Routledge
Taylor & Francis Group

Journal of Applied Security Research

ISSN: 1936-1610 (Print) 1936-1629 (Online) Journal homepage: http://www.tandfonline.com/loi/wasr20

Effects of Text Rotation, String Length, and Letter Format on Text-based CAPTCHA Robustness

Chatpong Tangmanee

To cite this article: Chatpong Tangmanee (2016) Effects of Text Rotation, String Length, and Letter Format on Text-based CAPTCHA Robustness, Journal of Applied Security Research, 11:3, 349-361, DOI: 10.1080/19361610.2016.1178553

To link to this article: http://dx.doi.org/10.1080/19361610.2016.1178553

JOURNAL OF APPLIED SECURITY RESEARCH
2016, VOL. 11, NO. 3, 349–361
http://dx.doi.org/10.1080/19361610.2016.1178553

Routledge
Taylor & Francis Group

TECHNOLOGY

Effects of Text Rotation, String Length, and Letter Format on Text-based CAPTCHA Robustness

Chatpong Tangmanee

Chulalongkorn Business School, Chulalongkorn University, Bangkok, Thailand

ABSTRACT

CAPTCHA, standing for Completely Automated Public Turing test to tell Computers and Humans Apart, has been adopted as a security check. One way to assess robustness of a text-based CAPTCHA test is to use optical letter reader (OCR) software to read it. If the reading fails, the design is robust. Though many design features have been examined, no research has investigated the effects of text rotation, string length, or letter format on CAPTCHA robustness. Fourteen hundred sixty four text-based CAPTCHA tests were created based on the three variables. The main effects of all three variables and few of the interaction effects on robustness were significant. The findings have both theoretical and practical contributions.

KEYWORDS

CAPTCHA; text rotation; string length; letter format; robustness

Problem statement

Short for Completely Automated Public Turing test to tell Computers and Humans Apart, CAPTCHA has been adopted as a security mechanism to minimize the chance of people abusing a Web-based service. Such abuse includes writing automated programming script to sign up for thousands of accounts at a free mail service or flooding a Web board with hundreds of commercial messages. Ideally, only a human being can correctly respond to CAPTCHA, while a computer (or automated software) should not. While there are many types, text-based CAPTCHA test is most widely accepted (Ahn, Maurer, McMillen, Abraham, & Blum, 2008, 2013; Lin, Huang, Bell, & Lee, 2011).

The best text-based CAPTCHA must be (a) easy for a person to solve and (b) robust if an attack script wants to break in. It is difficult to achieve a balanced CAPTCHA design. A fair amount of security-related research has examined the human side of its design (Bursztein, Bethard, Fabry, Mitchell, & Jurfsky, 2010; Gao, Wang, Fan, Qi, & Liu, 2014; Hsu & Lee, 2011). Yet, a relatively small number of studies have investigated CAPTCHA robustness which will be the main focus of the current study. The wide acceptance of text-based CAPTCHA is a result of an attempt to improve security in many online services such as free e-mail providers

CONTACT Chatpong Tangmanee chatpong@cbs.chula.ac.th Chulalongkorn Business School, Chulalongkorn University, Bangkok 10330, Thailand.

Figure 1. Noise added to a text-based CAPTCHA (from http://www.authenticeducation.com.au/blog/digital-marketing/how-to-add-facebook-comment-box-to-a-website).

or Web-board administrators or to disseminate knowledge through such a project as the reCAPTCHA (Ahn et al.).

A capable text-based CAPTCHA test must be robust to any attack from automated software, commonly known as a bot. An attacker will prepare an automated script to detect an array of texts on CAPTCHA. The robust CAPTCHA must, therefore, be sufficiently protected from attack, yielding an acceptable amount of security. A secure or robust CAPTCHA will be such that a bot will be unable to correctly interpret its letters.

To crack a CAPTCHA test, an attack script typically follows three steps: preprocessing, segmentation, and recognition (Bursztein, Martin, & Mitchell, 2011). First, preprocessing will ensure CAPTCHA is readable so that subsequent segmentation and recognition processes are feasible. Actions involved in preprocessing include adjusting the contrast between text and background, changing text color or removing noise such as small lines on texts. In Figure 1, one curly line is placed on the distorted string of dwaMSx. The curly line is an example of noise.

