Abstract—This paper presents a new semantic image CAPTCHA, called Multiple SEIMCHA. Multiple SEIMCHA system warps images by using geometric transformations and a 2d view of warped image is shown to user. Users should click on the upright orientation of warped image as a semantic. Multiple SEIMCHA evaluates user based on the idea of almost right response instead of completely right response. This idea uses hardness rate concept which is defined based on user response rate to images. The proposed system has a great response time and success rate as usability metrics and it is secure to bots.

Keywords-component; Multiple SEIMCHA; Almost Right Response based on Hardness Rate; Semantic Image based CAPTCHA; Upright Orientation; Geometric transformation.

I. INTRODUCTION

We, as user, need to prove ourselves as human being in the web because automated scripts or bots try to play the role of human for using free web services over the past few years. For example a script sends many spam emails for advertising. Completely Automated Public Turing Test to Tell Computer and Humans Apart (CAPTCHA) or Human Interaction Proof (HIP) offers a way to make distinction between a human and an artificial agent. In fact, CAPTCHA is a secondary authentication mechanism. The first CAPTCHA was presented in 2000 in Carnegie Melon University [1]. Today several types of CAPTCHAs included; text-based, image-based, voice-based and video-based are used. Researches try to design more secure and more usable CAPTCHAs every day.

Various criteria have been proposed in the literature for evaluating CAPTCHAs. We will consider the following four properties (originally reported in [2]):

- Automated: Tests should be easy to automatically generate and grade by a computer.
- Open: The underlying database(s) and algorithm(s) used to generate and grade the tests should be public.

This property is in accordance with Kerckhoffs’s Principle, which states that a system should remain secure even if everything about the system is public knowledge.

- Usable: Tests should be easily solved by humans in a reasonable amount of time. Furthermore, the effect of a user’s language, physical location, education, and/or perceptual abilities should be minimized.

- Secure: Tests should be difficult for machines to solve algorithmically.

Because of the attractiveness of image based CAPTCHAs and the difficulty of text-based CAPTCHAs for some people [5], image based CAPTCHAs are developed more and more as a good alternative for text based ones. There are several works on breaking text-based CAPTCHAs in various ways [3]. Identifying image is more difficult than identifying text for machines. There are different kinds of image based CAPTCHAs. To date, a lot of CAPTCHA designs are labeling based. In a tagging based or labeling based CAPTCHA, the label provided by the user side should matches the label already known to the CAPTCHA system. Labeling based CAPTCHA designs have some common weaknesses like unreliable labels for images on the web, small and fixed label set, misspelling in user side and etc. Also automated meaningful tagging by machine is very hard [4].

One of the big challenges in designing image based CAPTCHAs is the security of image database. Based on the properties of CAPTCHA the database of an ideal CAPTCHA should be open while the CAPTCHA should remain secure. Showing original images from a small database threatens security. Some common solution to avoid this are: (1) using a large database of images in scale of web, (2) updating database regularly and (3) using a transformed version of original image with many transformation parameters. These solutions are used in several works like [5, 6, 7, 8].
Another approach to design Image based CAPTCHAs is using semantic content in images. In these CAPTCHAs users should identify semantic cues instead of typing, selecting or mapping tags. One easy semantic for human is upright orientation. Currently, automatic detection of such concepts is possible only for a small subset of images [9, 10]. As reported in [11], 68.75% of users preferred rotating images as CAPTCHA, and 31.25% of users preferred deciphering text. And it confirms the attractiveness of this kind of image CAPTCHA for human. By focusing on orientation, we don’t face problems in tagging CAPTCHAs and can improve usability and security as two main properties for CAPTCHA.

This work presents a new non-tagging semantic image based CAPTCHA called Multiple SEIMCHA which is based on upright orientation concept, geometric transformations, almost right response strategy and hardness rate concept. Section 2 introduces related works, Section 3 describes Multiple SEIMCHA, contains transformations, heuristics and the right answer strategy applied to improve usability of Multiple SEIMCHA. Section 4 includes empirical study. This section presents usability and security analysis. Section 5 makes a comparison between proposed CAPTCHA and similar works and section 6 includes conclusion and future works.

