Character Image Semantic-Based CAPTCHA

Aziz Barbar and Anis Ismail

***Abstract*—CAPTCHA is almost a standard security technology, and has found widespread application in commercial websites. Image recognition CAPTCHAs faces many potential problems which have not been fully studied. It is difficult for a small site to acquire a large dictionary of images which an attacker does not have access to and without a means of automatically acquiring new labeled images, an image based challenge does not usually meet the definition of a CAPTCHA. They are either unusable or prone to attacks. In this paper, we propose the Character Image Semantic (CIS) based CAPTCHA that will combine both Image and Text into one CAPTCHA which will strengthen it and does not make it human solvable. Benchmarks will be discussed after conducting both Usability and Security tests with a clear set of results that show the accuracy of the proposed solution compared to other existing ones.**

***Index Terms*—CAPTCHA, security, image recognition, semantics.**

I. Introduction

With the development of the computer applications in different fields, internet has made a tremendous progress and become a special need in human life. It has applications in a wide range of daily affairs including trade, education, daily purchases and dialogues take place with the use of Internet. One of the common actions in the Internet web sites, especially commercial and administrative ones, is to fill out registration forms for certain purposes. Unfortunately, there are some programs which automatically fill out these forms with incorrect information to abuse the site, or automated programs which are usually written to generate spam. The notion of a machine imitating human intelligence was first addressed as early as 1950 by English mathematician and logician Alan Turing [1]. Acknowledged as the father of modern computing, Turing recognized that computers might eventually be able to imitate human thought in very convincing ways. Therefore, he suggested what is now known as the Turing test, where a human converses with a computer without seeing it. If the human is convinced by the computer's answers that it is human, then the machine passes the test and is deemed to have some level of human-like intelligence.

HIPs [2] are a slight modification of a reverse Turing test, where the challenge is administered by a machine and taken by a human. The burden is on the human participant to convince the machine that he is human. Furthermore, the challenge should not be solvable by any machine. Notice the paradox that this creates: the machine can automatically create, administer, and grade a test that it itself cannot pass. Tests developed to differentiate these programs from real humans took the form of what would come to be known as CAPTCHAs. CAPTCHAs generate and grade tests that most humans can pass but current computer programs can't. Such tests, often called CAPTCHA challenges are based on hard, open artificial intelligence problems. To date, the most commonly used CAPTCHAs are text-based, in which the challenge appears as an image of distorted text that the user must decipher and retype. These schemes typically exploit the difficulty for state-of-the-art computer programs to recognize distorted text.

In particular, we recently found that we could break a widely deployed CAPTCHA, carefully designed and tuned by Microsoft with a success rate of higher than 60 percent, even though its design goal was that automated attacks shouldn't achieve a success rate of higher than 0.01 percent. We expect that CAPTCHA will go through the same process of evolutionary development as cryptography, digital watermarking, and the like, with an iterative process in which successful attacks lead to the development of more robust systems. So, we are going to improve and enhance the CAPTCHA by providing a new strategy, providing some form of a challenge "easy questions" that seems easy to a human but hard to a robot. Then, the user is expected to provide an answer as a proof of his humanness.

II. Related Work

Many CAPTCHA implementations were designed by different companies (Microsoft, Yahoo, AltaVista) in order to offer a more secure online environment. An environment that distinguishes internet communications originating from humans from those originating from software robots. This section is going to present the different types of CAPTCHAs trying to defeat advanced computer programs or bots, discussing the limitations and drawbacks of each. In character labeling based CAPTCHA designs, the computer renders a sequence of letters after distorting them and adding noise. The user is asked to tell what characters they are in order, and will pass the test if the characters typed (new labels) match exactly those known to the server (known labels). Character labeling CAPTCHAs are the most widely used CAPTCHAs. In 2001 Allison Coates, Henry S. Baird and Richard Fateman of UC Berkeley developed Pessimal Print: that is low–quality of printed text images used certain rate of distortion [3]. The popularity of such schemes is due to the fact that they have many advantages [4].

Research on CAPTCHA mechanisms has received significant attention with the aim to improve their usability and at the same time prevent adversarial attacks by malicious

Manuscript received January 4, 2015; revised March 24, 2015.

Aziz Barbar is with American University of Science & Technology, Beirut, Lebanon (e-mail: abarbar@aust.edu.lb).

