Chimera CAPTCHA: A Proposal of CAPTCHA Using Strangeness in Merged Objects

Masahiro Fujita, Yuki Ikeya, Junya Kani,
and Masakatsu Nishigaki(✉)

Graduate School of Informatics, Shizuoka University, Hamamatsu, Japan
nisigaki@inf.shizuoka.ac.jp

Abstract. In this paper, we propose "Chimera CAPTCHA" that requests users to select only a *chimera object*, merged from two 3D objects, in a question image, which consists of some 3D objects and the chimera object. The Chimera CAPTCHA is easy for humans to solve because chimera objects, whose appearance are different from ones judged by common sense, cause a feeling of strangeness. Usability survey suggests that the correct response rate is 90.5 % and the average response time is about 5.7 s. In addition, the CAPTCHA system is able to generate questions countlessly and easily by using 3DCG technologies. We also describe threats to its security.

Keywords: CAPTCHA · A feeling of strangeness · Automatic generation · 3DCG

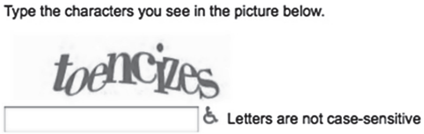
1 Introduction

With the expansion of Web services, denial of service (DoS) attacks by malicious automated programs (malware) are becoming a serious problem. The Turing test plays an important role in discriminating humans from malicious automated programs and the Completely Automated Public Turing test to tell Computers and Humans Apart (CAPTCHA) [1] system developed by Carnegie Mellon University has been widely used.

Today most Web sites adopt text recognition based CAPTCHA (Fig. 1) or image-based CAPTCHAs such as Asirra (Fig. 2) [2] and YUNiTi's CAPTCHA (Fig. 3) [3] to protect themselves from malware. Many researchers, however, have recently pointed out that the malware with Optical Character Reader (OCR) and/or machine learning could solve those CAPTCHAs [4–6]. In order to take measures against these malware, CAPTCHAs using a human higher recognition ability are needed [7–9]. However, it is not such a straightforward process. This is because CAPTCHAs contain a contradiction: Web servers (computers) should be able to automatically generate the CAPTCHA questions that malware (computers) cannot answer.

Actually, the authors have also proposed the CAPTCHAs using "feeling that something is wrong", one of human higher recognition abilities [8, 9]. When a human faces an unusual situation different from a situation judged by common sense, the scene

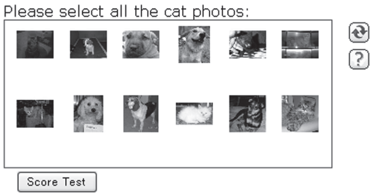
T. Tryfonas and I. Askoxylakis (Eds.): HAS 2015, LNCS 9190, pp. 48–58, 2015.
DOI: 10.1007/978-3-319-20376-8_5



Type the characters you see in the picture below.

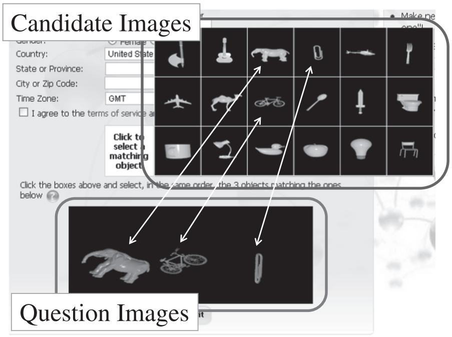
Letters are not case-sensitive

Fig. 1. CAPTCHA used by google



Please select all the cat photos:

Fig. 2. Asirra



□ I agree to the terms of servic

Click to select a matching object

Click the boxes above and select, in the same order, the 3 objects matching the ones below

