UNITECH 2017 GABROVO

INTERNATIONAL SCIENTIFIC CONFERENCE
17-18 November 2017, GABROVO

THE STATISTICAL CORRELATION OF DEMOGRAPHIC FACTOR INFLUENCE TO SOLVING FACIAL EXPRESSION CAPTCHAS

Darko Brodić
Technical Faculty in Bor
V.J. 12, 19210 Bor, Serbia

Branislav Ivanov
Technical Faculty in Bor
V.J. 12, 19210 Bor, Serbia

Milena Jevtić
Technical Faculty in Bor
V.J. 12, 19210 Bor, Serbia

Abstract

The paper studies the influence of different facial expression CAPTCHA types to users concerning their solving time. The experiment includes over 100 participants that use a smartphone to solve 4 different image type CAPTCHAs with the element of various facial expressions. All four tested CAPTCHAs are freely available on the Internet link. Furthermore, the obtained results are statistically processed. According to the analysis, the obtained results regarding Pearson's and Spearman's correlation coefficients are discussed and related conclusions are drawn.

Keywords: CAPTCHA, facial expression, artificial intelligence, statistical analysis.

INTRODUCTION

Completely Automated Public Turing test to tell Computers and Humans Apart (CAPTCHA) is a program based puzzle which is proposed to be a test that should be accurately solved by Internet users (humans) and hardly solved by computer programs commonly abbreviated as bots. Up to now, many different CAPTCHA types have been developed. One of them is the image-based CAPTCHA. It consists of different image elements integrated into a CAPTCHA [1]. Because it is based on image details, it represents an almost impossible task to be solved by bots [2]. As an extension to image recognition CAPTCHA the various face recognition CAPTCHAs have been developed [3]-[10]. Essentially, this type of image-based CAPTCHA is its subgroup which includes the elements of a face detection bringing a higher level of security.

FR-CAPTCHA is based on finding matching pairs of human faces in an image [3]. The easiness of solving proposed CAPTCHA confirms that humans can accurately recognize matched faces even under severe distortion. Essentially, each of FR-CAPTCHA images has embedded multiple face images. The user's assignment is to mark the location of two matching faces, which represents the face images belonging to the same individual.

The face CAPTCHA, which is more accustomed to solved by the touch screen devices like smartphones and tablets is proposed in [4]. This CAPTCHA contains face and non-face images. Particularly, the user should select the correct distorted human face image. Distortion is included in the CAPTCHA by: (i) erosion, (ii) increase brightness, (iii) periodic noise, (iv) speckle noise, (v) resolution modification, (vi) width scaling, and (vii) height scaling. If the user selects the non-face image, then the solving of CAPTCHA is failed.

FaceCAPTCHA that includes the elements of gender-indiscernible faces are proposed in [5]. Particularly this CAPTCHA asks the users to identify the gender of face images whose gender is difficult to be recognized by bots. FaceCAPTCHA incorporates gender-indiscernible faces that are sequentially displayed including four response buttons, which are labeled as: (i) male, (ii) female, (iii) not sure, and (iv) not human.

A new CAPTCHA, which also works on the principle of facial expression detection was

proposed in [6]. It exploits a real-time interaction from the laptop's front camera to get an input image. The proposed CAPTCHA implements three basic emotions: (i) neutral (or normal), (ii) happy, and (iii) sad. If the desired facial expression is matched with the input expression of user, then the CAPCTHA is solved.

Another face-based image CAPTCHA called fgCAPTCHA that includes the touch-based input methods are proposed in [7]. It is quite clear that it prefers the solving process by mobile devices like tablets and smartphones, which is allowed by genetically optimized face detection tests.

Another face recognition CAPTCHA, which includes a pair of human face images is proposed in [8]. In this case, the pair of images belongs to the same subject. Typically, they are embedded in a complex background for better security.

A new face image-based CAPTCHA, which is generated by a specific algorithm is proposed in [9]. In this case, the users should correctly identify visually-distorted human faces without selecting any other part of the image. The visual distortion of human faces is achieved using the stripes, strikeout, rotation, and blending into the background. Particularly, these faces are embedded in a complex background. In this way, the proposed algorithm generates a CAPTCHA that offers better recognition by the humans and lower level of the machine attacks compared to existing approaches.

A new approach to the face-based CAPTCHA is proposed as AgeCAPTCHA [10]. It applied a scheme to annotate images of human faces using their age groups. The results showed that users can quickly and accurately solve this type of CAPTCHA.

EXPERIMENT

The CAPTCHA experiment is conducted on 100 Internet users. The Internet users differ by the following demographic features: (i) age, (ii) years of Internet use, (iii) daily Internet use, and (iv) gender. The testing process includes the solving of four different face expression CAPTCHA publicly available on the site [11].