Anis Ismail is with University Institute of Technology, Lebanese University, Sidon, Lebanon (e-mail: anismaiil@ul.edu.lb).

software. Researchers promote various CAPTCHA designs based on text and speech-recognition challenges, and image puzzle problem [5]. Nevertheless, a variety of studies have been reported that underpin the necessity for improving the usability of CAPTCHA mechanisms. Result from a recent study, which investigated users' perception towards CAPTCHA challenges; claim that current implementations do not provide an acceptable trade off solution with regards to CAPTCHA usability [6]. Another large-scale study, which evaluated CAPTCHA on the Internet's biggest websites, revealed that CAPTCHAs are difficult for humans to solve [7].

The algorithms and data used to automatically generate these CAPTCHA challenges are publicly available. But with the advancement of OCR and sophisticated image processing algorithms and tools, these text-based CAPTCHAs can no longer provide the secure access to authenticate users from malicious computer programs. For instance, researchers have developed an attack against Microsoft's Hotmail CAPTCHA that yields a 60% success rate [8]. Also more complex image distortion to make it difficult for programs to crack, makes this text based method increasingly hard for human users to recognize the text, causing usability issues [9]. Thus the need for new form of CAPTCHA that is automated, open, usable, and secure is of urgent need. There are three categories of implementations of CAPTCHA schemes: text, sound, and image based schemes. Audio CAPTCHA usually pronounces letters or digits in randomly spaced intervals. Background noises may be added to make the tests more robust against bots. These systems are dependent on some sort of audio hardware to produce the sound clearly, and these sounds are sometimes difficult to perceive for locality reasons. Also persons with hearing difficulties cannot use this scheme. Furthermore, the basic principle to attack this CAPTCHA remains similar as text-based ones, which is to extract the feature and recognize the letters. Hence, the audio based CAPTCHA scheme does not provide any more user-friendliness or robustness against bots than text based CAPTCHA [10]. In [11], a new game theorem based CAPTCHA system is proposed.

Image based CAPTCHAs inquire users to perform some forms of image recognition tasks. These systems are developed to overcome the shortcomings of previously discussed schemes of CAPTCHAs. There are some schemes that use human ability to perceive and semantically analyze images to perform a task [12]. There are also some methods that ask users to adjust the orientation of 3D images or to identify semantic meaning from it. Microsoft's Asirra [13] was designed to use the existing database of petfinder.com and prompts users to identify images of cats out of other pets.

III. Proposed Solution

Our new CAPTCHA approach named "CIS", the Character Image Semantic based CAPTCHA will combine both types, in addition to a question based on the image, that is going to increase the strength of its security and will make it harder to be broken. We are aiming to a harder CAPTCHA solution with no harder problems for people.

A. The Image Approach

Image recognition CAPTCHAs face many potential problems which have not been fully studied. It is difficult for a small site to acquire a large dictionary of images which an attacker does not have access to and without a means of automatically acquiring new labeled images, an image based challenge does not usually meet the definition of a CAPTCHA. Thus, these set of images may be vulnerable, where the attacker can reach them. Also, if the storage concept is eliminated, rendering and drawing the images will be costly from performance point of view.

So, the new approach must provide a compromise. The "CIS" can provide any set of identical geometrical shapes. In this implementation we present to the user a set of circles, we will call them "balls" that will replace the concept of keeping a set of images in a dictionary, because balls are easy to draw at runtime and they are simple shapes (costless according to time). The key in CIS image implementation is the color of balls. Balls will have different colors. And the questions asked for the user will be based mainly on the color (the semantics part is explained in the next section).

There are many points that strengthen this approach. Balls are randomly distributed among the whole area of the CAPTCHA background, the location of balls will differ on each generation. This will make the segmentation "Splitting the image into regions which each contain a single object", because each pixel is going to be checked. Thus, Balls may interleave, or intersect with each others. While the process is easy for humans to identify such balls, the automated software will face difficulties (again in the segmentation phase). Another modification which can be added to the CIS which raises its security is changing the size of balls or the area occupied on different generations. Also, we can generate any set of shapes (rectangles, squares, and triangle) or objects. Image processing will not be effective, because image segmentation object detection and extraction will be impossible, thus, the cracker cannot know what is the concerned shape that it's both must recognize.

B. The Semantics Approach

The CIS interacts with the user by providing a specific question. The user must provide the correct answer based on what he sees. He passes if he entered the expected answer. The question consists of different words that contain a meaning, so we can benefit from playing with the words in order to create a set of questions to add a semantic approach to the CAPTCHA.