Fig. 3. YUNiTi's CAPTCHA

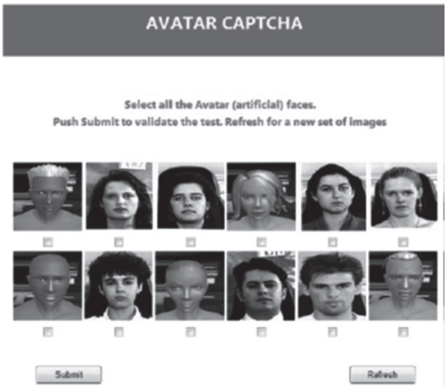
causes a feeling of strangeness such as unnaturalness or uncanniness. The more we gain experience of something, the better this ability becomes. On the other hand, since even state-of-the-art malware have no common sense, it is expected that malware cannot imitate this ability. Therefore by using "a feeling of strangeness" questions that humans can solve, but malware cannot, might be realized. However, in literatures [8, 9], there remains a problem of difficultly in generating questions automatically.

This paper proposes a new image-based CAPTCHA, called "Chimera CAPTCHA", which satisfies two requirements: (i) using a human higher recognition ability for improving attack tolerance and (ii) easily generating questions automatically. This CAPTCHA consists of some 3D objects and a *chimera object* generated by merging two 3D objects. If a user clicks only chimera objects in a question image, the system identifies the user as a human. It is easy for humans to identify chimera objects, because each chimera object, whose appearance is different from ones judged by common sense, causes a feeling of strangeness.

The organization of this paper is as follows. Section 2 describes related works using "feeling that something is wrong". Section 3 introduces Chimera CAPTCHA and Sect. 4 shows basic experimental results of the CAPTCHA. Section 5 discusses the effectiveness of the CAPTCHA. Finally, Sect. 6 presents our conclusions.

2 Related Works

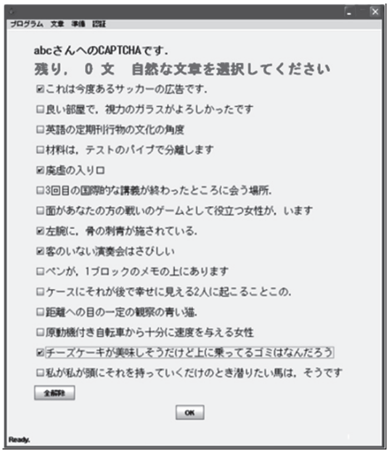
So far, there have been several CAPTCHAs using "feeling that something is wrong". One is Avatar CAPTCHA [10], which requests users to identify avatar faces from a set of 12 images consisting of human and avatar faces (Fig. 4). According to the phenomenon known as "The Uncanny Valley [11]", humans feel uncanniness for avatar faces. That means it is easy for humans to select only avatar faces from the image set. However the authors of [12] show that using modern object recognition and machine learning, malware can also solve Avatar CAPTCHAs as often as humans can.



AVATAR CAPTCHA

Select all the Avatar (artificial) faces.
Push Submit to validate the test. Refresh for a new set of images