Figure 1 shows these four CAPTCHAs.

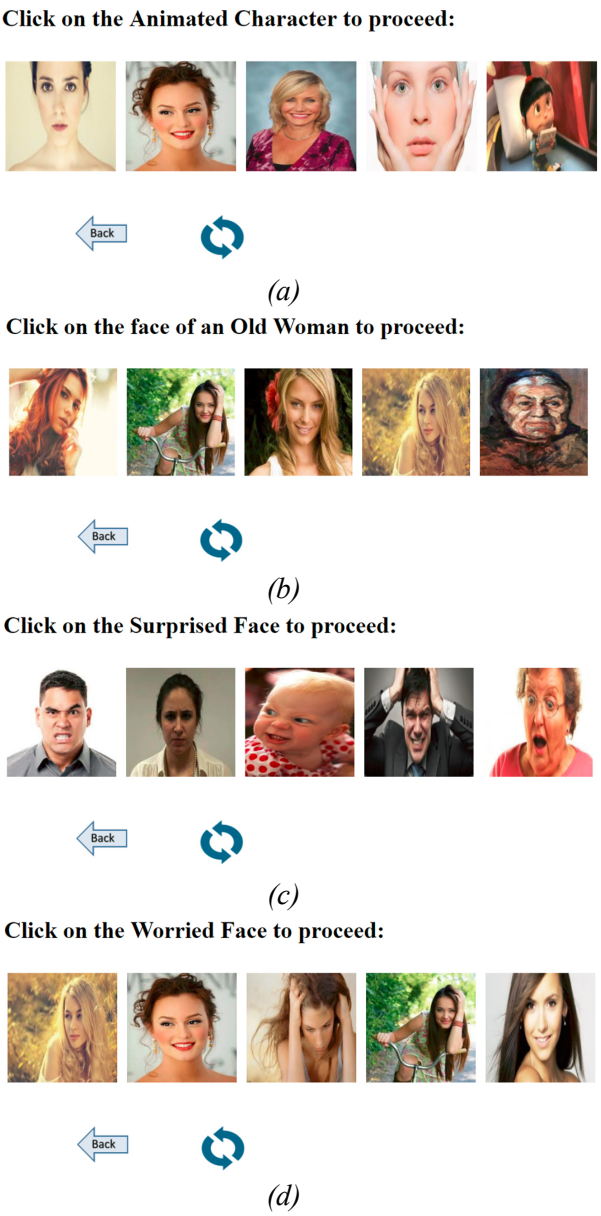

Fig. 1. *Face expression CAPTCHAs: (a) CAPTCHA 1 -animated character, (b) CAPTCHA 2 -face of an old woman, (c) CAPTCHA 3 - surprised face, and (d) CAPTCHA 4 - worried face*

Each of CAPTCHA includes five different images with the different facial expressions of people, but only one is the correct for solving the CAPTCHA test.

Internet users have 5 times to accurately solve the given CAPTCHA. Accordingly, it is measured their time to accurately solve certain CAPTCHA (in seconds). The experiment is performed on smartphone with 5" wide screen and Android operating system. All Internet users are agreeing to participate in the experiment and that obtained experimental results can be used for the research purposes.

HYPOTHESIS AND STATISTICAL TESTS

In this study, we are going to give the answers to the following hypotheses (H1 to H4):

1. Hypothesis 1 (H1): Does the age of the participants influence the solving time of the facial CAPTCHA?
2. Hypothesis 2 (H): Do the participant's years of Internet use influence the solving time of the facial CAPTCHA?
3. Hypothesis 3 (H3): Does the participant's daily use of the Internet's influence the solving time of the facial CAPTCHA?
4. Hypothesis 4 (H4): Do the gender influence the solving time of the facial CAPTCHA?

To prove or disprove these hypotheses, the analysis of the relation between specific input and output variables is immanent. The following relations should be researched:

H1 - relation between age and CAPTCHA 1 to 4,

H2 - relation between years of Internet use expressed in years and CAPTCHA 1 to 4,

H3 - relation between daily Internet use expressed in hours and CAPTCHA 1 to 4,

H4 - relation between gender and CAPTCHA 1 to 4.

The correlation between input and output variables will be investigated using Pearson's linear correlation coefficient. It represents a parametric technique for testing the connectivity of data that can be used to (dis)prove a null-hypothesis H_0 and a research hypothesis H_1. Basically, this test is used to compare linear connection between two variables. The p-value is the first crucial measure in this statistical test. Its value can be interpreted as follows:

- $p < 0.05$ shows a strong evidence against the null-hypothesis. As a consequence, the null-hypothesis of the test H_0 is disproved, while research hypothesis H_1 is proved;
- $p >= 0.05$ shows a weak evidence against the null-hypothesis of the test. As a consequence, the null-hypothesis of the test H_0 is proved, while research hypothesis H_1 is disproved.