II. RELATED WORKS

Several image CAPTCHA have been developed until now. A lot of them are tagging based and don’t use semantics. For Example ESP- PIX [16] is the first image based CAPTCHA which provides a set of related images in a GUI and asks user to select a related tag from a proposed list. As there are several differences between tagging CAPTCHAs and semantic CAPTCHAs, we don’t mention them in this paper. Semantic image CAPTCHAs vary from selecting wanted objects in an image by user [8] to identifying upright orientation of an image which is a hard concept for machine [5,7,11]. These CAPTCHA include a wide range of image CAPTCHAs. In this paper we just introduce CAPTCHA based on image orientation.

Gossweiler et al. [11], in the paper “What’s Up CAPTCHA” introduced an orientation based image CAPTCHA for the first time. Their CAPTCHA system presents a series of random rotated images. The user should use a slider to rotate the image continuously until it is in its upright position. Figure 1-a shows a screenshot of this CAPTCHA. In [7] authors introduced a simple orientation based image CAPTCHA which shows a set of images to user and user just needs to identify rotated images and click on them. This CAPTCHA transforms the original images to a line drawing black and white one. Figure 1-b shows a screenshot of this CAPTCHA. As an extension of What’s Up CAPTCHA, Ross et al. [5] introduced a new CAPTCHA based on upright orientation of line drawing rendered from 3D models called Sketcha. They downloaded their models from Google 3D Warehouse and rendered a collection of images from various angles (figure 1-c). The user should click on the image to rotate it until it is upright.

Multiple SEIMCHA also uses the idea of upright orientation as a semantic cue in combination with geometric transformations. In Multiple SEIMCHA we create a set of 3D shapes such as sphere, cone and etc. by geometric transformations and wrap the original image on to one of them. Then a 2D projection from a random viewpoint gives the final image. User should click on the upper area of the final image. The usability and security of these transformations without applying upright orientation is studied in [6]. Also we studied the combination of upright orientation and geometric transformations in a simple system, showing 1 image per each challenge, and called this simple system as SEIMCHA. We found the ideas fit to apply to a real image CAPTCHA.

In Multiple SEIMCHA we show 8 images to user in each challenge instead of 1 image, and study the behavior of users in the interaction with the system. We define a hardness rate concept for images based on response rate and grade users by using an almost right response strategy instead of applying completely right response based on this concept. Showing multiple images instead of one image is for improving security of SEIMCHA. And using almost right answer strategy based on hardness rate concept is for improving usability of Multiple SEIMCHA. Usability and security of Multiple SEIMCHA is more satisfactory than similar works.

III. MULTIPLE SEIMCHA

We present Multiple SEIMCHA as a new semantic image CAPTCHA. We use several approaches to design a more secure and more usable semantic image CAPTCHA. These approaches are: Geometric transformations, upright orientation, hardness rate and almost right answer.

The GUI shows several warped images. Users should click on the logical upright orientation of an image as right answer which is a specific area for server since images are transformed. Note to fleshes in Figure 2, part c, main output images. These fleshes show the area user should click.
We show some warped images to the user from a fixed database of 30 identifiable input images for human. These transformations are implemented in Matlab by using some geometric functions. Applying geometric transformations for images is a new solution for using in CAPTCHAs.

A. Transformations

In Multiple SEIMCHA we apply some transformations like simple rotations, corruptions and geometric transformations to warp an input image. Then convert 3D object to 2D images by capturing from a random viewpoint. This helps to generate many new various images of each original image. The below algorithm describes the approach better:

1. Randomly select an input image
2. Randomly rotate input image
3. Randomly select a 3D geometric transformation and transform image on it
4. Capture a 2D image from a random viewpoint

There are a lot of geometric transformations with several variables. We just uses a subset of them includes 6 fixed 3D objects. Figure 2 shows some example of these objects.

