A Study of the Users on the Security Codes

Darko Brodic[1], Alessia Amelio[2], Ivo R. Draganov[3]
[1]Technical Faculty in Bor, University of Belgrade
Vojske Jugoslavije 12, 19210 Bor, Serbia
{brodic@tfbor.bg.ac.rs}

[2]DIMES, University of CalabriaVia P. Bucci Cube 44
87036 Rende (CS), Italy
{aamelio@dimes.unical.it}

[3]Technical University of Sofia, 8 Kl. Ohridski Blvd
Sofia 1000, Bulgaria
{idraganov@tu-sofia.bg}

ABSTRACT: *CAPTCHA perhaps, the most known security issue among the general users. Web users have awareness about the security use of it for which no fundamental education is required. The usability elements in CAPTCHA are addressed in this work. We have used the disc CAPTCHA for an experimental analysis. We did it by contacting some two hundred users from many regions through online. We have collected the data and analysed it with interesting results.*

Keywords: Artificial Intelligence, CAPTCHA, Demographic Characteristics, Internet, Response Time, Statistical Analysis

Received: 12 January 2020, Revised 3 March 2020, Accepted 1 April 2020

DOI: 10.6025/isej/2020/7/1/1-8

1. Introduction

CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart) is a program representing a challenge-answer to the given test, which is used to realize if the solver is an Internet user (human) or a computer program (computer robot). This process includes the following elements: (i) The computer as a server, which generates the CAPTCHA test, (ii) Internet users or computer program which try to correctly solve the given task, (iii) The computer which evaluates the answer to the CAPTCHA in the format Yes/No (correctly/incorrectly solved). Typically, the CAPTCHA task is accustomed to the humans. Hence, there is a greater possibility that humans will solve this task compared to computer robots abbreviated as bots. Hence, the aim of the CAPTCHA program is to differentiate Internet users from bots [1].

The application of CAPTCHA program is useful in the following areas: (i) Online systems, (ii) The creation of free email accounts, (iii) Online pooling, and (iv) Online system for buying tickets, etc [2].

Still, the CAPTCHA should fulfill certain elements, such as: (i) The solving of CAPTCHA should not rely on the user's knowledge of certain language, (ii) The solving of CAPTCHA should not depend on the user's age, (iii) CAPTCHA should make an automatic evaluation of the correctness, (iv) The user's privacy should not be violated, and (v) It should be easy for Internet users to be solved unlike bots [3].

The related works on CAPTCHA often employ statistical approaches treating their various aspects. They can be partitioned taking into account their properties in the following areas: (i) Security, (ii) Practicality, and (iii) Usability [4].

Security represents the main concern to the CAPTCHA programmers. It represents a central problem of CAPTCHA, but it is not the only one that is of a great importance. Practicality is connected to the way of creating certain types of CAPTCHA. Again, it has greater concerns of programmers than CAPTCHA users. The usability represents the main problem related to the use of the CAPTCHA. Accordingly, it especially concerns the CAPTCHA users. Hence, this study is used to uncover the elements of CAPTCHA usability, which represents the main concern of the Internet users. In this way, an objective analysis of a certain type of CAPTCHA can facilitate better understanding the user-centric relation between computer and man, i.e. CAPTCHA and Internet user which will contribute to innovate and improve CAPTCHA elements to be more accustomed to the Internet users unlike bots.

This paper is organized in the following manner. Section II presents the CAPTCHA types. Section 3 describes the experiment. Section 4 gives the results of the experiment and discussed them. Section 5 draws conclusions and points out the direction of future works.

2. CAPTCHA Types

All CAPTCHA types can be divided into five typical groups: (i) Text-based CAPTCHA, (ii) Image-based CAPTCHA, (iii) Audio-based CAPTCHA, (iv) Video-based CAPTCHA and (v) Other types of CAPTCHA [5].

Text-based CAPTCHA asks the Internet users to input exact combination of the given characters. This type of CAPTCHA is the most widespread one. In order to reduce its vulnerability to bot attacks, many distorted elements are incorporated. Unfortunately, the text-based CAPTCHA can be successfully attacked by bot due to the solid OCR (Optical Character Recognition) programs. Figure 1 shows an example of the text-based CAPTCHA.

