Open sesame: Lessons in password-based user authentication

Received (in revised form): 27th August, 2020

Bahman Rashidi
Senior Cyber Security Researcher and System Architect, Comcast Cable, USA

Bahman Rashidi is a Senior Cyber Security Researcher and System Architect at Comcast Cable. At Comcast, his responsibilities focus on conducting cyber security research and development in the areas of network security, Internet of Things (IoT) devices, applications of machine learning (ML) in cyber security, privacy and data protection and quantum computing/key distribution. He designs and builds strategic cyber security technologies and tools on the bleeding edge, including building networks, led teams, architected security infrastructure and security solutions.

Comcast Cable, 1800 Arch Street, Philadelphia, PA 19103, USA
Tel: +1 215 286 3583; E-mail: bahman_rashidi@comcast.com

Vaibhav Garg
Senior Director of Cybersecurity Research & Public Policy, Comcast Cable, USA

Vaibhav Garg is the Senior Director of Cybersecurity Research & Public Policy at Comcast Cable. He has a PhD in security informatics from Indiana University and a Master's in information security from Purdue University. His research lies at the intersection of information security, technology policy and economics. He is the co-author of over 20 peer-reviewed publications and received the best paper award for his paper on the economics of cybercrime at eCrime 2011. He also received ACM Computer and Society's Outstanding Service award in 2015 for his contributions as editor-in-chief of *ACM Computers and Society Newsletter*.

Comcast Cable, 1800 Arch Street, Philadelphia, PA 19103, USA
Tel: +1 571 409 0993; E-mail: vaibhav_garg@comcast.com

Abstract The cost of unusable password policies in the wild is well documented. These costs impinge both business and security. The alternative is to move to multi-factor and risk-based authentication, which include software authenticators, hardware tokens, and biometrics. This paper provides an overview of the research in this area and concludes with guidance on how to best leverage password-based authentication. We recommend that designers should only implement efforts backed by empirical evidence, offer solutions to reduce user effort, and use compensating controls to address the underlying limitations of passwords.

KEYWORDS: passwords, biometrics, 2FA, MFA, authentication

INTRODUCTION

User authentication is the first line of defence when ensuring the security of information systems and protection of online accounts as well as personal data from unauthorised access. All authentication is based on one of three criteria: 1) what you know; 2) what you have; or 3) what you are. From ancient fables to modern IT systems, all identity verification both relies on these three considerations and is equally hampered by their inherent limitations.

The first of these — what you know — refers to a secret that only the authorised

party should know. Consider the fable of *Ali Baba and the Forty Thieves*. The latter secured their assets, ie loot, in a cave to be accessed only if the right password was presented. Unfortunately for the thieves, they all shared a single password that was never changed. Thus, their security was compromised by a simple eavesdropping attack. This same tale also warns us of the cost of forgetting passwords. Cassim, Ali Baba's brother, forgets the password, is unable to leave the cave and is killed by the thieves on their return. Thus, in a single fable we are made aware both of the challenges of ensuring good cyber hygiene for knowledge-based user authentication as well as the associated cost imposed on user experience.[1]

Different proposals have addressed the inherent limitations of password-based authentication. For example, users are commonly advised not to share passwords across multiple accounts.[2] In practice this results in numerous passwords that the user may have a difficult time remembering. Users may then use password managers, but these both concentrate risk in a single point and may themselves pose a security risk.[3,4]

A different line of defence is to consider passwords inadequate in isolation. The solution then is to employ secrets-based authentication in combination with one or more additional authentication mechanisms, two-factor authentication (2FA) or multi-factor authentication (MFA). These factors further affect user experience and also embody distinct challenges. For example, facial recognition systems may be less accurate for women and individuals with darker skin tones.[5] Other factors, such as fingerprint recognition systems, may be defeated by synthetic reproductions.[6]

All user authentication mechanisms, then, have inherent challenges. Yet, system designers must still implement solutions to verify user identity. This paper is intended to provide guidance for practitioners and empower decisions that will improve both user experience as well as security. Thus, the paper presents a holistic summary of the lessons from decades of research in password usability and corresponding impact on security.

We begin by reviewing the various proposals intended to improve the reliance on text-based passwords and their effectiveness. This is followed by an exploration of alternative secret-based authentication and its constraints. The next section addresses authentication mechanisms based in 'what you have' or 'what you are', which can be used to complement passwords or password-adjacent schemes. Finally, the paper concludes with a list of considerations for designing password-based authentication systems, policies or processes.