A key image is transforming exactly like the input image. This key image is divided into 3 parts and the corresponding top part of input image is considered as correct answer area which could be clicked by user as right answer. Figure 2-a and b show main and key images. The top part in black in key image is the correct answer area.

We applied 2 heuristics to improve usability of Multiple SEIMCHAS. First heuristic uses a little circle in the middle of top part of a white image (figure 3) and this image is transformed like input and key images. If the final heuristic image has some part of this circle, the main final image is usable.

Second heuristic uses the key image. A program calculates the percentage of black part of final key image as visible correct answer and this should be more than 20% of whole image (of colorful parts, not white margin parts) as a usable image.

B. Almost Right Response Strategy

Asirra [12] uses this strategy in image CAPTCHA for the first time. It applied a Partial Credit Algorithm (PCA) as a mechanism to use almost right answer strategy. This CAPTCHA asks users to identify cats out of a set of 12 images of both cats and dogs. They considered a response as almost right answer if 11 out of the 12 images showed are identified correctly. Asirra awards partial credit to a user who answers a test almost right, then the user moves to an intermediate state. And finally the user moves to the verified state if zero or one image is misclassified in next round. Two “almost right” answers means the user is human.
In Multiple SEIMCHA we present another algorithm to use almost right response strategy based on a hardness rate for every image. This hardness rate is measured by user response rate and saved in database previously. Hardness rate means the difficulty of the image for human to identifying the upright orientation of it. We classified our images into 4 groups based on their hardness rate: Simple, Intermediate, Hard and Very hard. Table 1 shows this classification in detailed.

<table>
<thead>
<tr>
<th>Group name</th>
<th>Response rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple</td>
<td>90% to 100%</td>
</tr>
<tr>
<td>Intermediate</td>
<td>80% to 90%</td>
</tr>
<tr>
<td>Hard</td>
<td>70% to 80%</td>
</tr>
<tr>
<td>Very hard</td>
<td>60% to 70%</td>
</tr>
<tr>
<td>Rejected</td>
<td>Less than 60%</td>
</tr>
</tbody>
</table>

In every round we show 2 simple images, 2 intermediate, 2 hard and 2 very hard images to a user. In this way, challenges will be produced in a normal range of difficulty. For using almost right answer strategy, imagine a human judge scores user response instead of computer in Multiple SEIMCHA. For example, our human judge could see the user clicks on 7 out of 8 images, she might say, “I think the user is a human but he can’t identify a very hard image or he clicked an image wrongly because he is careless. I’ll let him pass.” An almost right answer in Multiple SEIMCHA is defined in this way; correct clicking on a special subset out of 8 images based on their hardness rate. We will present these special states based on the empirical studies in next section.

IV. EMPIRICAL STUDY

Multiple SEIMCHA is implemented in ASP.net. The designed GUI shows 8 images in one page to user and asks her/him to click all images in any order she/he wants. After clicking all images, he/she could submit the test and go to the next round. We performed some studies on users and logged their interaction with the system and evaluate Multiple SEIMCHA. We asked 30 users to take part in Multiple SEIMCHA, 15 male and 15 female, 22 to 30 years old. They study as undergraduate, graduate and Ph.D students in engineering faculty of Ferdowsi university of Mashhad 1. 5 rounds were presented to each user and feedbacks were logged into an Access database containing 4 tables:

- User Information Table: Contains user information.
- Picture Information Table: Contains image name, transformations, resulting of applying H1 and H2.
- Multiple SEIMCHA Image Table: Users feedback of each image which contains image name, passed or failed, priority of clicking and response time.
- Multiple SEIMCHA Round Table: Users feedback of each round which contains round number, total time and order of clicking.

These records analyzed by a program.