Fig. 4. Avatar CAPTCHA



abcさんへのCAPTCHAです．

残り， 0 文 自然な文章を選択してください

☑これは今度あるサッカーの広告です．

□良い部屋で，視力のガラスがよろしかったです

□英語の定期刊行物の文化の角度

□材料は，テストのパイプで分離します

☑廃虚の入り口

□3回目の国際的な講義が終わったところに会う場所．

□面があなたの方の戦いのゲームとして役立つ女性が，います

☑左腕に，骨の刺青が施されている．

☑客のいない演奏会はさびしい

□ペンが，1ブロックのメモの上にあります

□ケースにそれが後で幸せに見える2人に起こることこの．

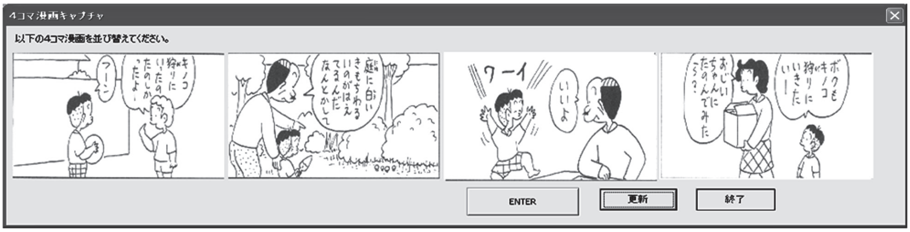
□距離への目の一定の観察の青い猫．

□原動機付き自転車から十分に速度を与える女性

☑チーズケーキが美味しそうだけど上に乗ってるゴミはなんだろう

□私が私が頭にそれを持っていくだけのとき滑りたい馬は，そうです

Fig. 5. SS-CAPTCHA



4コマ漫画キャプチャ

以下の4コマ漫画を並び替えてください。

キノコ狩りにいったのたのしかったよ!
フーン
庭に白いきもちわるいのがはえてるんだなんとかして
ワーイ
いいよ
ボクもキノコ狩りにいきたい!
おじいちゃんにたのんでみたら?
ENTER
更新
終了

Fig. 6. Four-panel cartoon CAPTCHA (Source: From left: 1st image: 1st panel of four-panel cartoon on p.25 of bibliography [13]; 2nd image: 4th panel; 3rd image: 3rd panel; 4th image: 2nd panel of the cartoon).

Yamamoto et al. proposed SS-CAPTCHA [8], which requests users to distinguish natural sentences by humans from machine-translated sentences (Fig. 5). Although current machine-translation techniques have progressed a great deal, it is impossible even for state-of-the-art machine translators to automatically generate perfect natural sentences that will not make a human feel as though something is wrong. That means it is too difficult for malware to identify whether a sentence is natural or not. However, there has not been a framework for self-production of a sufficient number of natural sentences.

Another interesting CAPTCHA is Four-panel Cartoon CAPTCHA [9], which requests users to arrange the four-panel cartoon rearranged randomly in the correct order (Fig. 6). The four-panel cartoon rearranged randomly causes a feel of strangeness. Then it is easy for humans to arrange the panels in the correct order, since humans can understand the meaning of the pictures and utterances in each panel and hidden humor in the cartoons. Four-panel Cartoon CAPTCHA needs a large volume of four-panel cartoons for generating questions. However, as well as SS-CAPTCHA, Four-panel cartoon CAPTCHA also has the difficulty in self-producing a large number of four-panel cartoons.

3 Chimera CAPTCHA

3.1 Concept

Humans have commonsense through their daily life. “Feeling that something is wrong” is caused by an unusual situation different from a situation judged by common sense. So far computers with common sense have not been realized. Thus feeling that something is wrong must be one of human higher recognition abilities that malware cannot imitate. CAPTCHA systems are able to identify a user, whether a human or malware, by asking a user to discriminate “objects judged by common sense” from “objects different from ones acknowledged by common sense”.

However, representing “objects judged by common sense” and “objects different from ones acknowledged by common sense” have been inherently very difficult. To solve this problem, this paper uses 3D models. As services using 3D models have been increasing rapidly in recent years, an indefinite number of 3D models would appear in the future. Most 3D models are generated by modeling objects from the real world. In it, humans are supposed to have seen “the objects that are subject to modeling” at least once. It is expected that humans recognize the objects as a part of common sense. Therefore 3D models function essentially identically to common sense, and 3D models can be used as objects judged by common sense.

Unusual objects which are different from objects judged by common sense are generated by deforming 3D objects. Although there are various ways of deforming 3D objects, this paper uses “merging”. In detail, the objects about which humans feel strangeness (described as “chimera objects”) are generated by merging two objects picked from a 3D model database. For example, when a dog merges with a chair, a chimera object is generated as seen in Fig. 7.