Once preprocessing is successful, the attack script will further attempt to segment each letter individually. As a result, a robust CAPTCHA design must be able to prevent segmentation. Such a design may either connect letters together so that they are largely overlapping and the attack script will not be able to effectively separate one from the other or use similar colors on both text and background. In Figure 2,

Figure 2. Overlapping letters in CAPTCHA (from http://4.bp.blogspot.com/-fkkyCwsnbuc/T-KN8SoVIzI/AAAAAAAAIDE/uZLRTgWjYhA/s1600/Captcha.png).

Figure 3. Phags-pa-like english CAPTCHA ("LtWYUB8QxZ").

letters are overlapping. If the attempt to segment the CAPTCHA word is successful, the final step is for the attack script to properly recognize each letter. A robust design would make the recognition task impossible. Examples of antirecognition techniques are slanting letters, adopting various font types with different sizes or using special designs such as Korean Hangul-like English characters. An attack bot would have difficulty recognizing letters in Figure 3 (i.e., LtWYUB8QxZ) designed to mimic the characters.

These three common steps in a bot attack imply that the two major preventive techniques for designing CAPTCHA should be antisegmentation and antirecognition technology (Hernandez-Castro & Ribagorda, 2009; Yan & Ahmad, 2008). Antisegmentation techniques will ensure that the attack bot is incapable of separating one letter from others while antirecognition devices will prevent it from correctly identifying the letters in a CAPTCHA. Connecting the letters that form a CAPTCHA is an example of the former technique and arranging all letters in a wave style is one of the latter (Burstein et al., 2011; Yan & Ahmad, 2008).

The ideal way to assess CAPTCHA robustness is the effectiveness of its defense in an actual attack to see if the CAPTCHA is broken. The most robust design must be one that survives such an attack. Nevertheless, automated scripts which can attack CAPTCHA are not publically available. Nonetheless, there is a highly recommended software in the market for testing CAPTCHA robustness called optical character reader (OCR) software (Hsu &Lee, 2011; Singh, Bacchuwar, & Bhasin, 2012). The OCR will segment and identify all letters in the CAPTCHA which enables researchers to analyze whether the CAPTCHA may survive the segmenting and the recognition tasks (Bentley & Mallows, 2006; Ferzli, Bazzi, & Karam, 2006; Singh et al., 2012).

Three design issues of text-based CAPTCHA are of particular interest: text rotation, string length, and letter formats. Text rotation refers to a condition where letters in a CAPTCHA are rotated or tilted to a certain degree (Bursztein et al., 2011; Soumya, Abraham, & Sawati, 2014). The importance of CAPTCHA's rotated design is detailed in Burstein and colleagues (2011). When a letter is rotated, it imposes a certain difficulty of recognition on the attack script (Bursztein et al., 2014; Soumya et al., 2014). In addition, Sakkatos, Theerayut, Nuttapol, & Surapong (2014) remarked that automated spam software may not correctly detect a skewed text. Although text rotation is cited as having critical implications for evaluating CAPTCHA robustness, no empirical studies are available to verify such claims. In fact, the study by Sakkatos and colleagues (2014) who did a simulation, did confirm that skew in the range of 10–15 degrees enhanced CAPTCHA robustness. Chandovale and colleagues (2009) who developed a specific application to break a CAPTCHA, remarked on the practicality of text rotation in improving CAPTCHA robustness. In the current study, a typical letter stands at a zero degree. If tilted

Figure 4. Examples of −45 and +45 degree rotation of a CAPTCHA letter.

counterclockwise, the rotation degree is positive while clockwise tilt is negative. See Figure 4 for examples of −45 and +45 degree rotations.

String length of text-based CAPTCHA refers to the number of characters in one CAPTCHA test. According to Bursztein and colleagues (2011), the longer the string, the less likely the CAPTCHA can be successfully attacked and Starostenko, Cruz-Perez, Uceda-Ponga, and Alarcon-Aquino (2015) also note that CAPTCHA with long strings may survive an attack better than short ones. In other words, the longer the string of letters, the more robust the CAPTCHA. Moreover, a long CAPTCHA while being more secure than short ones is still able to be interpreted by actual human viewers (Gao et al., 2014). Other design variables may provide CAPTCHA robustness but users may have difficulty deciphering them. As such, a rather longer letter string per CAPTCHA will lead to an optimal design since it is both secure and user friendly. Despite several critical remarks (Gao et al., 2014; Soumya et al., 2014; Starostenko et al., 2015) on string length of CAPTCHA, empirical studies verifying such criticism are extremely rare. There is, therefore, a serious need for research on the issue.