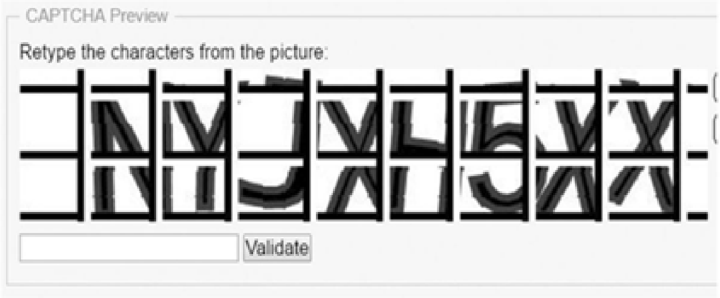


Figure 1. An example of the text-based CAPTCHA

Image-based CAPTCHA is considered as one of the most advanced and safest types of CAPTCHA. It requires from the users to find out a certain image from a list of images and point to it. Due to that, its elements include the image details.

It represents a relatively easy task to be solved by Internet users unlike bots. Figure 2 illustrates an example of the image based CAPTCHA.

Audio-based CAPTCHA includes an "audio element" whose purpose is an audio reproduction of characters that the user should have to input in order to solve the CAPTCHA. This type of CAPTCHA is especially designed for the people with disabilities. Unfortunately, the audio-based CAPTCHA is mostly attacked by speech and recognition algorithms in approximately 70% of cases. Fig. 4 illustrates an example of the audio-based CAPTCHA with "audio element" in the top right corner.

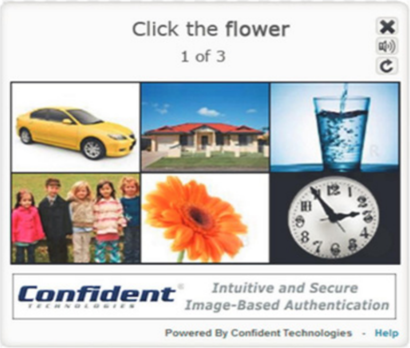

Click the flower
1 of 3
Confident
TECHNOLOGIES
Intuitive and Secure
Image-Based Authentication
Powered By Confident Technologies - Help

Figure 2. Illustrates an example of the image based CAPTCHA

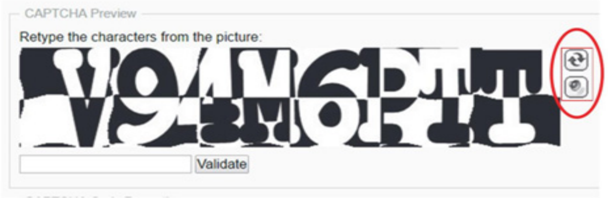


Figure 3. An example of the audio-based CAPTCHA

Video-based CAPTCHA contains text information embedded into the video. Hence, it is a video which includes a passing text given in specific color compared to video background. The user should recognize the given passing text and type it. The modern OCR programs challenge this task, making this CAPTCHA vulnerable to bot attacks. Figure 4 illustrates an example of the video-based CAPTCHA.

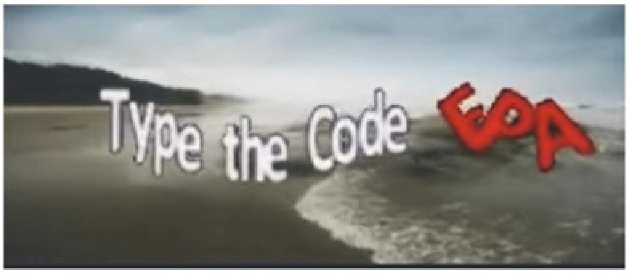
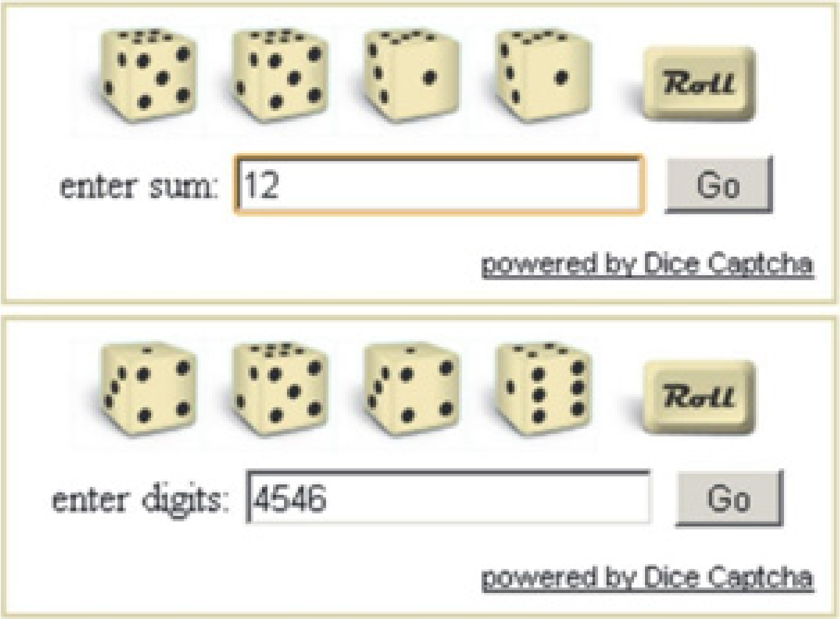


Figure 4. An example of the video-based CAPTCHA

Other types of CAPTCHA represent those CAPTCHAs that cannot be part of the previous standardization. Fig. 5 illustrates the examples of such types of CAPTCHA.