PASSWORDS: THE GOOD, THE BAD, AND THE COMPLICATED

Secrets-based authentication, ie passwords, are the first line of defence in many systems; however, password-based authentication has some inherent challenges. Passwords by definition must be difficult to guess, which often makes them difficult to remember as well. This difficulty is compounded by the fact that users must choose unique passwords for every account. As the number of accounts continues to increase, retaining all the associated passwords in rote memory becomes impossible. This results in users employing various compensating behaviours.

The first of these is password reuse, ie users do not choose unique passwords for each account but instead use the same password across multiple domains. Second, users choose unique passwords but store them outside of their memory, eg writing them on paper, saving them in a digital document, etc. Users come up with an 'algorithm' which allows them to choose the same basic password but adds some entropy based on the site. For example, they may use a combination of a base password and the account name. So, the passwords will then be 'passwordsite1', 'passwordsite2' and so on.

Many proposals, both technical and non-technical, have been made to address the inherent challenges of password-based authentication as well as the associated compensating strategies.[7] In this section we present a review of these proposals and differentiate the good from the less effective ones.

Password strength

The foremost issue with using passwords is that users choose weak passwords. A list of most common passwords claims that 91 per cent of all sampled passwords are limited to only one thousand passwords.[8] Furthermore, users tend to choose common words that are easy to guess as part of their passwords. For example, Qataris often use 'qatar' or 'doha' in their password.[9] To mitigate the risk of user-chosen passwords having low entropy, system designers have tried the strategies below.

Composition (complexity) policies

Composition policies are a set of pre-defined rules that the password string should comply with. Examples of such policies are the use of lower/upper case characters, digits, special characters, or not allowing certain common words and names. A composition policy is used to increase resilience of a password against guessing or brute force attacks; however, stricter composition rules make it more difficult for users to remember their passwords.[10] Consequently, most (US) users keep track of their online passwords by either using password managers or writing them down on a piece of paper.

Alternatively, users may reuse the same password (or a combination of the same password) across multiple websites. Ironically, password composition policies have limited impact on password strength. Even the strictest password policies can be manipulated to choose a weak password. For example, studies show that users mostly used '@' and '!' characters out of 28 special characters at either the beginning or end of the passwords.[11,12] Thus, instead of increasing the entropy of passwords, such policies only serve to reduce an attacker's search space.

Password expiration

Another solution to mitigating the risk from weak passwords and corresponding guessing attacks is password expiration. This in theory also protects against accounts compromised due to other password disclosure attacks, eg phishing, keyloggers, etc. The payment card industry (PCI), for example, recommends that passwords should be changed every 90 days. The 2003 version of National Institute of Standards and Technology (NIST) 'Special Publication 800-63 Appendix A' similarly recommended password expiration, as back then it was an industry best practice.[13] As with many things security, the best practice was really a common practice. Since then research has shown that users tend to change their passwords in predictable patterns, such as adding a single character to the end of their last password or replacing one letter with another.[14]

Forced password resets also lead to other coping strategies, such as: 1) choosing a weak but easy to remember password; 2) password reuse; or 3) simply writing down passwords.[15] While NIST has since changed its position (as well has the original author of the guidance), others like PCI have continued to stick to this increasingly antiquated idea. PCI requires that the new password should not be the same as or similar to the last four passwords.

Strength meter

A password strength meter is an indicator, either in graphical or text format, that tells users how strong their passwords are. Strength meters compare passwords against a list of rules and policies or compute the strength through running mathematical formulations on the client side. Strength meters come in various formats, strictness

and colours. Some of the strength meters show users a level of strength (weak, fair and strong) in real time as users type in their passwords. Another meter, known as a 'peer-pressure-motivator', tells users how strong their password is compared to other users' passwords.[16] Research shows that strength meters help users select stronger passwords when setting a password for important (sensitive) accounts. The same study found no observable difference in the password strength for less important accounts. One other challenge with strength meters is that the stronger the passwords are, the harder it is to memorise them. A different research study showed that the scoring stringency and having a visual component are the most informative features. They found that colour, font and size of texts, and the shape of password meters have an effect on the usability.[17]

Passwords through persuasion

The challenges and issues with passwords have opened the door to psychological factors as well. There have been some efforts in the field that take personality, behavioural and social psychology factors into consideration. Such approach is called persuasive technology, aiming at interacting with users to change users' attitudes and behaviours. This technology is built upon the belief that users are not enemies of security, but collaborators who need effective guidance to choose strong and memorable passwords. This technology focuses on behavioural factors in helping users to set strong passwords while keeping the system usable. For instance, researchers have designed an approach called persuasive text password (PTP) in which users set a password of their choice and the PTP system inserts additional characters at random places. Users have the option to shuffle the characters and find a combination they feel is memorable. The study shows that this solution helps with generating stronger passwords while memorability of passwords remained at an acceptable level.[18]