The final design variable of CAPTCHA is letter format. In this study, format refers to whether the letters in a CAPTCHA are typed or in script (handwritten) style. Gao and colleagues (2014) remarked that script format letters often connect to each other making this design so robust that an attack is less likely than with typed ones. Not only does the script format enhance antisegmentation strength, it also imposes a remarkable barrier if the attack software tries to identify each letter in a CAPTCHA (Starostenko et al., 2015). However, Starostenko and colleague's (2015) experiment did not compare robustness between typed and script CAPTCHA letter formats. Indeed, Soumya and colleague (2014) mention that handwritten letters (which are visually similar to script style) are considered an effective technique to protect CAPTCHA from bot attack. While there is a fair amount of research on potential impacts of letter formats on CAPTCHA robustness, none are empirical examinations to ascertain impacts of this design feature.

Research objectives

A review of the literature on text-based CAPTCHA robustness indicates two gaps. First, the importance of CAPTCHA's antisegmentation and antirecognition attributes has been greatly noted (Bentley & Mallows, 2006; Bursztein et al., 2011; Ling-Zi & Yi-Chun, 2012; Yan & Ahmad, 2008). However, only a few projects have

empirically verified the effects of specific antisegmentation or antirecognition techniques on the robustness of CAPTCHA technology. Moreover, empirical research on antirecognition techniques is relatively rare as compared to that on antisegmentation techniques (Bursztein et al., 2011; Hernandez-Castro & Ribagorda, 2010). Second, research discussing or verifying the effects of antirecognition techniques does exist, however, studies that evaluate all three variables of antirecognition technology is scant.

As a result, the current study's main objective is to analyze the extent to which text rotation, string length, and the letter format affect the robustness of text-based CAPTCHA.

Methodology

There are four methodological issues to be discussed: research design, variable operationalization, experiment execution, and data analysis and hypothesis testing.

Research design

To test the effects of text rotation, string length, and letter formats on the robustness of text-based CAPTCHA, we adopted an experimental approach. We intended to manipulate all three independent variables and observe their effect on CAPTCHA robustness. The experiment is designed in such a way that we could have sufficient control on the manipulation of variables and observation of effects. If there is an observed effect, it is considered reliable and valid (Babbie, 2013).

Variable operationalization

In the current study, there are three independent variables and one dependent variable. The first independent variable is text rotation which has two possible values: +45 and −45 degrees of rotation or tilt. The reasoning behind the application of text rotation in these two values is based on a limited amount of prior research and to some extent, fairly intuitive (Sakkatos et al., 2014). That is, past research on CAPTCHA robustness and text rotation is extremely rare. Thus, we had no empirical grounds on which we could have operationalized text rotation. Among the very few studies on this variable, Chellapilla, Larson, Simard, and Czerwinski (2005) suggested that the greatest degree of rotation on text-based CAPTCHA that humans are still able to read is between −45° and +45°. We then wanted to test if there was a statistically significant difference in robustness when all letters in a CAPTCHA are tilted at −45° and +45°. Figure 3 shows how the −45 and +45 degree rotations are manipulated in our experiment.

The second independent variable is string length, referred to in the current study as the number of characters in one CAPTCHA test. Yan and Ahmad (2008) commented that a CAPTCHA with a long string length should be more robust than those with short strings. This is perhaps due to the higher chance that the attack may fail

to properly read through the long list of letters of the string. Thus, the more letters in one CAPTCHA, the more difficulty the attack bot has to interpret the CAPTCHA (Gao et al., 2014). Moreover, Rice, Kanai, and Nartker (1993) confirm that the more letters printed on paper texts, the less accurate the OCR reader. Despite the fact that Rice and colleagues did not focus on CAPTCHA technology, their finding supports the idea that a CAPTCHA with many letters may survive a bot attack better and is more robust than one with fewer letters.