Question images of Chimera CAPTCHA consist of a chimera object and some ordinary 3D objects. Our CAPTCHA requests users to click the chimera object in the question image. Humans can click it easily, because chimera objects, whose appearance is different from ones judged by common sense, cause a feeling of strangeness. An

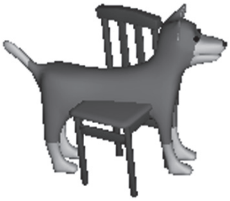


Fig. 7. A chimera object merged from a dog and a chair

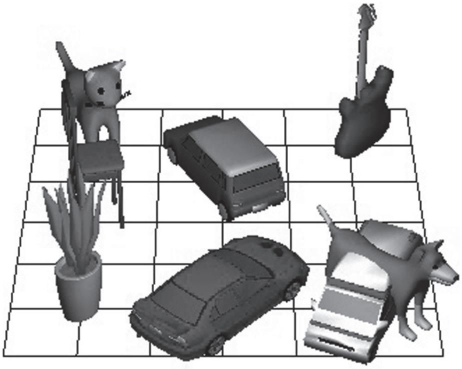


Fig. 8. An example question image of Chimera CAPTCHA

example image of Chimera CAPTCHA in the case of one chimera among seven objects is shown in Fig. 8. In the lower right of Fig. 8, there is a chimera object; a dog and an ambulance are merged.

3.2 Automatic Generation of Chimera Objects and CAPTCHA Questions

As Fig. 9 depicts, chimera objects can be generated by putting two 3D objects on the same point. The detail is as follows:

Step 1. The system picks two 3D objects (called "object A" and "object B" respectively)
Step 2. The system puts object A on an arbitrary point P on the three-dimensional plane α
Step 3. The system puts object B on the point P

When the two objects, A and B, are projected onto the two-dimensional plane, the objects merge with each other and humans can recognize them as a chimera object. Therefore our method makes it possible to generate chimera objects easily. In addition, as Sect. 3.1 describes, it is believed that an indefinite number of 3D models would appear in the future. Therefore our method makes it possible to generate chimera objects countlessly.

Once we can assemble chimera objects, the rest is simple. The question images of Chimera CAPTCHA can be produced just by putting a chimera object and a specified number of ordinary objects on a three-dimensional plane. This is how Chimera CAPTCHA has solved the problem of difficulty in generating questions automatically. This is a significant advantage over SS-CAPTCHA and Four-panel cartoon CAPTCHA.

3.3 Automation Procedure

It is assumed that the Chimera CAPTCHA system has a 3D model database, in which enough numbers of 3D models are archived. Authentication procedure of Chimera

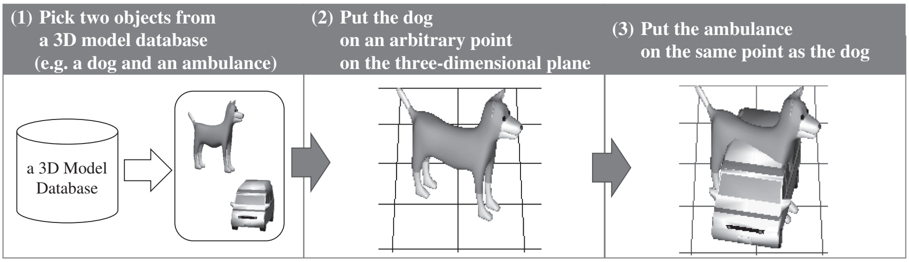

(1) Pick two objects from a 3D model database (e.g. a dog and an ambulance)
(2) Put the dog on an arbitrary point on the three-dimensional plane
(3) Put the ambulance on the same point as the dog
a 3D Model Database

Fig. 9. How to generate a chimera object automatically

CAPTCHA is as follows, where N is the number of 3D objects in an image of Chimera CAPTCHA, or security parameter.

Step 1. The system picks up N objects at random from the 3D model database
Step 2. The system transforms each 3D object picked in Step 1 individually and arbitrarily by Affine Transformation
Step 3. The system puts each object transformed in Step 2 on the three-dimensional plane α
Step 4. The system picks up one more 3D object at random from the 3D model database
Step 5. The system transforms the object picked in Step 4 arbitrarily by Affine Transformation
Step 6. The system selects an object at random out of the N objects put in Step 3
Step 7. The system puts the object transformed in Step 5 on the same position as the object selected in Step 6
Step 8. The system projects the one chimera object and the N-1 ordinary objects, placed on the three-dimensional plane α, onto a two-dimensional plane for the question image
Step 9. The system shows a user (Web page visitor) the question image generated in Step 8
Step 10. The user clicks a part of the question image that feels strange, or the two objects merged with each other
Step 11. If the clicked position on the question image is correct, the user is identified as a human, and if the position is incorrect, the user is identified as malware

In Chimera CAPTCHA, it is difficult for malware, which don’t have common sense, to identify the chimera object in the question image. On the other hand, our system knows the position of the chimera object in Step 7. Because this knowledge forms a trapdoor, our system (Web server) can automatically generate the challenges that malware cannot answer, and then the system can determine whether the user (Web page visitor) clicked the correct object or not.