3. Experiment

The CAPTCHA experiment is conducted on 190 Internet users. It is divided in two different experiments solving two different Dice CAPTCHAs (Dice 1 and 2). The first experiment is based on Dice CAPTCHAs tested on a community of 90 laptop users aged from 29 to 62 years. The laptop used for the experiment is Lenovo B51 with the following characteristics: (i) 15.6" wide screen, (ii) CPU Quad-core 2.4 GHz Celeron, (iii) 4 GB of RAM, (iv) 500 GB of internal memory, and (v) Operating system Microsoft Windows 7. The second experiment is based on Dice CAPTCHAs tested on a community of 100 tablet users aged from 28 to 55 years. The tablet used for the experiment is Lenovo IdeaTab A3000 with the following characteristics: (i) 7" wide screen, (ii) CPU Quad-core 1.2 GHz Cortex-A7, (iii) 1 GB of RAM, (iv) 16 GB of internal memory, and (v) Operating system Android.



enter sum: 12

enter digits: 4546

Figure 5.. An Example of other type of CAPTCHA on the dice CAPTCHA Samples (Dice 1 at the top, and Dice 2 at the bottom)

All Internet users represent volunteer students, employees, clerks, teachers and engineers, who signed an online consent form before starting the experiment. Accordingly, they gave their consent to anonymously provide and use their data only for research and study purposes. Each of them was required to solve Dice CAPTCHA, and the response time to find the solution to the CAPTCHA was registered. The users are partitioned taking into account their various demographic factors: (i) Age, (ii) Education level, (iii) Gender. The experimental results are then statistically processed. The obtained results are then compared to evaluate (dis)advantages of using Dice CAPTCHA by the different Internet users' groups (laptop or tablet).

4. Results and Discussion

4.1. Hypotheses

It is worth noting that the solution time for "ideal" CAPTCHA should not depend on the age, education and gender differentiation. However, if any CAPTCHA can satisfy these elements, then it doesn't mean that it can be solved quickly and easily. According to previous facts, the following four hypotheses are proposed according to the given demographic characteristics:

Hypothesis 1 (H1) - There exists a statistically significant difference between users' groups (laptop vs. tablet) in average response time to solve the CAPTCHA.

Hypothesis 2 (H2) - There exists a statistically significant difference between age groups in solving the CAPTCHA,

Hypothesis 3 (H3) - The group of Internet users with higher education will have a faster response time in solving the CAPTCHA,

Hypothesis 4 (H4) - There exists a statistically significant difference between gender groups in solving the CAPTCHA,

4.2. Experimental Results

The first results of the experiment are given in Tables I-II. These tables give a descriptive analysis of the obtained results for 90 laptops' and 100 tablets' Internet users concerning CAPTCHA Dice 1 and They are obtained by *KolmogorovSmirnov* test, which test the unknown distribution and check the normality assumption in the analysis of variance [6].

	Gender	Age	Education	Dice 1	Dice 2
N	90	90	90	90	90
Mean	1.16	1.70	1.88	7.644	6.514
SD	0.364	0.461	0.329	2.554	2.203
Asymp. Sig. (2-tailed)	0.000	0.000	0.000	0.046	0.003

Table 1. One-sample Kolmogorov-smirnov Test for Laptop Users

The most important information represents the measure Asymp. Sig. (2-tailed). It defines the statistical significance of the analyzed data. Because it is smaller than 0.05, then obtained results are statistically significant. Also, it is worth noting that the average time to solve CAPTCHA Dice 1 is 7.6444, while CAPTCHA Dice 2 is solved in 6.514 seconds by laptop users.