Once we decided to include string length as an independent variable, the next question was how to apply it. Previous work addressing both string length and CAPTCHA robustness is extremely rare. Since effectively designed CAPTCHA must not only be robust but should be easy for humans to read, it was decided the current study's operationalization of string length be based on empirical insight into human's recognition of English characters and its use for text-based CAPTCHA (Gao et al., 2014; Miller, 1955; Thongkamwitoon, Asdornwised, Aramvith, & Jitapunkul, 2002; Wimmer & Goswani, 1994). Gao and colleagues (2014) reported (a) poor robustness of 22% on Yahoo's text-based CAPTCHA that had six to eight characters in one CAPTCHA test and (b) fair robustness of 66% on Baidu's with the average of four characters per test. Based on the findings in Gao and colleagues (2014), we had made three decisions. First, the selected values of the string length must be lower than six or higher than eight characters. They should not fall in the range of six to eight for the low amount of robustness reported in Gao and colleagues (2014). Second, the low value of the string length was four characters. This is a result of the improved robustness on the Baidu's CAPTCHA which has in average four characters. The selection of four characters is also in line with the word recognition task in Wimmer and Goswami (1994) where the array of four characters were most recognized. Finally, we chose the 10 characters for the high value of the string length. The selection was fairly arbitrary. Gao and colleagues (2014) implied the number of characters per CAPTCHA be higher than eight in order to gain the acceptable level of robustness. We were however unable to locate previous literature suggesting what the value should be. As such, the 10 characters per CAPTCHA was selected and it is much intuitive. In conclusion, the two values of string length in the current study are four and 10 letters per CAPTCHA test.

The final independent variable is CAPTCHA letter format. The two formats examined in the current study are typed and script (like handwriting). Following Gao and colleagues (2014), who discovered that CAPTCHA's overlapping letters have the highest survival rate in attacks as compared to other types of distortion, it was deemed the CAPTCHA should adopt the script format for the current study since it may be more resistant to a bot attack than the typed style. A number of information system security researchers also agree on the importance of letter format on CAPTCHA design to enhance robustness (Soumya et al., 2014; Starostenko et al., 2015; Yan & Ahmad, 2008).

We assessed CAPTCHA robustness in this study using an OCR. For purposes of the experimental approach, we created a set of CAPTCHA tests based on the three design variables discussed and tested to see if the differences in robustness are

statistically significant across the variables. In this study, a CAPTCHA test is robust if the OCR incorrectly read all of its letters. As such, the robustness criterion adopted is whether a given CAPTCHA was correctly read by the OCR and measured as the proportion of the CAPTCHA incorrectly read by the OCR to that of a CAPTCHA which could be completely read by the same OCR.

After a survey of OCRs presently available in the market and a review of recommendations in academic journals and trade magazines (Ahmad, Yan, & Marshall, 2010; Chellapilla et al., 2005; Hsu & Lee, 2011; Singh et al., 2012), it was decided to use the Omni (Professional 18) OCR to assess the robustness of our designed CAPTCHA tests.

Experiment execution

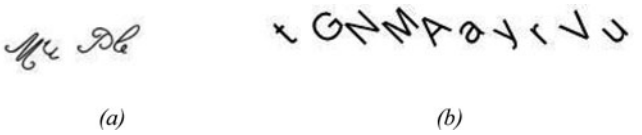
We developed 1,464 text-based CAPTCHAs based on eight conditions of the three design variables (2 rotations × 2 lengths × 2 formats). Table 1 shows the number of CAPTCHA tests thus created in each condition. Initially, we intended to have at least 180 in each condition to comply with statistical suggestions on the optimal number of samples in an experiment (Roscoe, 1975). However, following the recommendation of one of the research assistants, it was decided that all CAPTCHAS designed for data collection would be included for study.

Following the eight conditions, the design randomly used upper and lower cases. Neither numbers (i.e., 0, 1, 2 to 9) nor special symbols (e.g., ∗, %, $, or;) were used in our CAPTCHA design. This follows from reports noting user confusion between certain numbers or special symbols and the text (Yan & Ahmad, 2008). For example, 5 and s, 2 and z or 1, % and 7 can be misleading. Although our experiment did not involve human users, we want the validity of our findings to be robust for actual application of CAPTCHA.

Arial typeface was used for the typed condition and cursive standard typeface for the script condition. The selection of the former is due to Arial font being highly legible and often used (Bernard, Chaparro, Mills, & Halcomb, 2003) and the selection of the latter was based on lack of format recommendations in the literature and our own survey with fifty participants between 20–29 years of age who had used a number of online services and admitted to having seen CAPTCHA. Also, Cursive Standard font was voted most popular among the three most cited script fonts

Table 1. The number of CAPTCHA tests created based on eight conditions of three factors.