4 A Basic Experiment

We conducted basic experiments to evaluate the authentication rate of the proposed method. After the experiment, we conducted a survey on the subjects for usability.

4.1 Experimental Method

The subjects included seven volunteers, subjects S1–S7, who are all college students majoring in computer security at Shizuoka University. The number of 3D objects in an image (security parameter N) was 24, where the image contained 23 ordinary 3D objects and a chimera object. First each subject could solve as many tutorial challenges as they want. After that, each subject had to solve three challenges. We only evaluated the three trials. The 3D objects used in the tutorials were different from the ones used in the three challenges. Due to an insufficient number of 3D models we had collected, there might have been some objects used more than once in three question images. We instructed the subjects to click on a "merged object (chimera object)" in each question image. For each challenge, we recorded success or failure, the response time and the click position. After completing all of the CAPTCHA challenges, we had the subjects respond to the following questionnaire. Questions 1, 3, 5 were answered on a 5-point scale.

Question 1. Is it easy solving the CAPTCHA? (Easy): Yes (5) – No (1)

Question 2. If you choose 1 or 2 in Question 1, please write why you think that it is not easy

Question 3. Is it user-friendly? (User-friendly): Yes (5) – No (1)

Question 4. If you choose 1 or 2 in Question 3, please write why you think that it is not user-friendly

Question 5. Is it pleasant? (Pleasant): Yes (5) – No (1)

Question 6. If you choose 4 or 5 in Question 5, please write why you think that it is pleasant

Question 7. How many challenges would you be able to consecutively solve? Also, please write why you think that

Question 8. Which would you choose: text recognition-based CAPTCHA or Chimera CAPTCHA in a real Web service? Also, please write your reason

4.2 Correct Response

The experimental results are shown in Table 1, which summarizes the correct response rate and the average response time for each subject. From Table 1, the correct response rate of Chimera CAPTCHA is 90.5 % on average (a total of 21 times, 19 successes, 2 failures).

From Table 1, the average response time per challenge is 5.7 s; the shortest time is 2.7 s, and the maximum time is 12.2 s. The expected response time for text recognition-based CAPTCHAs is around 10 s at the most [14]. From these data, it can be concluded that the proposed CAPTCHA can be solved in a shorter time compared to text recognition-based CAPTCHA.

We analyzed why the two subjects failed. First is a mistake because of overlooking a chimera object. We showed the subject S3 the image he had failed, to which he replied, "I overlooked the answer (chimera) object". Second, in the image the subject S7 had failed, the two objects consisted of the chimera object were almost the same

Table 1. The experimental results for each subject

	Subject							
	S1	S2	S3	S4	S5	S6	S7	Average
Correct response rate	3/3	3/3	2/3	3/3	3/3	3/3	2/3	90.5% (19/21)
Average response time [sec]	3.4	4.6	12.2	5.2	4.5	2.7	7.5	5.7

color and almost the same size. For this reason, it seems to be difficult to find the chimera object in the question image. There is room for improvement in generating chimera objects.

4.3 Usability

The results of the survey are shown in Table 2.

In Question 1, most subjects answered 4 (5 if easy), and the average value was 4.0. The subjects who answered difficult (1 or 2) were asked to write the reason in Question 2. The subject S2 selected 2 and stated “There are some difficult images to find a chimera object”. As discussed in Sect. 4.2, this evaluation of S2 might be caused by presenting a difficult question image to solve.

