M. Voelker, and S. Savage, "Re: CAPTCHAs—Understanding CAPTCHA-Solving Services in an Economic Context," in Proc. USENIX Security, 2010, pp. 435-452.

[3] M. Szydlowski, C. Kruegel, and E. Kirda, "Secure input for web applications," in Proc. ACSAC, 2007, pp. 375-384.

[4] G. Wolberg, "2-pass mesh warping," in Digital Image Warping. Hoboken, NJ, USA: Wiley, 1990.

[5] HP TippingPoint DVLabs, New York, NY, USA. (2011). The Mid-Year Top Cyber Security Risks Report [Online]. Available: http://h20195.www2.hp.com/v2/GetPDF.aspx/4AA3-7045ENW.pdf

[6] S. Kim, X. Cao, H. Zhang, and D. Tan, "Enabling concurrent dual views on common LCD screens," in Proc. ACM Annu. Conf. Human Factors Comput. Syst., 2012, pp. 2175-2184.

[7] S. Li, S. A. H. Shah, M. A. U. Khan, S. A. Khayam, A.-R. Sadeghi, and R. Schmitz, "Breaking e-banking CAPTCHAs," in Proc. ACSAC, 2010, pp. 1-10.

[8] H. Gao, X. Liu, S. Wang, and R. Dai, "A new graphical password scheme against spyware by using CAPTCHA," in Proc. Symp. Usable Privacy Security, 2009, pp. 760-767.

[9] L. Wang, X. Chang, Z. Ren, H. Gao, X. Liu, and U. Aickelin, "Against spyware using CAPTCHA in graphical password scheme," in Proc. IEEE Int. Conf. Adv. Inf. Netw. Appl., Jun. 2010, pp. 1-9.

[10] J. Bonneau, "The science of guessing: Analyzing an anonymized corpus of 70 million passwords," in Proc. IEEE Symp. Security Privacy, Jun. 2012, pp. 20-25.

[11] John the Ripper Password Cracker [Online]. Available: http://www. openwall. com/john

[12] Openwall Wordlists Collection [Online]. Available: http://www. openwall.com/wordlists.

[13] K.Krol et.al, Simple Generic Attacks on Text Captchas, 2016 researchgate.com

[14] Piyush Gupta and Sanjay Gupta, A Comparative Analysis of MD5 and SHA algorithms, IJCSIT, 2014.

[15] N Roshanbin, J Miller – Future Generation Computer Systems, ADAMAS: Interweaving uni code and colour to enhance CAPTCHA security 2016, Elsevier

[16] M Khan, T Shah, SI Batool - Signal, Image and Video Processing ,A New Implementation Of Chaotic S-Boxes In CAPTCHA, 2016 Spring.