---

1. www.um.ac.ir

A. Usability Analysis

In this section we report some important usability analysis of Multiple SEIMCHA. Some other analyses contain hardness rate of input images, some example of unusable images, usability of 3D objects, relationship of success rate and expired time and the effect of the heuristics on these are not reported because of the lack of space in this paper.

1) Response Time and Response Rate

Success rate and response time are two usability metrics for CAPTCHA. The average response time to Multiple SEIMCHA is 32 seconds for each round. Users can optimize their response time after 3 rounds. Table 2 shows the improvement of response time per rounds. If we consider expiration time of 60 seconds, the average response time will improve to 26 seconds.

<table>
<thead>
<tr>
<th>Round Number</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Response Time to a Round</td>
<td>58.7 s</td>
<td>42 s</td>
<td>34.8 s</td>
<td>29.7 s</td>
<td>30.2 s</td>
</tr>
</tbody>
</table>

Also the response time of each group of images based of their hardness rate is reported below in table 3. You can see the average response time of simple images is less than hard ones.

<table>
<thead>
<tr>
<th>Group Name</th>
<th>Simple</th>
<th>Intermediate</th>
<th>Hard</th>
<th>Very Hard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Response Time to 1 Image</td>
<td>4.5 s</td>
<td>4.6 s</td>
<td>4.7 s</td>
<td>5.9 s</td>
</tr>
</tbody>
</table>

Success rate of Multiple SEIMCHA based on the minimum number of right clicked images is reported in figure 4. As you see, we can have a 98% success rate with considering 4 right answers out of 8 images. But this margin will increase random guessing which will be discussed later.

2) User Behaviour Analysis

We analyzed the interaction of user with Multiple SEIMCHA. The results show the users prefer first click on simple images. Figure 5 shows the occurrence of users clicks based on the priority and hardness rate of images. Also users could response to 6 and 7 images better than other states. They also can identify simple images better than hard images. We can use these results to define the almost right response strategy rules:

- 8 correct clicked images is a completely right response.
- 7 or 6 correct clicked images are almost right response.
- 5 or 4 correct clicked images that contain all 4 simple and intermediate images are almost right response.
- Completely right response and almost right response are accepted to the system.
The statistics shows that after applying these rules, success rate is Multiple up to 92% and the probability of random guessing will be decreased into a considerable ratio. We will discuss it later.

**B. Security Analysis**

There are 3 kinds of attacks to a CAPTCHA: random guessing, direct matching and machine learning attacks.

First is random guess attack. The probability of random guessing for one image is discussed in equals to the ratio of response area to all image area. According to statistics on our test images, the random guessing probability of one image after applying heuristics is 0.175. Multiple SEIMCHA system which displays 8 images would achieve a guess success rate of less than 0.88×10^-6. Note that because of using the defined rules for almost right answer strategy, the best state for an attacker for random guessing is the state that 4 simple and intermediate is clicked. The attacker doesn’t know what 4 out of 8 images are simple and intermediate. So the number of 4 selection out of 8 images is C_8^4=10. The probability of right clicking on these 4 images is (0.175)^4=0.9×10^{-3}. And the final probability is equal to 0.1×0.9×10^{-3}=0.9×10^{-4}, a standard acceptable computer success rates for CAPTCHAs [13].

Second attack is direct matching attack. In this attack, an attacker saves all possible final images of SEMCHA in a database. The number of images that a program has to traverse for a direct matching is simply enormous and we can calculate it. The needed operations for direct comparison of two images equal to logarithm_2 (the number of image pixels). The final images in Multiple SEIMCHA are 280×210 pixels. And 20% of all possible images are rejected after applying heuristics. Table 4 shows whole operations for direct matching attack. The operations needed for a direct matching is a factor if 10^{13} which is big enough as a problem even for a supercomputer.

Third attack is machine learning attack. There are two main ways to learn an image; using topology of feature points and shape of an image [14, 15]. Applying geometric transformation weakens the attacks based on these two. Also using upright orientation as a hard concept for machine makes Multiple SEIMCHA more difficult for a bot. We did not analyze it in practice and it is a good point for future works.