Condition	Text Rotation	String Length	Letter Format	Number of CAPTCHA Tests	Rate of Robustness
1	−45°	4	Typed	182	.154
2	−45°	4	Script	182	.473
3	−45°	10	Typed	198	.422
4	−45°	10	Script	180	.450
5	+45°	4	Typed	180	.406
6	+45°	4	Script	180	.439
7	+45°	10	Typed	180	.517
8	+45°	10	Script	182	.478
			Total	1,464	.397



(a) (b)

Figure 5. Examples of two CAPTCHA tests used in the experiment.

in the dafont.com website. Figure 5 shows (a) an example of a -45° tilted script CAPTCHA with four letters, and (b) an example of a $+45^{\circ}$ tilted typed CAPTCHA with 10 letters.

The process of creating a CAPTCHA for each of the eight experimental conditions had three steps. First, we randomly selected the letters based on the format and length conditions. Then, we placed them together as a word. For instance, we selected w, V, c, and X in the script format and we then formed a wVcX CAPTCHA word. Second, we checked if the selected word had a meaning in a dictionary. Only if it had meaning would we randomly change the order of the selected letters so that it had no meaning. Our attempt to use CAPTCHA without meaning results from certain bot attacks recognizing CAPTCHA from their meaning. Finally, we slanted each selected letter, either $-45°$ or $+45°$ and saved it as an image file for use in the experiment.

To evaluate the robustness of our CAPTCHA designs, we used the OmniPage OCR purchased directly from the vender's website. This selection was based on a review of previous literature (Ahmad et al., 2010). After configuring the program, we tested the process of reading a few of our CAPTCHA tests selected from each of the eight conditions. The CAPTCHA tests used in the test of the OmniPage OCR configuration were not used in the actual experiment. The outcomes confirmed the reliability and validity of our tests after which we started the reading process on 1,464 pieces of CAPTCHA. To ensure validity of the collected data, the experiment used a desktop computer using 1 GHZ Intel Pentium multicore processor with the WINDOWS XP operating system and 1 GB of RAM.

Analysis and hypothesis testing

In addition to the report of descriptive statistics, we employed an analysis of variance (ANOVA) to test if the main effects of the three factors as well as all possible interaction effects on the robustness rate were statistically significant (Babbie, 2013; Roscoe, 1975).

Results

Five hundred eighty one of the 1,464 CAPTCHA tests were incorrectly read by the OmniPage OCR, yielding a robustness rate of 0.397. Table 2 shows the rates across variable conditions, text rotations, string lengths, and letter formats. Table 3 displays the output of the ANOVA tests.

Table 2. Robustness rates grouped by three variables.

Variables	Number of CAPTCHAs Incorrectly Read by OCR	Number of CAPTCHAs Correctly Read by OCR	Rate of Robustness
Rotation (degree)			
−45	249	742	0.336
+45	332	722	0.460
Length (Number of Letters)			
4	266	724	0.367
10	315	740	0.426
Format			
Typed	248	740	0.335
Script	333	724	0.459
Total	581	1464	0.397

As seen in Table 2, the differences in robustness between two categories of the three factors appear small. For instance, the robustness rates of the CAPTCHA with texts at −45° and +45° tilt are 0.336 and 0.460, respectively. Nonetheless, the results of the ANOVA in Table 3 confirm that (a) the effects of text rotation, string length, or letter format on robustness rates of CAPTCHA tests are all statistically significant (significance level .000), and (b) the interaction effect of text rotation and letter format and that of string length and letter format (significance levels .000, and .033, respectively) are both statistically significant. Yet, other interaction effects were not significant.

In addition, it appears in Table 3 that the letter format may have an effect on the correlation between the text rotation and the rate of robustness of CAPTCHA or on the correlation between the string length and the robustness rate. Further exploration of Table 4 reveals the moderating effect of the letter format. In other words, only when the CAPTCHA tests use the typed letter format are the CAPTCHA robustness rate between two angles of text rotation (significance level .000) and that between two sets of string length (significance level .001) statistically significant. Nevertheless, when the letters in the CAPTCHA tests were of the script format, the rates between these two variables were not significant (significance levels .823 and .941, respectively).

Table 3. ANOVA results.