In Question 3, most subjects answered 5 (5 if user-friendly), and the average value was 4.0. In Question 4, no subject answered user-hostile (1 or 2).

In Question 5, most subjects answered 4 (5 if pleasant), and the average value was 3.6. The subjects who answered pleasant (4 or 5) wrote a reason in Question 6. The subject S1 stated, “It is interesting because it likes Quiz” and S3 stated, “It is interesting because it likes Game”. The main proposal of Chimera CAPTCHA is to improve attack tolerance and generate questions automatically and easily. However, the result in Question 5 and 6 suggested that Chimera CAPTCHA improved the entertainment value as well. In general, solving CAPTCHA is a chore task for normal users. By improving the entertainment value with our proposed CAPTCHA, users could enjoy solving CAPTCHA. Thus, Chimera CAPTCHA has the potential of contributing to improve usability.

In Question 7, all subjects answered “three challenges”. The reasons are as follows: “Challenging three times is comfortable for me. If the number of challenges is greater than three, I may feel chore”; “I solved three challenges, but I required a little effort”. The number of challenges (three challenges) may be a good number for users. However, the evaluation may also depend strongly on the experimental condition.

In Question 8, six subjects chose Chimera CAPTCHA, while only one subject chose the text recognition based CAPTCHA. The main reasons for choosing Chimera can be broadly divided into two. First is reducing the response time. For example, the

Table 2. Result of survey

	Subject							Average
	S1	S2	S3	S4	S5	S6	S7	
Q1	5	4	4	4	5	4	2	4.0
Q3	3	4	5	4	5	4	3	4.0
Q5	4	4	4	2	2	5	4	3.6
Q7	3	3	3	3	3	3	3	3.0
Q8	C	C	C	C	T	C	C	-

Q1. Easy: Yes (5) – No (1)
Q3. User-friendly: Yes (5) – No (1)
Q5. Pleasant: Yes (5) – No (1)
Q7. How many questions
Q8. Which would you choose?
(T: Text recognition-based CAPTCHA
C: proposed CAPTCHA)

subject S1 stated “This CAPTCHA takes only a short time for response” and S4 stated “This is a fuss-free CAPTCHA”. Second is user-friendliness, for example, the subject S2 stated “It is easy because it takes only one click” and S4 stated “It can be done only with the mouse”. However, as many Web services uses the text recognition-based CAPTCHA, users are accustomed to solve the text recognition-based CAPTCHA. The subject S5 who chose the text recognition-based CAPTCHA stated “I feel Chimera CAPTCHA is as easy as the text recognition-based CAPTCHA, so I chose the latter”.

5 Discussion

5.1 Attack Tolerance

Finding Chimera Objects. A typical attack against the Chimera CAPTCHA it can be considered is that malware extract all objects from a question image and find a chimera object. Namely, malware may try to find “a merged object” in an image. However, for the following two reasons, our CAPTCHA has the tolerance against this attack.

First, there will be occluded objects in a question image. For example, in the upper left of Fig. 8, a cat is occluded by a chair. Humans have spatial reasoning ability. Using the ability, humans could distinguish between “merged objects” and “occluded objects”. In contrast, computers do not so far have an adequate level of this ability, so malware may not distinguish between them.

Second, there are many merged objects in the real world. For example, the object shown in Fig. 10 consists of grass and a pot. These objects are “usual” merged objects, in contrast with “unusual” merged objects (chimera objects) that our system generates. Even if the image analysis technologies are advanced and the computers can recognize whether an object is a merged object or not, malware could still not recognize whether the merged object is a usual merged object or unusual one.

Brute Force Attack. In Chimera CAPTCHA, there remains the problem of low tolerance against a brute-force attack. If malware extract all objects from a question image, they can solve the CAPTCHA with a probability of one in N, where N is the sum number of 3D objects used in an image (one chimera object and N-1 ordinary objects).

Increasing the security parameter N or number of questions is the simple way to improve tolerance against brute-force attack. At the same time, however, increasing users’ mental load may reduce usability. In addition, if our CAPTCHA system requests users to correct an M number of questions, it is concerned that the answer response rate will be the M power of the rate per question. An ingenuity for reducing response time and increasing correct response rate is needed, such as being easier to find chimera objects.

