Association for Information Systems

AIS Electronic Library (AISeL)

AMCIS 2020 Proceedings

Information Security and Privacy (SIGSEC)

Aug 10th, 12:00 AM

Evaluating the Usability of a Multilingual Passphrase Policy

Pardon Blessings Maoneke
University of Mpumalanga, blessings.maoneke@ump.ac.za

Stephen V. Flowerday
Rhodes University, s.flowerday@ru.ac.za

Merrill Warkentin
Mississippi State University, m.warkentin@msstate.edu

Evaluating the Usability of a Multilingual Passphrase Policy

Completed Research Full Paper

Pardon Blessings Maoneke
University of Mpumalanga
blessings.maoneke@ump.ac.za

Stephen Flowerday
Rhodes University
s.flowerday@ru.ac.za

Merrill Warkentin
Mississippi State University
m.warkentin@msstate.edu

Abstract

The literature shows that users struggle to generate secure passwords. This has led to systems administrators implementing password expiry policies that burden and frustrate users. This study explores the security and usability of a multilingual passphrase policy, as multilingualism has the potential to enhance security. A total of 224 participants were invited to participate in an experiment to generate and recall short passwords and multilingual passphrases. The findings of this study show that, although a multilingual passphrase policy made passphrase generation slightly more difficult, its use motivated users to generate unique memorable passphrases. Arguably, repeated use of passphrases promotes memorability and cognitive fluency. Furthermore, the multilingual passphrases in this study proved to be stronger than those reported in the literature.

Keywords

Authentication, passphrases, multilingual passphrases, password usability, passwords, password policy.

Introduction

It has been shown that short password policies are burdensome, as users spend one to two business days a year generating passwords (Choong, Theofanos, & Lui, 2014). These new passwords are mainly generated to replace expired and forgotten passwords (Becker, Parkin, & Sasse, 2018). Forgetting short passwords is more pronounced when users are required to generate complex passwords (Becker et al., 2018; Choong et al., 2014). Consequently, passphrases have been identified, with caution, as a solution to the security and usability predicament of short passwords (Komanduri, 2016; Shay et al., 2016). Users often base their passphrases on popular words and grammatical rules in a language, something that translates to weak passphrases (Bonneau & Shutova, 2012; Shay et al., 2016). The United States' National Institute of Standards and Technology (NIST) and many research articles in the literature are advocating for the use of a blacklist to restrict the use of common words in passphrases (Grassi, Garcia, & Fenton, 2017; Shay et al., 2016). Although useful, a blacklist may have a negative impact on usability and users may circumvent this security control by applying simple modifications or "workarounds" (Alter, 2014) such as inserting digits or using capital letters in blacklisted passwords (Habib et al., 2017; Shay et al., 2016). Furthermore, passphrases from a single language may share the same letter distribution, which could be leveraged by probabilistic password guessing algorithms (Wang, Cheng, Gu, & Wang, 2015; Weir, Aggarwal, Collins, & Stern, 2010). This study assessed the efficacy of a technique grounded in intrinsic difference in the character distribution of passwords that are generated using different languages (Maoneke, Flowerday & Isabirye, 2018; Wang et al., 2015) and proposes the use of a multilingual passphrase policy. Such policies have the potential to mitigate the threat posed by the dominance of words from a single language, thereby promoting randomness (Rao, Jha, & Kini, 2013) in character distribution. The use of multilingual passphrases is feasible, especially in Africa where the first spoken language is often different from the first written language or language of instruction (Deumert & Lexander, 2013). It can also be feasible throughout Europe and India, where large percentages of the populations are also multilingual. Authors of this study argued that if users can generate strong multilingual passphrases, systems administrators will not need to enforce regular password expiry policies. This has the potential to enhance memorability and reduce the burden imposed by having to change passwords or passphrases frequently. Hence, this study investigates the security and usability of a multilingual passphrase policy.

Literature review

Human memory functionality

Password use requires users to generate, learn, retain, and recall passwords (Woods & Siponen, 2018, 2019), activities which rely on the memory, which is responsible for storing and retrieving information acquired through the different senses. Atkinson and Shiffrin (1968) proposed a stages of memory theory that splits the memory into the sensory, short-term, and long-term memory. This theory suggests that passwords find their way into the short-term memory through the sensory memory. The short-term memory is limited in capacity to 7 ± 2 chunks of information and can hold new information for just 30 seconds before one starts to forget, a process known as memory decay (Miller, 1956). The longer a password can stay in the short-term memory, the more likely it is to be transferred to the long-term memory, where permanent and memorable information resides waiting to be retrieved (Atkinson & Shiffrin, 1968). However, complex passwords with more than seven characters that appear to be random to a user may not last for more than 30 seconds in the short-term memory. This possibly explains why users struggle to memorize system-assigned passwords (Shay et al., 2012; Warkentin, Davis, & Bekkering, 2004). Interference and a lack of cues between new information and the information that resides in the long-term memory is one of the reasons why users forget complex passwords (Cowan, 2014; Zhang, Luo, Akkaladevi, & Ziegelmayer, 2009). Password rehearsal or repeated use of the newly generated passwords is one of the ways of prolonging the time passwords are retained in the short-term memory in the hope that they will eventually be transferred to the long-term memory (Woods & Siponen, 2019), and, in some cases, behavioral automaticity or habit can form, which involves implicit memory (Vedadi & Warkentin, 2018). In addition, generating new passwords based on information in the long-term memory may reduce interference and improve memorability. This is consistent with chunking theory which postulates that meaningful information in the long-term memory can be grouped together into new information that is easy to memorize. Thus, "highly meaningful words are easier for a person to learn and remember than less meaningful words, with meaningful being defined by the person's number of associations with the word" (Newell, Shaw, & Simon, 1961 in Carstens, Malone, & McCauley-Bell, 2006, p. 100). Hence, passphrases can be seen as chunks of information that have been grouped together to formulate a single meaningful and memorable phrase (Woods & Siponen, 2019).