Source of Variance (SOV)	Degree of Freedom (Df)	Sum Square of Error (SSE)	Mean Square Error (MSE)	F Statistics	Significance Level
Text rotation	1	5.488	5.488	24.143	.000
String length	1	1.390	1.390	6.114	.000
Letter format	1	5.502	5.502	24.203	.000
Text rotation × String length	1	0.066	0.066	0.292	.589
Text rotation × Letter format	1	5.742	5.742	25.263	.000
String length × Letter format	1	1.041	1.041	4.578	.033
Text rotation × String length × Letter format	1	0.110	0.110	0.485	.486
Error	1456	330.692	0.227		
Total	1463	350.031			

Table 4. Robustness rates across text rotation and string length while letter format was controlled.

Letter Format			Significance Level
Typed			
Text rotation (angle)	−45°	+45°	
	.216	.461	.000
String Length (number of letters)	4	10	
	.279	.389	.000
Script			
Text rotation (angle)	−45°	+45°	
	.461	.459	.823
String Length (number of letters)	4	10	
	.456	.464	.941

Conclusion and discussion

Using an experimental procedure, we tested whether text rotation, string length, and letter format affects text-based CAPTCHA robustness. The results confirm that all three design factors have statistically significant effects on the robustness of CAPTCHA tests. As a whole, 581 out of 1,464 CAPTCHA tests (or 39.7%) were robust since the OCR failed to read them correctly. Roughly 40% of our tests were found to be robust whereas the OCR was able to successfully read the remaining tests. We were unable to locate previous research in which CAPTCHA robustness was reported. We then reviewed other publications examining similar robustness rates such as the success rate of recognition or the overall success rate of CAPTCHA attacks (Bursztein et al., 2011; Yan & Ahmad, 2008). Although the 40% success rate found in the current experiment may seem low, it is in line with similar research findings (Burstein & Matthien, 2011; Starostenko et al., 2015). Using a collection of CAPTCHA tests from a Google search, Yan and Ahmad (2008) found a successful recognition rate of 46% while Bursztein and Matthien (2011) reported a success rate (robustness) as low as 18%. Experimenting with CAPTCHA tests in the reCAPTCHA projects, Starostenko and colleagues (2015) reported that the overall success rate of the tests was 26%. In conclusion, the overall robustness of the CAPTCHA tests found in the current study is consistent with similar findings in previous studies, despite being somewhat low.

The robustness between the -45°-slanted and $+45^{\circ}$-slanted CAPTCHA tests are statistically different. Based on our experiment findings (Table 2), CAPTCHA using a positive angle (i.e., counterclockwise, $+45^{\circ}$) is thus rendered more robust as compared to the negative angle (clockwise, -45°). It was found that only Sakkatos and colleagues (2014) examined the effects of text rotation on CAPTCHA robustness. They reported that a rotation of 15° or greater will enhance CAPTCHA robustness. However, Sakkatos and colleagues (2014) do not indicate which direction of rotation their CAPTCHA was oriented (counterclockwise or clockwise). As a result, there is a need to examine the optimal degree of text rotation in order to maximize CAPTCHA robustness.

With respect to string length, 4-lettered and 10-lettered CAPTCHA tests were found to have statistically different rates of robustness. As expected, the string of 10 letters yields more robust CAPTCHA tests than does that of 4 letters (Bursztein

et al., 2011). Gao and colleagues (2014, p. 349) claim that the more letters, the more robust. However, their conclusion was not based on a controlled experiment. Our finding on the effect of string length is, therefore, a valid empirical confirmation of its effect on CAPTCHA robustness.

Regarding differences in robustness between typed and script letters in CAPTCHA, rates varied significantly. Table 2 indicates that CAPTCHA using script letters were more robust than those using typed. Although there is no empirical study examining letter format effects on CAPTCHA performance, our finding is not surprising. Gao and colleagues (2014) recommend security practitioners use written letter format (script) to design CAPTCHA connecting all letters together to make it difficult to segment or to recognize letters. Moreover, Starostenko and colleagues (2015) support that recognizing the written format letters requires more resources than typed ones. Nonetheless, their support was not an attempt to compare the degree of robustness between typed and script CAPTCHA. To verify the effects of letter format (typed, script, or others), we recommend further empirical research be done on this aspect of CAPTCHA design.