The Semantic approach will add different meaningful questions while the answer expected from the user is a number. Each time a different question based on the count of balls of specific color can be provided to the user (Even on one specific image). Now, the cracker has to consider a new obstacle and to change his mentality of thinking. His automated software will suffer and he must analyze and understand the question to pass. Some of the questions that could be provided by the CIS are "How many red balls you can count in the picture?", "How many distinct colors in the picture?", "What is the number of balls in the picture?", "Regardless of the color, how many balls you can count?".

Now, we can assume that the cracker is going to study the

syntax and the morphology (NLP techniques) of CIS questions. His software will depend on extraction of specific words, and then it analyzes the request and works on the target image to predict the answer. CIS can be extended and provide a new barrier that overcomes such deficiency (if exists) by providing another form of Semantic questions. Such questions need to be solved and may contain simple arithmetic operations are "What is the sum of yellow and green balls?", "What is the result of multiplying the red balls by 2?", "If we subtract blue balls from red balls, what is the result?", "Is the number of yellow balls, odd or even? If odd enter 1 otherwise 2." The semantics can be the key to reach a secure CAPTCHA which is able to stand against attacks from automated software. A good point is that questions are obvious to the user, but the bot will find difficulties to analyze the words or reveal the meaning.

C. The Characters Part

Like any text based character, CIS renders the characters in front of the balls, the characters and the balls interleave with each other. Also, all set of character are used to make the process of recognizing them more difficult.

D. User Interaction with CIS

The user should pass the CIS validation in order to proceed for the next task. The interaction can be summarized in Fig. 1. First, he should check the question and analyze it. Then, check the image and enter the correct answer (a number). Finally, the characters in CIS should be entered in the order shown.

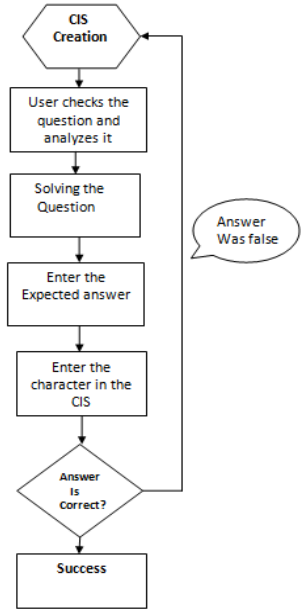

Fig. 1. User interaction with CIS.

E. CIS Workflow Overview

In Fig. 2, we show the flow of the process of preparing and creating the CIS CAPTCHA. The CIS depends on a random generator that picks a different question and set of characters each time. Noting that a question is a container of different attributes as shown.

Finally, the drawing phase comes where the balls are drawn behind the characters to provide more complexity and security. Below, we can see some screenshots taken for the CIS (Fig. 1). Notice how the balls interleave with each others, they distribute on different locations each time.

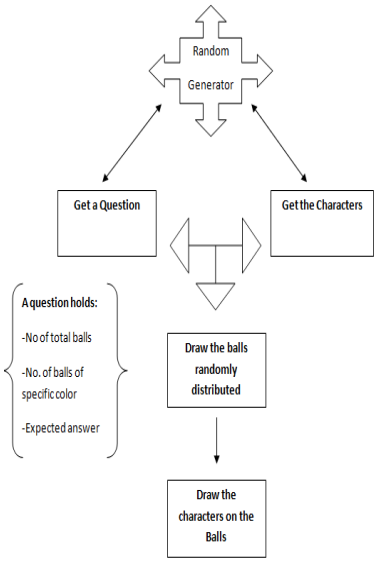

Fig. 2. CIS workflow.

The user is expected to answer the statement, "Enter number of blue balls". There exists only one blue ball in Fig. 3, also after he enters the correct characters, he will pass the CIS validation.

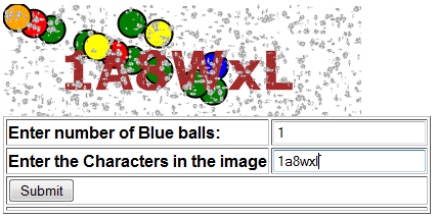

Fig. 3. Screenshot 1 from CIS.

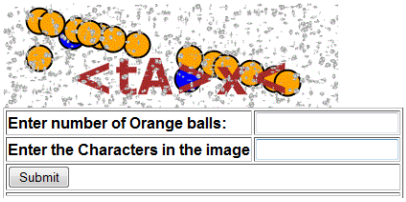

Fig. 4. Screenshot 2 from CIS.