Password memorability

If passwords are adequately random, they can be difficult for users to remember. Studies indicate that users need to remember more than 25 different passwords on average.[19] Thus, users employ compensating behaviours such as writing down passwords (˜55 per cent) and password reuse (˜40 per cent).[20]

Password manager

To alleviate the cognitive cost of remembering multiple passwords, users may delegate to a password manager, which stores a multitude of secrets and is itself protected by a master secret. Unfortunately, these password managers themselves then become a central point of attack and failure. Web-based password managers are vulnerable to, among other attacks, bookmarklet vulnerabilities, classic web vulnerabilities, authorisation vulnerabilities and user interface (UI) vulnerabilities.[21]

Additionally, the password manager may face compatibility challenges with different browsers, operating systems, applications and devices.[22] This makes password managers harder to use and limits adoption.[23] It has been noted that adoption is primarily driven by convenience rather than security gain even though most users of password managers tend to report higher computer proficiency.[24] Thus, password managers are not perceived by users as a tool to increase password security.

This intuition may be supported by research. Password managers may simultaneously reduce password reuse as well as increase the strength of the unique passwords, if used with a password generator.[25] They further note that Chrome's widely used password manager, if used without a generator, may in fact exacerbate the user's worst impulses. For example, in

their study, Chrome users were more likely to reuse passwords.

Federated authentication

As noted previously, password managers may have limited impact on mitigating the risks from passwords due to limited adoption as well as the vulnerabilities in password managers themselves. The alternative solution, then, is to reduce the overall number of passwords a user needs. Federated authentication enables entities under one or different organisations to work with each other without requiring users to be authenticated at every unit of service. In other words, a single authentication service is used to authenticate users' access to several accounts.

An implementation of this is the single sign-on service (SSO).[26] This can potentially improve user experience by reducing the number of times a user needs to authenticate as well as the number of identity/authentication pairs they need to create. As with password managers, SSOs also centralise the risk. Compromise of an SSO system can potentially allow an attacker to all associated services.

Previous research has identified vulnerabilities in OpenID and OAuth, two popular SSO protocols used by Microsoft, Google and Facebook.[27,28] Even in the absence of technical or implementation-based vulnerabilities, these protocols may be susceptible to human factors-based attacks. For example, many SSO system providers may be easily spoofed and lend themselves to phishing attacks.[29]

Unlike password managers, SSO systems are perceived to be more usable.[30] Attention to design can still improve user experience. Linden and Vilpola, for example, note that a single sign-on should also imply single logout.[31] In addition, previous studies have noted the distinct mental models of SSO.[32] The first is that of a master key, the second that of keyless entry. Clarifying which model is supported by the design is important for user experience. Waters notes that a poorly designed SSO system may result in users oversharing information with third parties.[33] While not a security risk, this does impinge a user's ability to adequately control their information exposure and can limit adoption of SSO.[34]

Password reuse

A final threat to password security is password reuse. This concern is not just driven by password memorability. As noted previously, users often reuse password even when using a password manager.[35] While in theory password complexity policies can prevent the same password being reused across multiple accounts, in practice researchers note that a well-constructed password may satisfy as many as 99 per cent of password policies.[36] One study shows that more than 50 per cent of users in an interview study reused passwords simply because it would be too hard for them to remember them all. Analysis of leaked password datasets showed that more than 43 per cent of identical usernames in two different datasets had the same password. Password reuse introduces a security vulnerability because if credentials for one account are compromised, an attacker can easily get access to other accounts.[37] The primary approaches to mitigate the risk from password reuse are password blacklists and password rate limiting.

Password disallow list

Employees often tend to use the same password that they set for their personal accounts for accounts related to their job. Thus, if a personal account is compromised, attackers may potentially get access to non-personal accounts. While many organisations explicitly prohibit password reuse in their policies, it is impossible to verify. The closest solution has been to incorporate password disallow lists, which prevent users from choosing known weak or compromised

passwords. This approach is becoming an industry best practice furthered in part by NIST, which in their latest iteration of SP 800-63 recommends the use of password disallow lists.

There is limited research on the effectiveness of password disallow lists at preventing account compromise. A recent study noted that users who were prevented from using a disallowed password ultimately created passwords that were significantly easier to guess, partly because they modified the disallowed password and partly because they chose passwords with lower entropy. In order for password disallow lists to be truly effective, researchers recommended that checks should strip all candidate passwords of digits and symbols to perform a case-insensitive search.[38] This will prevent users from choosing a modified version of the disallowed password.