5.2 Automatic Generation

As shown in Sect. 2, the existing CAPTCHAs using the ability of feeling strangeness have the difficultly in generating questions automatically. On the other hand, as shown

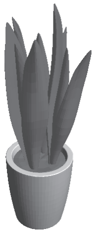


Fig. 10. A usual merged object; grass and a pot are merged

in Sect. 3.2, the Chimera CAPTCHA system makes it possible to generate questions automatically and easily by using 3DCG technologies. In detail, archiving enough 3D models in a 3D database and changing used objects or the parameters such as the scale of object, the system can generate question images countlessly.

6 Conclusion

In this paper, we proposed Chimera CAPTCHA, which is an image-based CAPTCHA focusing on the advanced human-cognitive-processing ability of feeling that something is wrong. The main feature of the CAPTCHA is that computers (Web server) can automatically and easily generate questions that are difficult to solve by computers (malware). While developing a prototype system of the CAPTCHA, we have carried out basic experiments. Seven human subjects solved the challenges of our CAPTCHA in the experiment. The results showed that the correct response rate is 90.5 % and the average response time per one challenge is 5.7 s. Our survey of the results of usability is satisfactory.

It is expected that solving Chimera CAPTCHA is a very difficult task for computers (malware). However, the attack techniques of malware vary and Chimera CAPTCHA's resistance to decipherment is not proven theoretically. We will conduct studies to determine whether our CAPTCHA is truly resistant to malware attacks.

References

1. The Official CAPTCHA Site. http://www.captcha.net
2. Elson, J., Douceur, J., Howela, J., Saul, J.: Asirra: a CAPTCHA that exploit interest aligned manual image categorization. In: 2007 ACM CSS, pp. 366–374 (2007)
3. YUNiTi.com. http://www.yuniti.com/
4. PWNtcha-Captcha Decoder. http://caca.zoy.org/wiki/PWNtcha
5. Yan, J., Ahmad, A.S.E.: Breaking visual CAPTCHAs with naïve pattern recognition algorithms. In: 2007 Computer Security Applications Conference, pp. 279–291 (2007)
6. TechnoBabble Pro: How they'll break the 3D CAPTCHA. http://technobabblepro.blogspot.jp/2009/04/how-theyll-break-3d-captcha.html

7. Chellapilla, K., Larson, K., Simard, P., Czerwinski, M.: Computers beat humans at single character recognition in reading-based human interaction. In: The 2nd Conference on Email and Anti-Spam (2005)
8. Yamamoto, T., Tygar, J.D., Nishigaki, M.: CAPTCHA using strangeness in machine translation. In: The 24th International Conference on Advanced Information Networking and Applications, pp. 430–437 (2010)
9. Yamamoto, T., Suzuki, T., Nishigaki, M.: A proposal of four-panel cartoon CAPTCHA. In: The 25th International Conference on Advanced Information Networking and Applications, pp. 159–166 (2011)
10. D'Souza, D., Polina, P.C., Yampolskiy, R.V.: Avatar CAPTCHA: telling computers and humans apart via face classification. In: 2012 IEEE International Conference on Electro/Information Technology, pp. 1–6 (2012)
11. Mori, M., MacDorman, K.F., Kageki, N.: The uncanny valley. IEEE Robot. Autom. Mag. **19**(2), 98–100 (2012)
12. Korayem, M., Moharmed, A.A., Crandall, D., Yampolskiy, R.V.: Solving avatar CAPTCHAs automatically. In: 2012 International Conference on Advance Machine Learning Technologies and Applications, pp. 102–110 (2012)
13. Ueda, M.: Shin Kobo-chan 8, Houbunsha (2006). (in Japanese)
14. Kani, J., Suzuki, T., Uehara, A., Yamamoto, T., Nishigaki, M.: Four-panel cartoon CAPTCHA. IPSJ J. **54**(9), 2232–2243 (2013). (in Japanese)