The linear correlation between continuous input and output variables is validated using Pearson's coefficient of correlation *r*:

$$r = \frac{\sigma_{xy}}{\sigma_x \sigma_y}, \quad (1)$$

where σ_{xy} is covariance, while σ_x and σ_y are standard deviation of the *x* and *y* variables, respectively. If there exists a correlation between two variables *x* and *y*, then for H_0 follows $r = 0$, and for H_1 follows $r \neq 0$.

The Pearson's correlation coefficient *r* can have values between +1 and -1. If the value is equal to 0, then there is no correlation between the two variables. Table 1. shows the strength correlation between the variables.

Table 1. *Strength of connection (correlation)*

Coefficient r	Strength
0.70 <= |r|<= 1.00	Strong connection
0.30 <= |r|< 0.70	Medium connection
0.00 < |r|< 0.30	Small connection
r = 0	There is no linear connection (does not exclude the existence of a nonlinear form of connection)

Because the hypothesis H4 use non-continuous, i.e. scaled variables, then the correlation between the two variables should be examined by non-parametric technique using Spearman's rank correlation coefficient. Spearman's rank correlation coefficient r_s is calculated as:

$$r_s = 1 - \frac{6\sum d_i^2}{n(n^2-1)}, \quad (2)$$

where d_i is the difference between the two ranks of each observation, *n* is the number of observations.

RESULTS AND DISCUSSION

The obtained data are subjected to the statistical processing (Pearson the correlation coefficient and Spearman's rank correlation coefficient) in MATLAB R2017a.

Table 2 presents the results obtained by statistically processing of experimental data regarding Pearson's correlation coefficient for the age and corresponding CAPTCHAs 1 to 4.

Table 2. *Pearson's correlation coefficient (Age vs. CAPTCHA solving time)*

	CAPTCHA 1	CAPTCHA 2	CAPTCHA 3	CAPTCHA 4
Pearson's correlation coefficient	0.432	0.322	0.409	0.250
Correlation	medium positive	medium positive	medium positive	weak positive
Sig. (2-tailed)	0.000	0.001	0.000	0.010

The first relevant measure, which should be evaluated is Sig. (2-tailed). If it has smaller value than 0.05, then the data samples are interpreted as representative sample leading to statistically significant result. It proves that there is a correlation between the participant's age and the speed of correctly solving the CAPTCHA. However, there exists a weak correlation in the case of CAPTCHA 4, and medium correlation in the case of CAPTCHA 1, 2, and 3. However, in all cases there exists a positive correlation, which means that the increase of users' age designates the increase in CAPTCHA solving time.

Table 3 presents the Pearson's correlation coefficient obtained by statistical processing concerning years of Internet usage and corresponding CAPTCHAs 1 to 4.

Table 3. *Pearson's correlation coefficient (Internet use vs. CAPTCHA solving time)*

	CAPTCHA 1	CAPTCHA 2	CAPTCHA 3	CAPTCHA 4
Pearson's correlation coefficient	**-0.221**	**-0.308**	-0.167	-0.074
Correlation	weak negative	medium negative	-	-
Sig. (2-tailed)	**0.023**	**0.001**	0.087	0.452

The first relevant measure to be assessed is again Sig. (2-tailed). It has value less than 0.05 for CAPTCHA 1 and 2. Hence, only the results obtained for CAPTCHA 1 and 2 are statistically significant. It proves that there is a correlation between the year of use the Internet and the solving time for the CAPTCHA 1 and 2, only. The obtained correlation is weak for the CAPTCHA 1, and medium for CAPTCHA 2. It is worth noting that correlation is negative in all cases. It can be interpreted that participants who use the Internet longer (more years) need less time to solve CAPTCHA 1 and 2. Because of Sig. (2-tailed) is higher than reference value of 0.05 for CAPTCHA 3 and 4, then the analysis for these two CAPTCHAs is not statistically significant. Furthermore, it means that the years of Internet use cannot be interpreted to affect the solving time of CAPTCHA 3 and 4.

Table 4 presents the Pearson's correlation coefficient obtained by statistical processing linked between daily Internet usage expressed in hours and corresponding CAPTCHAs 1 to 4.