We can see the complexity of knowing the number of balls (see Fig. 4) where the orange balls intersect with each others in continuous way.

Since designing a CAPTCHA is a synthesis task, we have benefited from already existing solutions in order to provide a new way of thinking in CAPTCHA world. CIS is a combination that will defend strongly against bots. The Semantic part allows the CIS to provide dynamic challenges to the user that cannot be easily processed and analyzed by intelligent programs.

IV. Testing and Verification

In order to study and test the effectiveness of CIS CAPTCHA, we need to make a comparative study with already existing implementations in terms of two metrics: usability and security. We quantify the usability of a CAPTCHA using the success rates of humans in an

experimental setting and the security of a CAPTCHA as the probability that a specific, automated attack passes the challenge.

A. Usability Tests

CIS first prototype system was implemented in order to carry out real user experiments based on it. The purpose is to check usability (score hits and time taken to pass the test) and to explore the ways to improve it.

Forty participants from different environments were chosen randomly to accomplish the CIS test. Each person has 10 trials, the total number of trials was 400. The participants were divided into groups according to their educational levels. Group 1 represents individuals in the primary school years. Group 2, represents individuals in the high school, group 3 represents individuals who have earned a B.S., and group 4 represents the individuals who have earned a master degree. In Table I, the results of the trials for the four groups are presented. The second column represents average time to solve the CIS for each group, and the third column represents the average hit scores or success ratio (success in passing the CIS) for each group.

TABLE I: USABILITY TEST RESULTS

Group Number	Average time spent to solve CIS	Success Ratio
1	18	88%
2	15	90%
3	11	93%
4	13	92%

The results were promising, the average response time for successfully passing a test was 14.25 seconds and the success ratio was 91%.

According to the Image Based CAPTCHAs, the anomaly CAPTCHA achieved the best scores with success ratio of 90%. However the expected time to take this CAPTCHA was 51 seconds.

B. Security Tests

Most of the already existing CAPTCHAs were broken at high success rates. Table II shows various implementations with the attack success rates.

TABLE II: BROKEN CAPTCHAS

CAPTCHA Name	Ez-Gimpy	Gimp-r	Gimpy	Microsoft	Others
Success rate	92%	78%	38%	60%	49% - 100%

CIS is considered secure comparing it to others. It is a combination that makes it less vulnerable. We need to put it in the real environment "The Internet" in order to figure out its strength. However, to break the CIS we will approximate the results theoretically; the attacker must break the 3 parts. The probability that an attack passes the each part is P (A).

According to the character part, we can use one of the existing hard to break labeled CAPTCHA with P (A) = 0.5. The image part with P (A) = 0.6 and the semantic part with P (A) = 0.5. Thus, the success rate of attacks on CIS is approximated as 15% in the worst scenarios.

Striking a balance between security and usability is a difficult task. It is most often the case with CAPTCHAs that the two metrics vary inversely, if you wish to have higher security, the usability will suffer and if you wish to have higher usability, the security will suffer. However, the CIS is secure and usable as proved in the tests.

V. CONCLUSIONS AND FUTURE WORK

We have proposed the first CAPTCHA that uses Image, Characters and some of form of semantic understanding to distinguish between legitimate humans and machines and showed it to be a viable alternative to existing CAPTCHAs. A new method for differentiating between humans and machines was developed, the "CIS". CIS adds a new factor (the semantics), this is to reduce the vulnerability of the system against one or more attacks. CIS may be a more enjoyable alternative to text and image based CAPTCHAs. We have provided a set of techniques that would allow for the system to be secure and less vulnerable to bot attacks. It is a well synthesized CAPTCHA, where the attacker should pass three obstacles in order to bypass it. The only automation that should be able to pass the CIS is the one generating it. Also, the usability and security of our CIS CAPTCHA is good compared to existing text-based and image-based CAPTCHAs. In previous implementations hard un-broken CAPTCHAs were not human solvable, that decreases their usage because they affect human performance due to the time required and low hit scores.

Also, any machine can generate and grade the CIS easily. The Image part is simple to generate and render without the need for a storage or even querying Google searching for a set of images. It has its own set of geometrical shapes that constitutes the image part. CIS satisfies the basic properties of a CAPCTCHA and provides a secure mechanism for securing online environments.

