# A Case Study of Text-Based CAPTCHA Attacks

Xiao Ling-Zi
Dept. of Digital Media Technology
Communication University of China
Beijing, China
E-mail: ttkdn_ling@cuc.edu.cn

ZHANG Yi-Chun
Dept. of Digital Media Technology
Communication University of China
Beijing, China
E-mail: zhangyichun@cuc.edu.cn

***Abstract*—CAPTCHA (Completely Automated Public Turing Test to tell Computers and Human Apart) is widely used than before, which becomes the common part of current website login system. However, the CAPTCHA implementation is tricky and risky without deliberate design. In this paper, we give a study case of the vulnerabilities in current login website using text-based CAPTCHA. Our target is a website of mainstream bank of china. We show that with some specialized methods, the CAPTCHA scheme in its website can be easily cracked. Finally, we give some advices for CAPTCHA designers to revise our CAPTCHA implementation security in the future.**

***Keywords- CAPTCHA (Completely Automated Public Turing Test to tell Computers and Human Apart); security; CAPTCHA attacks; Image segmentation***

## I. INTRODUCTION

CAPTCHA (Completely Automated Public Turing Test to tell Computers and Human Apart)[1][2], also known as Human Interactive Proof (HIP), was firstly developed to solve the "chat room" problem for Yahoo. The “chat room” problem is that automatic program designed to talk with humans at the chat rooms, market products, and gathering personal information. Luis Von Ann introduce an new idea [2] using hard AI (Artificial Intelligence) problems to distinguish human from machine, which the problems are easy for people to solve but difficult for computer programs. Nowadays, CAPTCHA are widely used in bot detection [3], anti-spam [4], and web authentication.

The text-based CAPTCHA has been widely used because of the easily accessible and rich image resource on the Internet. However, the CAPTCHA implementation is tricky and risky without deliberate design. As far as we know, there’re many attacks against text-based CAPTCHA schemes, many of which resulted from careless design and security overlook.

In this paper, we propose a CAPTCHA scheme on a mainstream bank website of china as a study case to analyze the existed security weaknesses in it, and show a detailed attack process step by step. Our case study can be used for CAPTCHA designers to build a robust CAPTCHA scheme. For such a purpose, we suggest the rules for CAPTCHA construction.

The paper is organized as follows. In section Ⅱ, we analysis the security vulnerabilities of our target CAPTCHA scheme. The detailed attack process is introduced in section Ⅲ. The results and its related analysis are shown in section Ⅳ. Finally, we give our conclusions for CAPTCHA design.

## II. VULNERABILITIES ANALYSIS

Our target is the text CAPTCHA which is collected from the website of one of the biggest bank in China. The bank is famous and influential in mainland china, however the CAPTCHA scheme is not secure as it seems. Fig 1 gives a sample of the bank CAPTCHA scheme.

Fig 1. Sample of the CAPTCHA of the bank

From those CAPTCHA based challenge images from the bank website, we found that the webpage is transmitted using HTTPS (Secure Hypertext Transfer Protocol) transmission protocol. It’s one of the methods that they use to protect their webpage. However, we still can get CAPTCHA image by saving the entire page in the local. We use tools of browser (any main stream browser has this function) to see the source code of the website and find a fixed JaveScript link[23], to the CAPTCHA image. Then we find the file with fixed name ” com.icbc.inbs.person.servlet.Verifyimage2” in those files we saved from the bank website. We analyze the contents of the fixed name file. Fig 2 shows the structure of the file.

```
          0  1  2  3  4  5  6  7
00000000h: FF D8 FF E0 00 10 4A 46
00000010h: 00 01 00 00 FF DB 00 43
00000020h: 07 07 07 09 09 08 0A 0C
00000030h: 13 0F 14 1D 1A 1F 1E 1D
00000040h: 22 2C 23 1C 1C 28 37 29
```

Fig 2 Target File Data Structure Analysis

It can clearly see this file’s identifier is "FF D8". It comes to that it is actually a jpeg image file. If we change its file extension to ”.jpg” , we can get the CAPTCHA image.

We randomly download 500 CAPTCHA challenge images, and find the following features:

978-0-7695-4810-4/12 $26.00 © 2012 IEEE
DOI 10.1109/CyberC.2012.28



IEEE computer society

- The character amount is fixed. Four characters are used in every challenge and only lower-case letters are used.
- Characters in one challenge image are different in color. That means the regular attack based on color track will be difficult.
- There's a significant difference in pixel between the foreground characters and the background image.