Password policies: Security and usability

There are different password generation policies in the public domain (Wang et al., 2015; Weir et al., 2010). This study focuses on the password composition policy. This is a widely used password policy that arguably emanated from the old version of the NIST's Special Publication (SP) 800-63 Electronic Authentication Guideline. Also known as an "explicit password creation policy," this policy defines that which constitutes an acceptable password (Weir et al. 2010, p. 171) in terms of the minimum length, use of symbols, numbers, and uppercase and lowercase letters. The policy works best with proactive check mechanisms to monitor the use of different character classes and length (Shay et al., 2012) or a dictionary check and a blacklist to prevent users from adopting common passwords (Grassi et al., 2017). The short password policy is an example of a password composition policy. Short passwords are often based on 8-characters and have different character classes to enhance security e.g. "P@55word" (Wang et al., 2015; Weir et al., 2010). In addition, the password composition policy can be adjusted to accommodate the generation of passphrases where emphasis is on increasing length by using several words within each passphrase in order to enhance security and usability. Passphrases can have a minimum length of 12 characters and the use of different character classes is not always mandatory (Komanduri, 2016; Melicher et al., 2016; Shay et al., 2016).

The short password policy is associated with a number of security and usability concerns. When this policy is implemented restrictively, users find password generation and learning frustrating, and often fulfil password requirements in a predictable manner (Becker et al., 2018; Komanduri et al., 2011). For example, users find it difficult to meet the character class requirement and sometimes the required password length. This was demonstrated by users who required more password generation attempts (2 to 3.35 attempts) to generate a password that is "at least 8 characters" long with a "lowercase English letter, uppercase English letter, digit and symbol" than when generating a passphrase of at least 12 characters in length (Komanduri et al., 2011; Shay et al., 2016, p. 12). Furthermore, it has been noted that users are more likely to make minor changes to existing passwords, to use existing passwords as is, or to use old passwords when they find the short password policy complex and burdensome (Choong et al., 2014).

In addition, users are more likely to adapt semantic information when faced with a restrictive short password policy (Wei, Golla, & Ur, 2018). According to Tam, Glassman and Vandenwauver (2010), users may adopt poor password generation techniques in order to maximize convenience. Semantic information that is often adapted during short password generation includes keyboard patterns, dictionary words, website information or personal information such as hobbies, personal names or names of loved ones, date of birth, address, phone numbers and identification numbers (AlSabah, Oligeri, & Riley, 2018; Wang et al., 2015; Li, Han, & Xu, 2014; Wei et al., 2018). The passwords that are generated can be easily guessed by probabilistic guessing algorithms (Weir et al., 2016). By contrast, in passphrase policies, users may base their passphrases on common words in a language and predictable grammatical rules (Bonneau & Shutova, 2012; Rao et al., 2013; Shay et al., 2016).

Another usability challenge faced by users is the failure to memorize passwords (Becker et al., 2018). The memorability of a password is the most important factor users consider when generating a password (Choong et al., 2014, Ur et l., 2015). Users often opt to memorize a password if they "perceive [that] the benefit" of memorizing the password outweighs the cost of writing it down in the event of "a security breach" (Duggan, Johnson, & Grawemeyer, 2012, p. 416). Where users find it difficult to memorize passwords, they resort to writing them down, reusing them and using password managers (Choong et al., 2014; Komanduri et al., 2011; Shay et al., 2016; Stobert & Biddle, 2014). The fact that users resort to writing down passwords and storing them in insecure locations or saving them in unencrypted devices or web browsers (12% to 81%), instead of using formal password managers, is worrying (Choong et al., 2014; Stobert & Biddle, 2014).

Lastly, users often face authentication challenges when logging in using short passwords that are generated following different character classes. For example, mistyping or typographical errors when logging in (Choong et al., 2014; Keith, Shao, & Steinbart, 2009; Melicher et al., 2016). Typographical errors occur when a user mistakenly strikes a nearby key or transposes the correct characters or "enters too few or too many keystrokes" (Keith et al., 2009, p. 68). These errors are not a result of memory loss, but of an error that occurs during the execution stage as one is keying in a password (Keith et al., 2009). Nevertheless, typographical errors contribute to user frustration during logging in.