	Gender	Age	Education	Dice 1	Dice 2
N	100	100	100	100	100
Mean	1.59	1.30	1.44	12.090	8.590
SD	0.494	0.461	0.499	5.874	4.360
Asymp. Sig. (2-tailed)	0.000	0.000	0.000	0.005	0.018

Table 2. One-sample Kolmogorov-smirnov Test for Tablet Users

From Table 2, the measure Asymp. Sig. (2-tailed) is again lower than reference of 0.05, which determines the statistical significance of the analyzed data. Furthermore, it is worth noting that the average time to solve CAPTCHA Dice 1 is 12.090, while CAPTCHA Dice 2 is solved in 8.590 seconds by tablet users.

From Tables 1-2, it is quite clear that exists a statistically significant difference in the response time to solve CAPTCHA Dice1 and Dice 2 between laptop and tablet users' group. Obviously, Dice CAPTCHA is more convenient to be solved on a laptop than on a tablet computer. It is proved by statistical significant population. Hence, H1 is proved.

4.3. Statistical Test

The Mann-Whitney U test is a non-parametric test which can be used to (dis)prove a null-hypothesis H0 and a research hypothesis H1. Essentially, this test is used to compare differences between two independent groups N1 and N2. To be used, some pre-assumptions should be valid: (i) Input should be composed of two categorical independent groups N1 and N2, (ii) Output should be ordinal or continuous, (iii) There should be no correlation between groups N1 and N2, and (iv) The input variables should not be normally distributed. The Mann-Whitney U test considers 3 important measures: (i) pvalue, (ii) U value, and (iii) Z value.

The p-value is the first crucial measure of this statistical test. Its value can be interpreted as follows: (i) p=0.05 shows a weak evidence against the null-hypothesis of the test. As a consequence, the null-hypothesis of the test is proved, while research hypothesis H1 is disproved. U value is calculated as:

$$U = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1 \ , \tag{1}$$

where U represents the result of the Mann-Whitney U test. Accordingly, n_1 is the size of the independent group N_1, n_2 is the size of the independent group N_2, and R_1 represents the sum of ranks of group N_1. If U value is higher than the critical U value, then the two groups N_1 and N_2 will have the same score distributions, otherwise the two distributions N_1 and N_2 will be different in some aspect. Critical value U is important only for small size distributions, where the number of their elements is up to 20. If the group is larger than 20, then U value approaches to normal distribution. In that case, the Z value has importance. It is calculated as:

$$Z = \frac{U - n_1 n_2 / 2}{\sqrt{n_1 n_2 (n_1 + n_2 + 1)/12}} \ . \tag{2}$$

If the absolute value of Z is lower than 1.96, then the two groups N_1 and N_2 will have the same score distributions, otherwise the two distributions of N_1 and N_2 will be dissimilar in some way. Accordingly, if Z is lower than 1.96 research hypothesis is disproved, otherwise it is proved.

4.4. Analysis of the Results and Discussion

The results obtained by statistically processing (MannWhitney U test) of experimental data for the age characteristic of the laptop/tablet users are given in Table 3.

The first relevant measure, which has to be evaluated is Asymp. Sig. (2-tailed). For laptop users as well as for tablet users concerning CAPTCHA Dice 1 and 2 it is higher than 0.05. Accordingly, this analysis is not statistically significant. Hence, H_2 is not proved.

The results obtained by statistically processing (MannWhitney U test) experimental data for the education demographic characteristic of the laptop and tablet users are given in Table 4.

Again, the measure Asymp. Sig. (2-tailed) is evaluated the first. For laptop users as well as for tablet users concerning CAPTCHA Dice 1 and 2 it is higher than 0.05. Hence, this analysis is not statistically significant. This leads that H3 is not proved.

Laptop	**Age (y/o)**	**N**	**Mean rank (1/2)**	**Z**	**Asymp. Sig. (2-tailed)**
Dice 1	27/63	90	49.04/43.98	98 -0.842	0.400
Dice 2	27/63	90	53.48/42.08 -	-1.899	0.058
Tablet	Age (y/o)	N	Mean rank (1/2)	Z	Asymp. Sig. (2-tailed)
Dice 1	70/30	100	48.21/55.83	-1.208	0.227
Dice 2	70/30	100	49.06/53.87 -	-0.764	0.445

*y-younger, o-older, 1-group 1 (younger), 2-group 2 (older)

Table 3. Mann-whitney u test (laptop/tablet users) for the age

The results obtained by statistically processing (MannWhitney U test) experimental data for the gender demographic characteristic of the laptop and tablet users are given in Table 5.