According to these features above, we could easily know several vulnerabilities the CAPTCHA have:

- The significant difference between foreground and background in pixel is weak to binarization attack. This attack converts a rich-color CAPTCHA image to a black-white image using a threshold method;
- The characters are not segmentation-resistant enough by using tilt techniques only. The CAPTCHA will be easily compromised if segmenting attack occurs.
- Only using standard character gives convenience to late recognition.

## III. THE PROCESS OF BREAKING CAPTCHA

In this section, we introduce how to design the attack system on the basis of vulnerabilities of the scheme and what we actually do in each step.

Fig 3 illustrate the attack procedures.

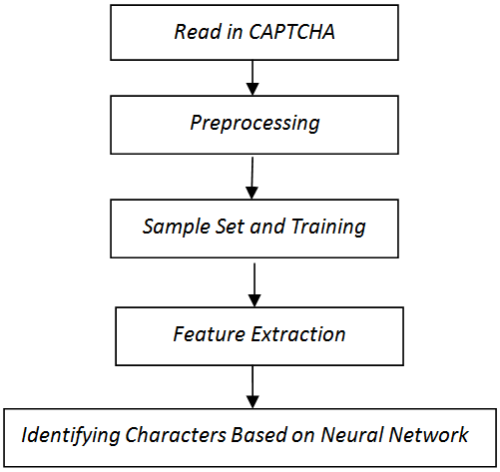


Fig 3. The Attack Flow Chart

We explain each step as below.

### A. *Preprocessing*

We first convert a rich-color CAPTCHA image to a black-white image using a threshold method: pixels with intensity higher than a threshold value are converted to white, and those with a lower intensity are converted to black. ( see Fig. 4(a) and (b)). The threshold is manually determined by analyzing the sample set, and the same value is used for each image in both the sample and test sets.

(a) (b)

Fig 4. Preprocessing.
(a) Original image, (b) Binarized image

After the binarization, the foreground characters and the background are separated. However, the edges of some characters are blurred and the image SNR (Signal-Noise-Ratio) decreases .The second step of preprocessing is image filtering in order to smooth edge and eliminate noise. We use an average filter [6] to remove edge burr. Fig 5 shows the basic theory of average filtering and Fig 6 shows the result of average filtering.

| | | |
|---|---|---|
| 1 | 1 | 1 |
| 1 | 10 | 1 |
| 1 | 1 | 1 |

| | | |
|---|---|---|
| 1 | 1 | 1 |
| 1 | 2 | 1 |
| 1 | 1 | 1 |

(a) (b)

Fig 5. (a)before processing;(b)after average filtering

(a)Before filtering (b)After filtering

Fig 6 The result of average filtering

### B. *Vertical Segmentation*

A vertical segmentation method is applied to segment a CAPTCHA vertically into several characters after preprocessing. According to observing the binarized image, the following simple method works well to segment each character. The vertical segmentation steps are described as below:

- The first step is to map the image to a histogram that represents the number of foreground pixels per column in the image.
- It then scan the histogram from left to right, vertical segmentation lines separately the image into single character by cutting through columns that have no foreground pixels at all and the neighbor columns have foreground pixels. This step achieves the left and right margin of each character.
- For each segmented area done by the previous step, we scan from the top and the bottom respectively and find the top and the bottom character margin in selected area.
- For some special characters which is hard for vertical segmentation, such as "j", "i", we choose an empirical method to separate a target area containing two characters. We predefine the width of character "i" and "j" , and set another character width equals

to difference between target area's width and the predefined width.

- In the end, we uniformize each separated character to 64*32 in pixels.

Fig 7 shows the segment result of the scheme.

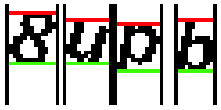

Fig 7. Character Segmentation

### C. Sample Set and Training

In order to get the training sample set, the normalized character image in each rectangular area is saved as separated image files. In the input encoding, we use PCA [7][8]to reduce the dimensionality of the original image, and use the dimensionality reduction of a uniform length feature vector as the input. The dimension of the feature vector of this method encoded image equals to the product of height and width number of pixel that means the characteristic dimension is 64*32.