Methodology

Experiments and password corpora in the public domain are often used as sources of passwords for various analyses (AlSabah et al., 2018; Melicher et al., 2016; Shay et al., 2016; Weir et al., 2010; Woods & Siponen, 2019). To the knowledge of the authors, there is no password corpora in the public domain that could be used to meet the research aims of this study. Hence, an experiment was designed to gather data for evaluating the usability of a multilingual passphrase policy. An ethics certificate was secured prior to the commencement of the experiment. The next section explains the experimental design, showing measures that were taken to enhance the ecological validity of passwords for this study (Shay et al., 2016). This is followed by an explanation of the techniques that were used to evaluate the usability of passphrases.

The experimental design and administration

This study used a web-based experiment that was built specifically for this study, the design of which was guided by Shay et al. (2016) and Komanduri (2016). On opening the password generation platform, participants were presented with a scenario encouraging the generation of a realistic short password or a multilingual passphrase. The experiment required every participant to generate and confirm a short password or multilingual passphrase, then to complete the first questionnaire on password/passphrase generation usability and re-enter the password/passphrase to test password recall after a cognitively demanding task. Participants were subsequently invited to log in to their profiles after three days for further password/passphrase recall tests. This is consistent with the literature (Melicher et al., 2016; Shay et al., 2016; Woods & Siponen, 2018, 2019). After the fourth and final logging in event, participants in this study were presented with the second questionnaire, which gathered data for evaluating password/passphrase recall usability, as well as data on user satisfaction which was in turn used to measure usability. Additional data for evaluating usability was gathered during password/passphrase generation and periodic logging in events by means of key logs. The activities of short password and multilingual passphrase generation and recall were done in succession, separated by a period of two weeks to allow participants to do each task on a fresh note**.** This study adopted and adjusted a two-word 16 (2word16) character passphrase policy reported in Shay et al. (2016). This policy has been shown to yield better security and usability results. Maoneke, Flowerday and Isabirye (2020) gave a theoretical overview of the multilingual passphrase policy. Below are the short password and multilingual passphrase policies that were used in this study:

Short password policy: participants were required to generate a short password that

- had to be at least 8 characters long, without any spaces
- had to contain at least one capital letter, one lowercase letter, one number and one special character (symbols like &, $, @, #, !, *)
- did not contain their username or their personal details.

Multilingual passphrase policy: participants were required to create passphrases with the following characteristics:

- Have at least two words.
- The words making the passphrase should be from at least two different languages, for example an English word and the other could be an African word.
- The words making up the passphrase should be separated by at least one non-letter sequence.

A dictionary check was used to ensure that participants generated multilingual passphrases. Words and password dictionaries (wordlist) were used in the background to check whether the passphrase entered by a user was oriented to a single Indo-European language. The wordlist was downloaded from OpenWall, the use of which as a wordlist source is common practice in the literature (Shay et al., 2015). Given that the research context of this study involved two Indo-European languages, English and Afrikaans, dictionaries with words from these languages were integrated in the web-application such that the use of English- or Afrikaans-only passphrases was restricted. The participants in this study comprised students from selected South African and Namibian universities. The behavior and attitudes of university students towards passwords are comparable to those of the general population (Alomari & Thorpe, 2019). Raw passwords were also gathered for analysis during password generation.

Password security and usability analysis

This study compared the security and usability contributions of the multilingual passphrase policy against that of the short password policy (Komanduri, 2016; Melicher et al., 2016; Shay et al., 2016). Guess numbers were used to estimate the strength of the passphrases and short passwords. Guess number shows the resistance of a short password or passphrase to password guessing and a probabilistic context-free grammar (PCFG) password-guessing algorithm proposed by Komanduri (2016) was used for guess number estimations. In addition, this study used statistical techniques to analyze the usability of the short password and the multilingual passphrase policy. Data that was gathered through the use of the questionnaires was analyzed using a one-sample t-test to establish the equality of means for user satisfaction. In addition, a paired samples t-test was used to compare participants' perceptions of usability between the short password and the multilingual passphrase policy. These usability comparisons were done on password or passphrase creation and recall. In addition, categorical data that was gathered by use of key logs was also used to support usability findings on password or passphrase creation and recall.

Sample profile

Two hundred and twenty-four (224) students participated in the experiment. However, while 224 participants created a short password, only 141 went on to participate in short password recall. In the case of the multilingual passphrase policy, 176 participants created a multilingual passphrase and 132 of these took part in the multilingual passphrase recall exercises. It should be noted that 112 participants completed both phases of the experiment: short passwords and multilingual passphrases. In other words, 50% of the participants (224) dropped out during the course of the experiment.

An analysis of gender distribution across participants shows that 54.85% was male and 44.15% was female. In addition, the majority (88%) of participants were within the age group of 18 to 25 years. This was followed by 11% of the participants who fell into the 26 to 35 age group.

Password/passphrase generation usability

This section presents findings on the short password and multilingual passphrase policy usability, focusing on the use of personal identification information, password reuse, the use of common passwords and the number of password generation attempts.