The data in Table 4 helps explain the significant interaction effects of text rotation and letter format as well as that of string length and letter format on CAPTCHA robustness. If a typed format is used, the effects of text rotation or string length are both statistically significant but the same effects became trivial when script letter format was used. From a security standpoint, script format is, therefore, recommended since regardless of the rotation angle or the string length, robustness is almost the same. On the contrary, should the typed format be used, the proper tilt angle of letters must be +45°or the CAPTCHA must have a ten letter string length in order to be relatively robust.

The results of the current study yield both theoretical and practical contributions to the study of CAPTCHA robustness. We are able to offer two unique theoretical contributions. First, our experiment has proved the significant effects of text rotation, string length, and letter format on CAPTCHA robustness. In particular, our analysis appears to indicate that text-based CAPTCHA with 10-typed letters clockwise-rotated at 45 degrees is most robust as compared to the other configurations (Table 1). This finding suggests empirical grounds on which future research into security issues could be pursued. Second, our analysis shows the significant interaction effect of letter format and text rotation in addition to the interaction between letter format and string length on robustness (Table 3). An exploration of these interaction effects in Table 4 would indicate the strong effects of script letter format on robustness.

Our study also offers three practical contributions. First, the overall rate of robustness in our experiment was 39.7% which means only four in 10 CAPTCHA tests passed the robustness assessment using the OCR. While in line with other studies, this low rate of success may alarm security practitioners who may conclude that adopting three design features (text rotation, string length, or letter format) may not be sufficient to create sufficiently robust text-based CAPTCHA. Since all

three factors are relevant to antirecognition techniques, practitioners must incorporate antisegmentation techniques into CAPTCHA design. Second, considering the three design issues of the antirecognition techniques studied, our experiment suggests that practitioners should be aware of the significant effect of the three issues on CAPTCHA robustness. Should a CAPTCHA designer want to rotate the text for instance, the recommended angle should be 45° clockwise. Finally, the significant interaction effect involving letter format suggests a practical value. As seen in Table 4, CAPTCHAs using script letter format result in robustness across categories of the other design variables being about the same. However, if typed letter format is used, the difference becomes statistically significant. As a result, the script format is more robust than the typed and strongly recommended for the practical design of text-based CAPTCHA.

As with other research, our study has two major limitations. The first one is methodological. Given the experimental approach, the scope of the current study is limited. It, therefore, constrains the finding's generalizability. Although internal validity was achieved by full control in the experiment, external validity may be compromised. Only future research on similar topics will enhance external validity. The second limitation is the dynamic context of the study. The online world is always changing. What researchers can properly claim today may not reflect accurately the state of technological circumstances tomorrow. This suggests the need for continual empirical examination in the field.

Funding

We are thankful for financial support, in part, from "Chulalongkorn Academic Advancement Into Its Second Century" Project.

References

Ahn, L.V., Maurer, B., McMillen, C., Abraham, D., & Blum, M. (2008). reCAPTCHA: Human-based character. *Science*, *321*(5895), 1465–1468.

Ahn, Y., Kim, N., & Kim, Y. S. (2013). *A user-friendly image-text fusion CAPTCHA for secure web services*. Proceedings of iiWAS2013 Conference, Vienna, Austria, pp. 550–554.

Ahmad, A. S., Yan, J., & Marshall, L. (2010). The robustness of a new CAPTCHA. *Proceedings of the 2010 EUROSEC Conference*, Paris, France, pp. 36–41.

Babbie, E. (2013). *The practice of social research* (13th ed.). Belmont, CA: Wadsworth: Cengage Learning.

Bentley, J., & Mallows, C. (2006). *CAPTCHA challenge strings: Problems and improvements*. Proceedings of the SPIE Conference, Document Recognition and Retrieval XIII3, January 15, pp. 1–7.

Bernard, M. L., Chaparro, B. S., Mills, M. M., & Halcomb, C. G. (2003). Comparing the effects of text size and format on the readability of computer-displayed Times New Roman and Arial text. *International Journal of Human-Computer Studies*, *59*, 823–835.

Bursztein, E., Bethard, S., Fabry, C., Mitchell, J. C., & Jurfsky, D. (2010). *How good are humans at solving CAPTCHAs? A large scale evaluati on*. 2010 IEEE Symposium on Security and Privacy, May 16–19, Berkeley, CA, pp. 399–413.

Bursztein, E., Martin, M., & Mitchell, J. C. (2011). Text-based CAPTCHA strengths and weaknesses. In Proceedings of the 18th ACM conference on Computer and Communications Security (CCS '11). ACM, Chicago, IL, pp. 125–138.