Laptop	Age (y/o)	N	Mean rank (1/2)	Z	Asymp. Sig. (2-tailed)
Dice 1	79/11	90	44.08/55.73	-1.387	0.165
Dice 2	79/11	90	43.86/57.27	-1.596	0.110
Tablet	Educ. (h/s)	N	Mean rank (1/2)	Z	Asymp. Sig. (2-tailed)
Dice 1	44/56	100	53.75/47.95	-0.997	0.319
Dice 2	44/56	100	50.63/50.40	-0.038	0.969

*h-higher, o-secondary, 1-group 1 (higher), 2-group 2 (secondary)

Table 4. Mann-whitney U Test (Laptop/Tablet Users) for the Education

Laptop	Age (y/o)	N	Mean rank (1/2)	Z	Asymp. Sig. (2-tailed)
Dice 1	14/76	90	52.79/44.16	-1.137	0.256
Dice 2	14/76	90	46.61/45.30	-0.173	0.863
Tablet	Gender (m/f)	N	Mean rank (1/2)	Z	Asymp. Sig. (2-tailed)
Dice 1	59/41	100	52.47/47.66	-0.820	0.412
Dice 2	59/41	100	48.58/53.27	-0.800	0.424

*m-male, f-female, 1-group 1 (male), 2-group 2 (female)

Table 5. Mann-whitney U Test (Laptop/Tablet Users) for the Gender

From Table 5 the measure Asymp. Sig. (2-tailed) is again higher than reference value 0.05. Hence, for laptop users as well as for tablet users concerning CAPTCHA Dice 1 and 2 the given analysis is not statistically significant. Accordingly, H4 is not proved.

From the aforementioned, the H1 is only proved, while H2, H3 and H4 are not proved. Because, the postulate of "ideal" CAPTCHA is to be solved in reasonable time (less than 30 sec. [5]), and the solution time should not depend on the age, education and gender differentiation, the Dice CAPTCHA represents a good direction toward creating an "ideal" CAPTCHA. However, it is worth noting that using CAPTCHA on different computer types should also diminish differences between solution time of certain CAPTCHA. In our case, solution time of Dice CAPTCHA between laptop and tablet users is almost 50% less in favor of laptop users. Taking into account this information, Dice CAPTCHA is more accustomed to the laptop than tablet Internet users. Hence, Dice CAPTHA can be considered only as the first step in right direction toward creating an "ideal" CAPTCHA.

5. Conclusion

The paper analyzed the response time of Internet laptop and tablet users in solving the Dice CAPTCHA version 1 and 2. To research the given topic, an experiment was conducted on 190 users. It was divided into two parts: (i) testing of 90 laptop users in solving Dice CAPTCHA 1 and 2, and (ii) testing of 100 tablet users in solving Dice CAPTHA 1 and 2. Then, the obtained results were statistically processed. According to the results, four hypotheses were established, which should be proved or disproved. All hypotheses were closely related to the elements of an “ideal” CAPTCHA. Using statistical tools, a descriptive statistical analysis and the results of Mann-Whitney U test were used for proving and disproving the given hypotheses. At the end, the H1 hypothesis was only proved, while the other ones were rejected. In spite of the obtained result, which represents the main elements of an “ideal” CAPTCHA, due to rather different time in solving Dice CAPTCHA between laptop and tablet users, this type of CAPTCHA cannot be used as an example of “ideal” CAPTCHA. But, because of some overlapping with the characteristics of an “ideal” CAPTCHA, the Dice CAPTCHA is a good start and a right direction toward creating the real “ideal” CAPTCHA.

Acknowledgement

This study was partially funded by the Grant of the Ministry of Education, Science and Technological Development of the Republic of Serbia, as a part of the project TR33037.

References

[1] Von Ahn, M. Blum., Langford, J. (2004). Telling Humans and Computers Apart Automatically, *Communication of ACM*, 47 (2), p 47-60, 2004.

[2] The CAPTCHA test. http://en.wikipedia.org/wiki/CAPTCHA

[3] von Ahn, L., Blum, M., Hopper, N., Langford, J. (2003). CAPTCHA: Using Hard AI Problems for Security, *In*: Proceedings of Eurocrypt, Warsaw, Poland, p 294-311, 2003.

[4] Baecher, P., Fischlin, M., Gordon, L., Langenberg, R., Lutzow, M., Schroder, D. (2010). CAPTCHAs: The Good, the Bad and the Ugly, Sicherheit, pp. 353-365, 2010.

[5] Brodic, D., Petrovska, S., Jevtic, M., Milivojevic, Z. N. (2016). The Influence of the CAPTCHA Types to its Solving Times, *In*: Proceeding of 39th MIPRO, Opatija, Croatia, p 1274-1277.

[6] Conover, W. J. (1999). *Practical Nonparametric Statistical*, p 428- 433, *John Wiley & Sons*, Inc. New York, 1999.