In the output encoding for an n-class problem, we choose the same class number n which equals to the number of output units. Each output corresponding to the category label, training for the i class sample i output unit sets a upper limit value (0.9), while the other unit is value (0.1) as the target output. The test takes the high output unit number as the predicted values of network. This program has total 36 output units (include 26 english letters and 10 arabic numerals).

Finally, we consider two situations that are the training image preprocessing and the test image preprocessing: when dealing with the training image under the category of each number or letter, they're categorically saved to the training classification file folder (we have already set 36 new different folders).

### D. Feature Extraction

This is to extract the useful data or information from the image such as descriptions, values, vectors, and symbols. Feature extraction is employed.

We desire to find out a feature which won't be changed in resize and rotation of the characters. The two-dimension invariant moments [9] are another group of features we have chosen which are not allergic to translation, scale and rotation [10].

The grayscale of image is a two-dimensional or three-dimensional density distribution function, and then the invariant moment method can be used in the field of image analysis and used for image feature extraction.

### E. Identifying Characters Based on BP Neural Network

We chose the BP neural network [12] to achieve the function of character recognition [13][14][15]. And the number of layers in the network, the number of neurotic units in each layer, the initial value and the learning rate, etc. should be considered in the design of BP neural network [20].

- The number of layers in the network: With at least one S-type hidden layer and a linear output layer network can approximate any rational function. Increasing the number of layers can further reduce the error, but makes the network become more complicate at the same time. So we decide to increase the number of hidden layer neurons to increase the accuracy.
- The number of neurotic units in each layer: Select of the number neurotic units in hidden layer nodes in the theory is not clear identified. We can decide certain quantity through trained compared.
- The initial value: Because the system is nonlinear, in order to ensure that the weight of each neuron is able to make adjustments in its activation function that changes the maximum. The initial value is a random number between (-1, 1). Usually choose the right value of the order of magnitude $\sqrt[5]{S1}$. S1 is the number of first layer neurons.
- The learning rate: In order to avoid a too large learning rate that may lead to oscillation and make the system unstable, we tend to select a smaller learning rate in order to ensure the stability of the system. General values between 0.01-0.8.

## IV. RESULTS

After a number of experiments, by adjusting the parameters of the BP network, we find that ideal identification rate is of 75%. However, the recognition rate is still rising. We analyzed all cases of failure of our attack in both the sample and test sets. We find several aspects that can improve in our system:

- Since all samples in the training sample libraries are not standard training library samples. Therefore, these preprocessing and segmentation of the sample have been different.
- In our attack, we use the BP neural network as a classifier to achieve recognition. Recently, the other classifiers, such as: the SVM[5][18] support vector machines, convolution neural network[19] recognition function can be used as the classifier. For instance, the trait of convolution neural network without the character segmentation will be able to identify after sample training.

## V. CAPTCHA ENHANCEMENT

### A. Apply our results to the Current text CAPTCHA Design

According to our case study and some literatures [21][22], we conclude that the most important part of the text-based CAPTCHA attacks is image segmentation and image recognition. Some key points should be considered in current CAPTCHA design:

- We can apply some anti-segment technologies to characters such as adhesion, interlace and using different font styles to increase the difficulty of character segmentation.
- There are many other things we can do to background to increase the difficulty of foreground extraction. Such as complex background, blurred, complex noise background to enhance the CAPTCHA anti-segment ability to conventional binarization attack or edge extraction attack .
- We can do asymmetric transformation and size transformation to characters so that it is more difficult to network getting sample collection. By that time the time of sample collection and the neural network judgment will have greatly increased.

### *B. Open issues on CAPTCHA design*

In order to eliminate the possibility of a computer attack, there is still a lot we could do. CAPTCHA can be further improved:

- We can enhance human participation and make full use of our human brain, such as simple mathematic test, sentence and image understanding, which could hardly learned by computer vision.
- We could use human other senses except visual decryption system and sound filtering system for CAPTCHA application. Current CAPTCHA have not realized how complicated human brain is. Other human senses, including touch, smell, quick learning or logical analysis ability, did not be exploited by CAPTCHA designer.
- We could create new form[17] for CAPTCHA application. With the development of the media technology, mobile phones and smart televisions are competing for users with computer. Creating CAPTCHA for other forms of media becomes necessary than before. Then other media would be protected by authentication.

## ACKNOWLEDGMENT

Thanks for the supporting of CUC Research Project (XNG1112).