The use of personal identification information: Participants were asked to indicate their password and passphrase generation strategy. In this regard, the findings show that adapting a name is by far (49%) the most common use of personal identification information in short password and multilingual passphrase generation. Few participants included details pertaining to their date of birth (15% in short passwords and 10% in multilingual passphrases) and very few (approximately 5%) included a phone number in their short passwords and multilingual passphrases.

Password reuse: Participants were asked to indicate whether they had arrived at their short passwords or passphrases by adapting existing passwords. Approximately 38% of the participants agreed to adopting an existing short password while 23% indicated that they had adopted an existing password when generating a multilingual passphrase. In addition, a comparison was done between a short password and a multilingual passphrase generated by the same participant to establish the practice of password reuse. A multilingual passphrase with a substring of at least four consecutive characters that matched those in the short password was regarded as a multilingual passphrase that was arrived at after password reuse as done by Shay et al. (2014). It was observed that 19% (33) of the passphrases had substrings that were similar to those found in short passwords. In cases where participants reused their short passwords, a substring on the far left of the generated multilingual passphrase was often (72%) a result of adapting a short password.

The study went on to investigate the magnitude of transformation that was done to a short password as it was transformed into a substring of a multilingual passphrase. Accordingly, Levenshtein's edit distance was used to measure the distance between a substring in a passphrase and a short password (Das, Bonneau, Caesar, Borisov, & Wang, 2014). In this case, the edit distance was used to show the number of characters that needed to be changed in order to convert a substring in a multilingual passphrase to a short password. Results from Levenshtein's edit distance showed that 6% of the multilingual passphrase corpora had a substring that was identical to a short password. One of the two substrings of these passphrases were based entirely on the original short password that was keyed in twice, separated by a space to come up with a multilingual passphrase that had two substrings. Furthermore, 7% of the multilingual passphrase corpora had a substring that differed from the reused short password by one to three characters. The remaining 5% had substrings that were four to seven characters away from the original short password.

The use of common passwords: A content analysis was carried out to establish the use of words or passwords that could be easily guessed, as done in the literature (Bonneau & Shutova, 2012; Shay et al., 2016). The focus was on establishing whether a subset of at least four successive characters or the whole of the password appeared among the top one hundred most common passwords of 2016, 2017, and 2018 according to SplashData. Subsequently, close to 3% of the short password corpora were found to have substrings that appeared in the list of the most common passwords, whereas only 4% of the passphrase corpora displayed a substring that is among the common passwords. Focusing on at least four successive characters (Shay et al., 2014) to identify common passwords might have exaggerated this result. Making comparisons based on the base word may give different results. A number of short passwords contained various mangling rules that distanced them from selected passwords in the list of most common passwords. For example, "P@55w0rd777," "P@5sword," and "Pa$$word2" found in the short password corpora could still be traced back to password or password1 which appear in the list published by SplashData.

Generation attempts: Data on key logs was gathered and used to ascertain the number of attempts required to generate a short password or multilingual passphrase. It was assumed that more passphrase or short password generation attempts or time spent would suggest that the policy is less usable (Shay et al., 2016). The study results showed that 4.5 attempts were required, on average, to generate a multilingual passphrase whereas it had taken 2.1 attempts, on average, to generate a short password. These findings were corroborated by the average of 82.2 seconds that was required to generate a short password and the 34.3 seconds that was spent on confirming the short password generated. Hence, it took 116.5 seconds in total to generate a short password. Inversely, it took participants 246.9 seconds, on average, to generate a multilingual passphrase and 74 seconds to confirm the generated multilingual passphrase. A non-parametric test using the Wilcoxon signed ranks test showed that multilingual passphrase generation required significantly more time compared to the time needed to generate short passwords.

Password/passphrase generation user satisfaction: Findings on user satisfaction show that participants found both the short password and multilingual passphrase policy user-friendly. For instance, the results of a one-sample test showed that participants rated short password generation user satisfaction (mean = 3.8880, SD = 0.97012, t = 13.639, p = < 0.0001) significantly higher than 3. Similarly, multilingual passphrase generation user satisfaction (mean = 3.7874, SD = 0.99814, t = 10.405, p = < 0.0001) was also rated significantly higher than 3. However, a mean rating comparison between short password and multilingual passphrase generation usability using a paired-samples t-test showed that the mean for short password generation user satisfaction (mean = 0.1701; df = 109; t = 2.053; p = 0.042) was significantly higher compared to that of multilingual passphrases.

Password/passphrase recall usability

This section presents the findings on the short password and multilingual passphrase policy recall usability. These findings are reported according to password/passphrase recall strategy, authentication challenges, login attempts, and user satisfaction with the researched policies.