**V. COMPARISON**

What’s Up [11], Sketcha [5] and Orientation based image CAPTCHA [7] are tree similar CAPTCHAs ask user to identify upright orientation of images. What’s Up displays input images with random rotations, Sketcha shows line drawing images rendered from 3D models and Orientation based image CAPTCHA shows line drawing black and white images transformed from color images. Multiple SEIMCHA show some color images transformed by geometric functions.

Success rate of these CAPTCHAs in order are 84%, 88% and 92%. And it is 92% for Multiple SEIMCHA too. Response time of Sketcha is 35 seconds and it is not reported for What’s Up. The time of success passing challenges of Orientation based image CAPTCHA is 14 seconds. However it is more reasonable to report a total response time for a CAPTCHA instead of a success response time. It is 26 seconds for Multiple SEIMCHA.

Sketcha and What’s UP ask users to set images in their main upright orientations. User response type of What’s Up is image rotation using slider or mouse movement or up-down control. In Sketcha user should rotate each image by clicking. These tasks need more time and effort than a single click on the image which is provided in Multiple SEIMCHA and Orientation based image CAPTCHA. The probability of random guessing of each CAPTCHA plus other comparisons is summarized in Table 5.
TABLE V. COMPARISON OF MULTIPLE SEIMCHA AND SIMILAR CAPTCHAS

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Response time</td>
<td>Not reported</td>
<td>14s for success passing and 12 images</td>
<td>35s for 10 images</td>
<td>Less than 3.8s for 1 image</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>26s for 8 images</td>
</tr>
<tr>
<td>Human success rate</td>
<td>84% for 3 images</td>
<td>92%</td>
<td>88%</td>
<td>92%</td>
</tr>
<tr>
<td>Random guessing</td>
<td>4.44% for 1 image, 0.009% for 3 images</td>
<td>50% for 1 image, 0.02% for 12 images</td>
<td>25% for 1 image, 0.001% for 8 images</td>
<td>17.5% for 1 image, 0.009% for 8 images</td>
</tr>
<tr>
<td>Response mechanism</td>
<td>Slider/ Moving mouse/ Up-down control</td>
<td>Single click for images Multiple click for a round</td>
<td>Single click for images Multiple click for a round</td>
<td></td>
</tr>
</tbody>
</table>

VI. CONCLUSION AND FUTURE WORK

Multiple SEIMCHA is introduced in this paper as a new semantic image based CAPTCHA. It is a non-tagging CAPTCHA which uses upright orientation and geometric transformations. Multiple SEIMCHA shows 8 warped images in a page and asks users to click on the upright part of all images. Showing 8 images in a page, makes it more secure to all attacks especially random guessing. Applying almost right response strategy which uses rules defined based on hardness rate of images, improves usability of Multiple SEIMCHA while keeping it secure. Multiple SEIMCHA has a simple response mechanism, single clicking, therefore it is faster than similar works. It has an excellent 92% response rate and a good response time, 26 seconds for 8 images. Using 3D objects to warp a 2D image provides a huge search space for final images which caused more security for Multiple SEIMCHA.

As future works, we are working on the idea of exploiting users interaction patterns due to a more usable and secure version of Multiple SEIMCHA. We will use these patterns for designing an intelligent system to a better recognition of humans and bots. This intelligent system will grade users based on their interactions. Better answers gain better credits. These credits will be applied to a dynamic Multiple SEIMCHA to update the database. Users can vote to not-evaluate input usability of Multiple SEIMCHA greatly.

We recommend the followings as other future works:

- Applying more mathematical functions for transforming images and finding the best subset of them for using in a image based CAPTCHA
- Working on a special attack for Multiple SEIMCHA which is designed exactly for this CAPTCHA. Such this system can uses different kind of attacks like machine learning attack and mechanical Turk attack.

REFERENCES