## REFERENCES

[1] Ahn, L. Von. "Human Computation", Ph.D. Dissertation, Carnegie Mellon University, CMU-CS-05-193. (2005)

[2] Ahn, L. Von., Blum M., Hopper N "CAPTCHA: Telling humans and computers apart automatically", communication of the ACM, Vol. 47, No.2. (2004)

[3] Kang H.W., Wang K.S., Soukal D. Behr F., Zheng Z.J., "Large-scale bot detection for search engines", Proc. 19th international conference on World Wide Web (WWW'10), ACM, pp. 501-510. (2010)

[4] B.S., Choudhury A.J., Lee, Kim T.Y., Lee H.J.,"An efficient password authentication method using CAPTCHA", Proc. 5th international conference on Convergence and hybrid information technology(ICHIT'11), pp. 456-463. (2011).

[5] Cortes C. and Vapnik V.,"Support-vector networks", Machine learning, Vol. 20, Issue 3, pp. 273–297.(1995).

[6] WU Ying, WU Hai-yong "Adaptive mixed filtering method for removing image noise" COMPUTER ENGINEERING AND APPLICATIONS, 2010,46(7)

[7] Park B.S., Choudhury A.J., Lee, Kim T.Y., Lee H.J.,"An efficient password authentication method using CAPTCHA", Proc. 5th international conference on Convergence and hybrid information technology(ICHIT'11), pp. 456-463. (2011)

[8] Kim Esbensen, Paul Geladi "Principal component analysis", Chemometrics and Intelligent Laboratory Systems,Volume 2, Issues 1–3, August 1987, Pages 37–52,2001

[9] Jan Flusser, Tomas Suk "Pattern recognition by affine moment invariants" Volume 26, Issue 1, January 1993, Pages 167–174

[10] Ming-Kuei HU, IRE "Visual Pattern Recognition by Moment Invariants" , IRE Transactions on Information Theory, Volume: 8 , Issue: 2 ,Page(s): 179- 187 1962

[11] Greg Mori, Jitendra Malik, Computer Science Division Computer Science Division University of California, " Recognizing Objects in Adversarial Clutter:Breaking a Visual CAPTCHA"

[12] Nielson, HR, Theory of the Backpropagation Neural Network, IEEE IJCNN, 1989. 1, 593~ 606

[13] Shifei Ding, Weikuan Jia, Chunyang Su, Xiaoliang Liu, Jinrong Chen," An Improved BP Neural Network Algorithm Based on Factor Analysis" Journal of Convergence Information Technology,Volume 5, Number 4, June 2010

[14] Yang Wei, Han Chuncheng," Improving the Function of BP Neural Networks Node and Apply the Improved Algorithm to Learn Library of Chinese Character"(in Chinese), Journal. Changchun, Inst. Opt. &Fine Mech.

[15] Widrow B, Mccol J, and Ball M. T he Complex LMS algorithm, pro c. IEEE 1995, 63: 719~ 720

[16] Zhang H.G, Hu Y., "Jigsaw puzzle: Construction and evaluation for scheme of CAPTCHA", Computer Engineering and Deign, Vol 31 (12), 2010

[17] Chow R., Golle P., Jakobsson M., Wang L.S., Wang X.F.," Making CAPTCHAs clickable", HotMobile '08, Proc. 9th workshop on Mobile computing systems and applications, ACM, pp 91-94. (2008)

[18] Golle, P. "Machine learning attacks against the Asirra CAPTCHA.", Proc. in Computer and Communication Security (CCS'08), ACM, pp.535-542 (2008)

[19] LeCun Y., Bottou L, Bengio Y., Haffiner P., "Gradient-based learning applied to document recognition", Proc. IEEE, Vol. 86, Issue 11, pp. 2278-2324. (1998)

[20] Tian Hua ,The Information engineering school of Yantai Nanshan College, longkou ,Shandong "Performing BP Caculation"

[21] Elie Bursztein, Matthieu Martin, John C. Mitchell "Text-based CAPTCHA Strengths and Weaknesses", Proceedings of the 18th ACM conference on Computer and communications security, New York, NY, USA 2011

[22] Ahmad S El Ahmad, Jeff Yan, Mohamad Tayara," The Robustness of Google CAPTCHAs", Proceedings of the Third European Workshop on System Security, New York, NY, USA ,2010

[23]/servlet/com.icbc.inbs.person.servlet.Verifyimage2?randomKey=13379 9165516520792&imageAlt=Click image to refresh（in Chinese）