Password/passphrase recall strategy: The literature review in this study showed that users resort to writing down passwords, password reuse, and the use of password managers with the aim of aiding memorability. The results of this study show that approximately 80% of the participants indicated that they had managed to memorize their short passwords and multilingual passphrases. However, 24% of the participants saved their short passwords on their phones and 27% wrote their short passwords down in case they forgot them. Inversely, approximately 29% of the participants wrote down their multilingual passphrases while 27% saved their multilingual passphrases in a mobile phone in case they forgot. A paired samples t-test was conducted in order to make comparisons relating to short password and multilingual passphrase recall strategy. No significant difference on mean ratings for password recall strategy was found (mean = 0.1111; df = 110; t = 1.273; p = 0.206). However, a Levene's test for homogeneity of variance revealed a significant difference in mean ratings on password recall strategy across gender. Males (mean = 4.3734; SD = 0.72413) showed a significantly higher mean rating for short password recall strategy than their female (mean = 3.8629; SD = 0.1.08245) counterparts (t = 3.195; Pr > |t| = 0.002). Similarly, males had significantly higher mean ratings for multilingual passphrase recall strategy (mean = 4.2524; SD = 0.78212; t = 2.523; Pr > |t| = 0.013) when compared to their female counterparts.

Authentication challenges: Data gathered from key logs was analyzed to establish authentication challenges faced by participants. Only 25% of all those who participated in short password recall failed to accurately key in their passwords. This is in stark contrast to the 40% of participants who failed to completely and accurately type in their multilingual passphrases. Those who failed to key in their multilingual passphrases accurately ended-up formulating candidate passphrases that were 4.2 characters, on average, away from the actual multilingual passphrase. Nevertheless, the 10% of cases that failed to accurately key in a short password were a result of typographical errors. Inversely, 26% of all the login failures for multilingual passphrases was due to typographical errors. Figure 1 shows the causes of password recall failure in this study, excluding typographical errors.

Logging in attempts: Key logs data was analyzed to establish password recall usability in terms of the number of login attempts that were required to login successfully. Participants needed 1.3 login attempts, on average, to accurately log into their profiles using a short password compared to 2 attempts that were needed, on average, when logging in with a multilingual passphrase. It should be noted that password recall was tested on four occasions with each occasion separated by three days. Research findings show that it took 30 seconds, on average, to successfully login to a profile using a short password on one's first return. However, it required 132.9 seconds on average for participants to accurately login to their user profiles during their first login session using a multilingual passphrase. The average login time dropped to 33, 26 and 21 seconds on the second, third and fourth login attempts, respectively, when using short passwords. This was in stark contrast to the 60.3, 28.2 and 48.7 seconds that were needed on average to accurately key in a multilingual passphrase on participants' second, third and fourth return, respectively.

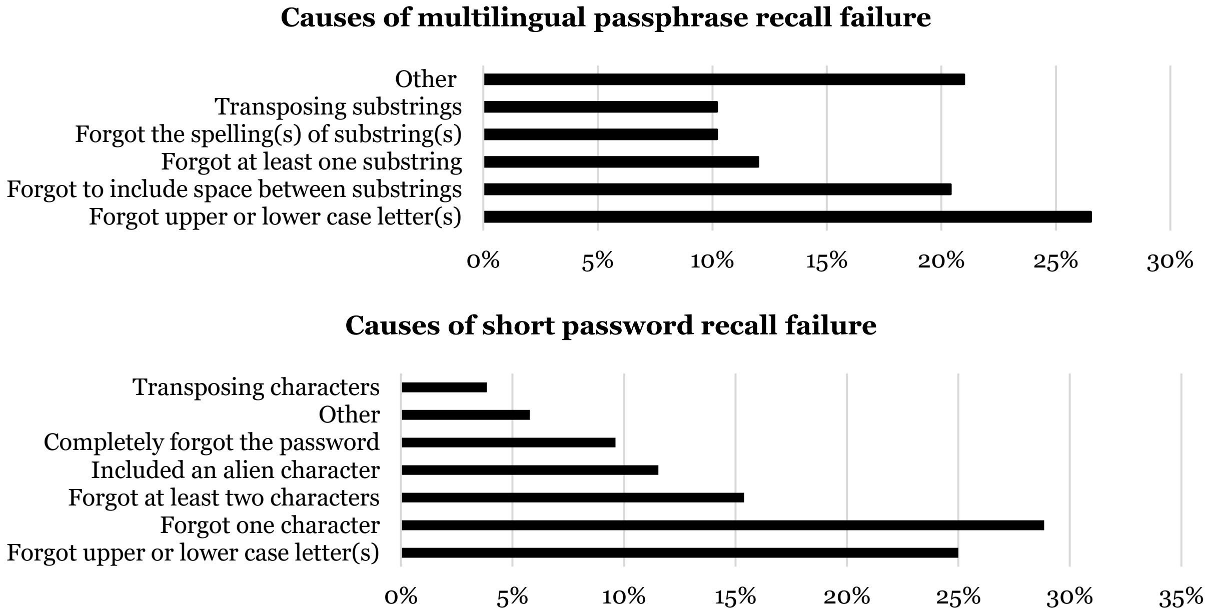

Figure 1. The common causes of short passwords and multilingual passphrase recall failure (expressed as a percentage of all the passphrase/password recall failure)

Password recall user satisfaction: The results of a one-sample test showed that, on average, participants rated user satisfaction with short password recall significantly higher than 3 (mean = 3.8905, SD = 0.85774, t = 12.284, p = < 0.0001). A similar result was observed for multilingual passphrase recall user satisfaction (mean = 3.8168, SD = 1.01229, t = 9.235, p = < 0.0001). A subsequent paired samples t-test comparing short password and multilingual passphrase recall went on to show that there was no significant difference on mean ratings for recall user satisfaction (mean = 0.1439; df = 100; t = 1.339; p = 0.183).