Table 4. *Pearson's correlation coefficient (Daily Internet use vs. CAPTCHA solving time)*

	CAPTCHA 1	CAPTCHA 2	CAPTCHA 3	CAPTCHA 4
Pearson's correlation coefficient	**-0.260**	**-0.235**	-0.158	**-0.194**
Correlation	weak negative	weak negative	-	weak negative
Sig. (2-tailed)	**0.007**	**0.015**	0.106	**0.046**

Sig. (2-tailed) is less than 0.05 in the example of CAPTCHA 1, 2 and 4. It proves that there is a statistically significant correlation between the daily Internet use and the solving time of CAPTCHA 1, 2 and 4. Still, the correlation is rather weak and negative in all three examples. Hence, if the participant uses Internet more hours per a day, then he/she will need less time to solve CAPTCHA 1, 2, and 4. On the contrary, CAPTCHA 3 has Sig. (2-tailed) higher than reference value 0.05. Accordingly, its analysis is not statistically significant leading to conclusion that the daily use of the Internet cannot be interpreted to affect the solving time of CAPTCHA 3.

Table 5 shows the Spearman's rank correlation coefficient obtained by statistical processing of experimental data between gender and solving time of CAPTCHAs 1 to 4.

Table 5. *Spearman's correlation coefficient (Gender vs. CAPTCHA solving time)*

	CAPTCHA 1	CAPTCHA 2	CAPTCHA 3	CAPTCHA 4
Spearman's correlation coefficient	0.114	0.009	0.035	-0.053
Sig. (2-tailed)	0.247	0.929	0.722	0.587

Unfortunately, in all cases Sig. (2-tailed) is higher than referent value of 0.05. Therefore, this analysis is not statistically significant, which leads to the conclusion that the gender cannot be interpreted to affect the solving time of CAPTCHAs 1 to 4.

CONCLUSION

Among four hypotheses, only the H1 hypothesis is completely proved. It shows that there exists a correlation between the age of Internet users and their CAPTCHA solving time. In that sense, the younger users usually solve facial expression CAPTCHAs more quickly than older users. Hypothesis H2 is not proved. Essentially, it is partially proved only for CAPTCHA 1 and 2 in the sense that these

CAPTCHAs can be more quickly solved by users with more years of Internet experience. Furthermore, the hypothesis H3 is not proved. Basically, it is partially proved for CAPTCHA 1, 2 and 4. At the end, the hypothesis H4 is not completely proved. It means that we have no evidence that the gender has any impact on facial expression CAPTCHA solving time.

ACKNOWLEDGEMENT

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

REFERENCE

[1] M. Chew, J. D. Tygar, Image Recognition CAPTCHAs, International Conference on Information Security (ISC 2004), LNCS, volume 3225 (K. Zhang and Y. Zheng (Eds.)), pp. 268-279, Springer-Verlag Berlin Heidelberg, 2004.

[2] D. Brodić, S. Petrovska, M. Jevtić and Z. N. Milivojević, The Influence of the CAPTCHA Types to Its Solving Times, 39th International Covention on Information and Communication Technology, Electronics and Microelectronics (MIPRO 2016), Opatija, Croatia, pp. 1274-1277, 2016.

[3] G. Goswami, B. M. Powell, M. Vatsa, R. Singh, A. Noore, FR-CAPTCHA: CAPTCHA Based on Recognizing Human Faces, PLoS One, 2014, Vol. 9, No. 4, e91708. doi: 10.1371/journal.pone.0091708.

[4] N. B. Kadu, Pushkar U. Wable, Swati K. Wale, Dipti B. Tambe, Amol P. Londhe, Face Captcha, International Journal of Engineering and Computer Science, 2015, Vol. 4, No. 3, pp. 10775-10778.

[5] J. Kim, S. Kim, J. Yang, J. Ryu, K. Y. Wohn, FaceCAPTCHA: a CAPTCHA that identifies the gender of face images unrecognized by existing gender classifiers, Multimedia Tools and Applications, 2014, Vol. 72, No. 2, pp. 1215–1237.

[6] A. Sinha, S. Tarar, Implementation of Facial Expression Detection as a CAPTCHA, International Journal of Computer Science And Technology, Vol. 7, No. 2, 2016, pp. 69-72.

[7] B. M. Powell, G. Goswami, M. Vatsa, R. Singh, A. Noore, fgCAPTCHA: Genetically Optimized Face Image CAPTCHA, IEEE Access, Vol. 2, 2014, pp. 473-484.

[8] G. Goswami, R. Singh, M. Vatsa, B. Powell, A. Noore, Face Recognition CAPTCHA, IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems (BTAS 2012), Arlington, USA, 2012, pp. 412-417.

[9] G. Goswami, B. M. Powell, M. Vatsa, R. Singh, A. Noore, FaceDCAPTCHA: Face detection based color image CAPTCHA, Future Generation Computer Systems, 2014, Vol. 31, 59-68.

[10] J. Kim, J. Yang, K. Wohn, AgeCAPTCHA: an Image-based CAPTCHA that Annotates Images of Human Faces with their Age Groups, KSII Transactions on Internet and Information Systems, Vol. 8, No. 3, 2014, pp. 1071-1092.

[11] http://www.captchasamples.altervista.org/