We have provided a new concept to the CAPTCHA world. Further work and enhancements can be achieved in order to enrich the CIS and make it more defendable. These enhancements can be summarized in the following points, drawing different geometrical shapes (squares, circles, triangles), using one of the existing strong labeled based CAPTCHAs, rendering Characters with the colors similar to balls, that will make a new barrier for image processing functions, and Enhancing the randomization and the distribution of balls. According to the semantic layer, we can make it more complex by creating more challenges and questions based on the shapes generated, such as adding questions that involve more analytic and solving skills e.g., asking questions with arithmetic operations or questions that involve different shapes (count of triangles and circles).

REFERENCES

[1] The Alan Turing Internet Scrapbook. (1950). The Turing Test. [Online]. Available: http://www.turing.org.uk/turing/scrapbook/test.html

[2] K. A. Kluever. (2008). Securely extending tag sets to improve usability in a video-based human interactive proof. Department of Computer Science Rochester Institute of Technology. Rochester. [Online]. Available: http://www.kloover.com/thesis/proposal.pdf

[3] H. S. Baird, A. L. Coates, and R. J. Fateman, "Pessimal print: A reverse turing test," *Int. Journal of Document Analysis and Recognition*, Seattle, WA, vol. 5, no. 2-3, pp. 158-163, April 2003.

[4] K. Chellapilla, K. Larson, P. Y. Simard, and M. Czerwinski, "Building segmentation based human-friendly human interaction proofs (hips)," in *Human Interactive Proofs*, H. S. Baird and D. P. Lopresti (Eds.), vol. 3517, Berlin Heidelberg: Springer, 2005, pp. 1-26.
[5] M. Belk, P. Germanakos, C. Fidas, G. Spanoudis, and G. Samaras, "Studying the effect of human cognition on text and image recognition CAPTCHA mechanisms," *Human Aspects of Information Security, Privacy, and Trust*, Berlin Heidelberg: Springer, 2013, pp. 71-79.
[6] C. Fidas, A. Voyiatzis, and N. Avouris, "On the necessity of user-friendly CAPTCHA," in *Proc. 29th ACM Conference on Human Factors in Computing Systems*, ACM Press, New York, 2011, pp. 2623-2626.
[7] E. Bursztein, S. Bethard, C. Fabry, J. C. Mitchell, and D. Jurafsky, "How good are humans at solving CAPTCHAs? A large scale evaluation," in *Proc. IEEE International Symposium on Security and Privacy*, IEEE Press, Washington, 2010, pp. 399-413.
[8] J. Yan and A. Ahmed, "A low-cost attack on a Microsoft CAPTCHA," in *Proc. CCS*, ACM Press, 2008, pp. 543-554.
[9] L. Yue. CAPTCHA, 89 - The 10th Zhejiang University Programming Contest – B. [Online]. Available: http://acm.zju.edu.cn/onlinejudge/showContestProblem.do?problemId=3714
[10] E. Bursztein, R. Bauxis, H. Paskov, D. Perito, C. Fabry, and J. Mitchell, "The failure of noise-based non-continuous audio captchas," presented at 2011 IEEE Symposium of Security and Privacy, Oakland, 2011.
[11] J. Kani and M. Nishigaki, *Gamified CAPTCHA. In Human Aspects of Information Security, Privacy, and Trust*, Berlin Heidelberg: Springer, 2013, pp. 39-48.
[12] S. Vikram, Y. Fan, and G. Gu, "SEMAGE: A new image-based two-factor CAPTCHA," ACSAC, Orlando, Florida, USA: ACM, pp. 237-246, 2011.
[13] J. Elson, J. Doucerur, J. Howell, and J. Saul, "Asirra: A CAPTCHA that exploits interest-aligned manual image categorization" in *Proc. the 14th ACM Conference on Computer and Communication Security*, New York, NY, USA: ACM, pp. 366-374, 2007.

Aziz M. Barbar is the dean of the Faculty of Arts and Sciences Science at the American University of Science & Technology (AUST), Lebanon. He has a Ph.D. in computer science from the University of Nice-Sophia Antipolis, France. His research interests include database reverse engineering, data mining and natural language processing. Dr. Barbar is currently the vice-president of the Lebanese Information Technology Association (LITA), and the vice-chair of the IEEE section, Lebanon.

Anis A. Ismail was born in Lebanon. He is an assistant professor at the Lebanese University, University Institute of Technology, Lebanon. He has a BS degree in telecommunication and networking engineering from the Lebanese University (LU), an MS in computer science and an MS in CCE from the American University of Science and Technology (AUST) in Lebanon, and a Ph.D. in computer science from the University of AIX-Marseille, France. His main research interests include data mining in P2P systems, arabic language processing, and multimedia information.