Password security

Short passwords and multilingual passphrases generated in this study were sent for password guessing using a PCFG. The cut-off for guessing passwords and passphrases was set at 10^{15} guessing attempts. Exactly 50.4% of the short passwords were guessed while none of the multilingual passphrases were guessed. This despite the presence of substrings that were among the popular passwords of 2016, 2017 and 2018; according to SplashData; in some of the passphrases. An analysis was done to establish the security benefits involved in orienting passwords towards native African languages. It was noted that 6% of the short passwords oriented towards the English language were guessed within the first one million guessing attempts compared to none of the short passwords oriented towards an African language. At the end of 10^{15} guessing attempts, 68% of the short passwords oriented towards an English language were guessed compared to 40% of the short passwords that were oriented towards an African language.

Discussion

Password policies that include passphrases are widely regarded as more secure and usable (Melicher et al., 2016; Shay et al., 2016). This study explored the security and usability contributions of a multilingual passphrase policy. Understanding both password policy security and usability is important given that, in a trade-off between password security and usability, security is most likely to be compromised in favor of usability (Choong et al., 2014; Woods & Siponen, 2019). The usability of passphrases was evaluated during passphrase generation and recall. In particular with regard to password generation, participants showed a heavy reliance on adopting names as short passwords and multilingual passphrases. This is a common practice and has been discussed extensively in the literature, where users do this to enhance usability (AlSabah et al., 2018; Choong et al., 2014). This tendency to adapt names for short passwords or passphrases confirms that users have personal password generation routines that they are unwilling to change even when subjected to different password policies (Renaud, Otondo, & Warkentin, 2019). Furthermore, password reuse, which is often done to enhance usability, was more pronounced under the short password policy (38%) than the multilingual passphrase policy (23%). The practice of password reuse in this study can be explained by the theory of interference (España, 2016; Zhang et al., 2009). As such, the difference in the levels of password reuse can be explained by the fact that participants in this study found it easy to exploit already known information when generating multilingual passphrases (España, 2016). In addition, results pertaining to the amount of time and the number of attempts needed to generate multilingual passphrases suggest that passphrases in this study were not so easy to generate when compared to short passwords. This was confirmed by the participants' perceptions on user satisfaction with the multilingual passphrase policy. Less exposure to passphrase policies and a high percentage of password reuse were reported under the short password policy, which might have exaggerated this result.

In particular, with regard to password recall, participants found both the short password policy and the multilingual passphrase policy to be user-friendly. Thus, participants did not significantly change their password recall routines as a result of migrating from a short password to a multilingual passphrase policy. It can be argued that multilingual passphrases are easy to learn and recall to the extent that participants had no reason to change their password recall strategy. However, females appeared to struggle with both short password and multilingual passphrase recall, with significantly more females than males showing their reliance on writing down passwords or storing them on their phones or sharing passwords with a colleague in case they forgot them. This finding could be explained by the fact that males are more likely to reuse "passwords than females," something that might have aided the ability of males to recall passwords (Helkala & Bakås, 2013, p. 352). In addition, participants experienced a constant drop in the amount of time needed to successfully login with each login session. This can be explained by the fact that participants were learning and getting used to their passwords. Studies show that the repeated use of passwords after generation can help facilitate the transfer of such passwords to the long-term memory, something that is supported by theories on memorability (Atkinson & Shiffrin, 1968; Woods & Siponen, 2019). Lastly, the study findings suggest that the multilingual passphrase policy successfully influenced participants to generate secure passphrases when compared to the short password policy used in this study and passphrase policies reported in the literature. Close to 25% of the 2word16 character passphrases in Shay et al. (2016) were guessed by the PCFG compared to none of the multilingual passphrases in this study. It can therefore be concluded that basing passphrases on different languages enhances security.

Limitations and future research

All participants in this study were aware that they were taking part in an experiment, something that may have negatively affected ecological validity. In addition, the password guessing algorithm for this study was not configured to be aware of the dictionary and grammar of the languages in the multilingual passphrases. Future research could explore the multilingual passphrase policy in a real-world setting.

Conclusion

This study proposed and evaluated the security and usability of a multilingual passphrase policy. The findings showed that multilingual computer users have the potential to generate secure and usable multilingual passphrases. However, challenges during password generation suggest that more effort should be put into guiding users when generating multilingual passphrases. In addition, the tendency to base passphrases on names calls for robust measures that could restrict the use of personal identification information as passphrases. Nonetheless, relatively low password reuse under the multilingual passphrase policy suggests that the policy gave users the room to be more creative during passphrase generation. However, in cases where there is password reuse, the first substring of the passphrase is likely to be similar or close to a known short password. Furthermore, it has been shown that repeated use of a multilingual passphrase will help users to learn and better recall their passphrases quickly. Lastly, passphrases in this study proved to be stronger than the short passwords and passphrases reported in the literature. Hence, systems administrators will not need to reset and change a multilingual passphrase as often as a short passwords and passphrases reported in the literature. This will save time and reduce user frustration.