Bursztein, E.; Moscicki, A.; Fabry, C.; Bethard, S. Mitchell, J. C. & Jurfsky, D. (2014). *Easy does it: More usable CAPTCHAs*. Proceeding of the SIGCHI Conference on Human Factors in Computing Systems, Toronto, Ontario, Canada, pp. 2637–2646.

Chandavale, A. A., Sapkal, A. M. & Jalnekar, R. M. (2009). *Algorithm to Break Visual CAPTCHA*. In Proceedings of the 2009 Second International Conference on Emerging Trends in Engineering & Technology in Nagpur, December 16–18, pp. 258–262.

Chellapilla, K., Larson, K., Simard, P., & Czerwinski, M. (2005). *Designing human friendly human interaction proofs* (HIPs). Conference on Human factors in computing systems (CHI), Portland, Oregon, pp. 711–720.

Ferzli, R.; Bazzi, R. & Karam, L. J. (2006). *A CAPTCHA based on the human visual systems masking characteristics*. Proceedings of the 2006 ICMA Conference, July 9–12, Toronto, Ontario, Canada, pp. 517–520.

Gao, H., Wang, W., Fan, Y., Qi, J., & Liu, X. (2014). The robustness of "connecting characters together" CAPTCHAs. *Journal of Information Science and Engineering*, *30*, 347–369.

Hernandez-Castro, C. J., & Ribagorda, A. (2010). Pitfalls in CAPTCHA design and implementation: The math CAPTCHA, a case study. *Computers & Security*, *29*, 141–157.

Hsu, C. H., & Lee, Y. L. (2011). J. C. Jacko (Ed.), Effects of age groups and distortion types on text-based CAPTCHA tasks. *Human-Computer Interaction, Part IV*, 453–455.

Lin, R., Huang, S.-Y., Bell, D. B., & Lee, Y.-K. (2011). A new CAPTCHA interface design for mobile devices. *Proceedings of the 12th Australasian User Interface Conference*, *117*, 3–7.

Ling-Zi, X. & Yi-Chun, Z. (2012). *A case study of text-based CAPTCHA attacks*. Proceedings of the 2012 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery, October 10–12, Sanya, China, pp. 121–124.

Miller, G. A. (1955). The magical number seven, plus or minus two some limits on our capacity for processing information. *Psychological Review*, *101*(2), 343–352.

Rice, S.V., Kanai, J., & Nartker, T.A. (1993). *An evaluation of OCR accuracy. Annual research report*. Las Vegas, NV: Information Science Research Institute.

Roscoe, J. T. (1975). *Fundamental research statistics for the behavioral sciences* (2nd ed.). New York, NY: Holt, Rineheart and Winston.

Sakkatos, P., Theerayut, W., Nuttapol, V., & Surapong, P. (2014). *Analysis of text-based CAPTCHA images using template matching correlation techniques*. Proceedings of the 4th Joint International Conference on Information and Communication Technology, Electronic and Electrical Engineering (JICTEE-2014), March 5–8, Chiang Rai, Thailand, pp. 1–5.

Singh, A., Bacchuwar, K., & Bhasin, A. (2012). A survey of OCR applications. *International Journal of Machine Learning and Computing*, *2*(3), 314–318.

Soumya, K. R., Abraham, R. M., & Sawati, K. V. (2014). A survey of difference CAPTCHA techniques. *International Journal of Advances in Computer Science and Technology*, *3*(2), 117–122.

Starostenko, O., Cruz-Perez, C., Uceda-Ponga, F., & Alarcon-Aquino, V. (2015). Breaking text-based CAPTCHAs with variable word and character orientation. *Pattern Recognition*, *48*, 1101–1112.

Thongkamwitoon, T., Asdornwised, W., Aramvith, S., & Jitapunkul, S. (2002). On-line Thai-English handwritten character recognition using distinctive features. *Proceedings of the 2002 Asia-Pacific Conference on Circuits and Systems (APCCAS)*, *2*, 259–264.

Wimmer, H., & Goswami, U. (1994). The influence of orthographic consistency on reading development: Word recognition in English and German children. *Cognition*, *51*, 91–103.

Yan, J., & Ahmad, A. S. (2008). *A low-cost attack on a Microsoft CAPTCHA*. Proceedings of CCS Conference, Alexandria, VA, pp. 543–554.