REFERENCES

Alomari, R., and Thorpe, J. 2019. "On Password Behaviours and Attitudes in Different Populations," *Journal of Information Security and Applications* (45), pp. 79–89.

AlSabah, M., Oligeri, G., and Riley, R. 2018. "Your Culture is in Your Password: An Analysis of a Demographically-Diverse Password Dataset," *Computers & Security* (77), pp. 427–441.

Alter, S. 2014. "Theory of Workarounds," *Communications of the Association for Information Systems* (34:55*)*, pp. 1041-1066.

Atkinson, R. C., and Shiffrin, R. M. 1968. "Human Memory: A proposed System and Its Control Processes," in *The Psychology of Learning and Motivation, Advances in Research and Theory,* K. W. Spence and J. T. Spence (eds.), (2), pp. 89–191.

Becker, I., Parkin, S., and Sasse, M. A. 2018. "The Rewards and Costs of Stronger Passwords in a University: Linking Password Lifetime to Strength," in *Proceedings of the 27th USENIX Security Symposium,* Baltimore: USENIX Association, pp. 239–253.

Bonneau, J., and Schechter, S. 2014. "Towards Reliable Storage of 56-bit Secrets in Human Memory," in *Proceedings of the 23rd USENIX Security Symposium,* San Diego: USENIX, pp. 607–623.

Bonneau, J., and Shutova, E. 2012. "Linguistic Properties of Multi-word Passphrases," in *International Conference on Financial Cryptography and Data Security,* J. Blyth, S. Dietrich, and J. L. Camp (eds.), Berlin: Springer, pp. 1–12.

Carstens, D. S., Malone, L. C., and McCauley-Bell, P. 2006. "Applying Chunking Theory in Organizational Password Guidelines," *Journal of Information, Information Technology, and Organizations* (1), pp. 97–113.

Choong, Y.-Y., Theofanos, M., and Lui, H.-K. 2014. "United States Federal Employees' Password Management Behaviors – a Department of Commerce Case Study," Retrieved from https://ws680.nist.gov/publication/get_pdf.cfm?pub_id=914843

Cowan, N. 2014. "Working Memory Underpins Cognitive Development, Learning, and Education," *Educational Psychology Review* (26:2), pp. 197–223.

Das, A., Bonneau, J., Caesar, M., Borisov, N., and Wang, X. 2014. "The Tangled Web of Password Reuse," in *Proceedings of The Network and Distributed System Security (NDSS),* San Diego: Internet Society, pp. 23–26.

Deumert, A., and Lexander, K. V. 2013. "Texting Africa: Writing as Performance," *Journal of Sociolinguistics* (17:4), pp. 522–546.

Duggan, G. B., Johnson, H., and Grawemeyer, B. 2012. "Rational Security: Modelling Everyday Password Use," *Journal of Human Computer Studies* (70:6), pp. 415–431.

España, L. 2016. "Effects of Password Type and Memory Techniques on User Password Memory," *PSI CHI Journal of Psychological Research* (21:4), pp. 269–276.

Everitt, K. M., Bragin, T., Fogarty, J., and Kohno, T. 2009. "A Comprehensive Study of Frequency, Interference, and Training of Multiple Graphical Passwords," in *Proceedings of SIGCHI Conference on Human Factors in Computing Systems,* Boston: ACM, pp. 889–898.

Grassi, P. A., Garcia, M. E., and Fenton, J. L. 2017. "Digital Identity Guidelines. NIST Special Publication," United States of America, 800-63-3.

Habib, H., Colnago, J., Melicher, W., Ur, B., Segreti, S., Bauer, L., and Cranor, L. 2017. "Password Creation in the Presence of Blacklists," in *USEC Conference*, San Diego: Internet Society, pp. 1–11.

Keith, M., Shao, B., and Steinbart, P. 2009. "A Behavioral Analysis of Passphrase Design and Effectiveness," *Journal of the Association for Information Systems* (10:2), pp. 63–89.

Komanduri, S. 2016. "Modeling the Adversary to Evaluate Password Strength With Limited Samples," Carnegie Mellon University.

Komanduri, S., Shay, R., Kelley, P. G., Mazurek, M. L., Bauer, L., Christin, N., and Egelman, S. 2011. "Of Passwords and People: Measuring the Effect of Password-Composition Policies". in *SIGCHI Conference on Human Factors in Computing Systems*, Vancouver: ACM New York pp. 2595–2604.

Li, Z., Han, W., and Xu, W. 2014. "A Large-Scale Empirical Analysis of Chinese Web Passwords," in *Proceedings of 23rd USENIX Security Symposium*, San Diego: USENIX Association pp. 559–574.

Maoneke, B. P., Flowerday, S., and Isabirye, N. 2020. "Evaluating the Strength of a Multilingual Passphrase Policy," *Computer & Security* (92), pp. 1–14.

Maoneke, P. B., Flowerday, S. and Isabirye, N., 2018. "The Influence of Native Language on Password Composition and Security: A Socioculture Theoretical View." in *IFIP International Conference on ICT Systems Security and Privacy Protection,* Poznań: Springer Cham, pp. 33-46.

Melicher, W., Kurilova, D., Segreti, S. M., Kalvani, P., Shay, R., Ur, B., and Mazurek, M. L. 2016. "Usability and Security of Text Passwords on Mobile Devices," in *Conference on Human Factors in Computing Systems*, San Jose: ACM New York, pp. 527–539.

Miller, G. A. 1956. "The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information," *The Psychologial Review* (63:2), pp. 81–97.

Rao, A., Jha, B., and Kini, G. 2013. "Effect of Grammar on Security of Long Passwords," in *Proceedings of the 3rd ACM Conference on Data and Application Security and Privacy*, San Antonio: ACM New York, pp. 317–324.

Renaud, K., Otondo, R., and Warkentin, M. 2019. "This is the Way 'I' Create my Passwords ... Does the Endowment Effect Deter People from Changing the Way they Create Their Passwords?," *Computers & Security* (82), pp. 241–260.

Shay, R., Kelley, P. G., Komanduri, S., Mazurek, M. L., Ur, B., Vidas, T., and Cranor, L. F. 2012. "Correct Horse Battery Staple: Exploring the Usability of System-assigned Passphrases," in *Symposium on Usable Privacy and Security (SOUPS)*, Washington DC: ACM New York, pp. 1–20.

Shay, R., Komanduri, S., Durity, A. L., Huh, P. (Seyoung), Mazurek, M. L., Segreti, S. M., and Cranor, L. F. 2016. "Designing Password Policies for Strength and Usability," *ACM Transactions on Information and System Security* (18:4), pp. 1–34.

Shay, R., Komanduri, S., Durity, A. L., Huh, P. S., Mazurek, M. L., Segreti, S. M., and Cranor, L. F. 2014. "Can Long Passwords Be Secure and Usable?," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Toronto, ON, Canada: ACM New York, pp. 2927–2936.

Shay, R., Ur, B., Bauer, L., Christin, N., Cranor, L. F., Forget, A., and Segreti, S. M. 2015. "A Spoonful of Sugar? The Impact of Guidance and Feedback on Password-Creation Behavior," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems,* Seoul: ACM New York, pp. 2903–2912.

Stobert, E., and Biddle, R. 2014. "The Password Life Cycle: User Behaviour in Managing Passwords," in *Proceedings of the Tenth Symposium On Usable Privacy and Security (SOUPS),* Menlo Park: USENIX, pp. 243–255.

Tam, L., Glassman, M., and Vandenwauver, M. 2010. "The Psychology of Password Management: a Tradeoff Between Security and Convenience," *Behaviour and Information Technology* (29:3), pp. 233–244.

Vedadi, A., and Warkentin, M. 2018. "Secure Behavior over Time: Perspectives from the Theory of Process Memory," *The DATA BASE for Advances in Information Systems* (49), pp. 39–48.

Veras, R., Collins, C., and Thorpe, J. 2014. "On the Semantic Patterns of Passwords and their Security Impact," in *NDSS Symposium,* San Diego: Internet Society, pp. 1–16.

Wang, D., Cheng, H., Gu, Q., and Wang, P. 2015. "Understanding Passwords of Chinese Users: Characteristics, Security and Implications," in *Proceedings of the ChinaCrypt*, pp. 1–14.

Warkentin, M., Davis, K., and Bekkering, E. 2004. "Introducing the Check-Off Password System (COPS): An Advancement in User Authentication Methods and Information Security," *Journal of Organizational and End User Computing* (16:3), pp. 41–58.

Wei, M., Golla, M., and Ur, B. 2018. "The Password Doesn't Fall Far: How Service Influences Password Choice. In *Who Are You?," Adventures in Authentication (WAY)*, Baltimore.

Weir, M., Aggarwal, S., Collins, M., and Stern, H. 2010. "Testing Metrics for Password Creation Policies by Attacking Large Sets of Revealed Passwords," in *17th ACM Conference on Computer and Communications Security* , Chicago: ACM New York, pp. 162–175.

Woods, N., and Siponen, M. 2018. "Too Many Passwords? How Understanding Our Memory Can Increase Password Memorability," *International Journal of Human - Computer Studies* (111), pp. 36–48.

Woods, N., and Siponen, M. 2019. "Improving Password Memorability, While not Inconveniencing the User," *International Journal of Human Computer Studies* (128), pp. 61–71.

Zhang, J., Luo, X., Akkaladevi, S., and Ziegelmayer, J. 2009. "Improving Multiple-Password Recall: An Empirical Study," *European Journal of Information Systems* (18:2), pp. 165–176.