Password rate limiting

Most passwords can be cracked with fewer than ten guessing attempts.[39] This is partly because users reuse passwords across multiple websites and partly due to use of minimally modified passwords across multiple accounts. Attackers' intuitive understanding of this has led to an increase in credential stuffing attacks, where attackers often use credentials disclosed in a data breach.[40] Yet, these attackers must attempt a certain number of tries on average to get a hit or a successful login.

To mitigate this risk, designers can limit the number of times a user can attempt to login. There are three common techniques to implement this kind of rate limiting.[41] First, users may be locked out of the account for a short time window after a certain number of login attempts. Attackers can use this to deploy a denial of service (DoS) attack against targeted accounts or services. Second, the user may be locked out of the account until they take some kind of 'unlocking' action. For example, the user may have to call a customer support desk to get the account unlocked or they may have to answer security questions on a web portal. The former can become expensive over time, while the latter may have an impact on security.[42] A third option is to have users solve CAPTCHAs, the limitations of which are well documented.[43] Most websites use a combination of the three approaches, along with other options such as asking for a second factor for verification.[44]

PASSWORDS++

As noted in the previous section, text-based passwords have certain underlying problems that have remained difficult to solve. Too many passwords can be difficult to remember, forcing users to adopt compensating strategies such as password reuse. Furthermore, it is difficult for users to construct passwords that are difficult for a computer to guess. Consequently, there has been an effort to develop alternatives that are both more usable and more secure.

Graphical passwords

One approach that targets usability is graphical passwords that typically refer to three distinct approaches.[45] The first requires users to choose between a password by selecting a subset of images, eg passfaces. A second approach requires the user to draw a unique character on a static image, eg passpoints. A third approach requires users to draw a gesture using a set of predefined points, eg android unlock pattern. In the past years, researchers have proposed authentication methods that use pictures as passwords. This type of authentication bases its reasoning on the fact that humans remember pictures better than text.

For instance, passfaces has a login failure rate of less than a third compared to that for text-based passwords.[46] Despite the additional usability advantages of easy recall, there are some concerns about guessability of graphical

passwords. A study on android unlock pattern, for example, notes high bias in the pattern selection process with users starting from the upper left-hand corner and simply drawing straight lines.[47] Graphical passwords then do not appear to be a significant improvement over text-based passwords, as they merely target the amount of memory required to store a password and not other concerns such as entropy.

Passphrases

In contrast to graphical passwords, passphrases target both memorability and entropy of passwords. Passphrases are space-delimited sequences of natural language words typically longer than ordinary passwords. Passphrases are created either through a system (randomly generated from a dictionary of words) or users. Study shows that the rate of memorability of passwords and passphrases are similar and were written down by a majority of participants in the study.[48] With respect to the performance, this study shows that it takes a longer time to enter passphrases than passwords (mainly because of greater numbers of words in each passphrase).

While there is a theoretical argument to made for additional security offered by passphrases, it is unclear whether passphrases can be much more secure than passwords. System-assigned passphrases may not provide advantage over system-assigned passwords of similar entropy.[49] User-chosen passphrases may be predictable based on word association or grammar.[50] Given that passphrases may both take longer to enter and incur typing mistakes, it is unclear whether they provide any overall advantage over passwords.

AUTHENTICATION BEYOND PASSWORDS

Passwords and adjacent solutions, eg graphical passwords, have inherent challenges that constrain their security guarantees. To mitigate the corresponding risk, the proposed solution has been to add additional authentication requirements. In this section we provide an overview of authentication mechanisms that complement traditional passwords.

What you have and 2FA

2FA is a method to enhance the resilience of password-based authentication by requiring users to provide more than one authentication factor. In 2FA, users are authenticated by verifying something that users know (password) and additionally a second factor other than what they know. An example of the second factor is repeating back something (eg a code generated by a security token) that was sent to them through an out-of-band mechanism. Hardware tokens, software tokens (codes generated by a dedicated smartphone app) and SMS are the most common practised forms of the second factor.

There have been several research efforts in the recent years to evaluate the usability and security of the 2FA solutions. A study on 2FA found that users' perceptions of the usability of 2FA is often correlated with their individual characteristics (eg age, gender, background), rather than with the actual technology or the context/motivation in which it is used. If the second factor is through a mobile device, larger touch areas for authentication are recommended. This study shows that when users have the choice to opt in, adoption rates will likely depend on 2FA usability.[51]

Another study that examined Yubico Security Keys, a 2FA hardware token implementing fast identity online (FIDO), found users' feedback consideration, clearer configuration instructions and communicating benefits significantly increased the usability. The same study concludes that even the best-designed hardware will not be used if the benefits are not apparent.[52] There is a different kind of

second factor offered by some banks and credit card companies. Every time there is a transaction on a customer's account, the responsible institution either sends a text message or an email to the customer in question. If the transaction is fraudulent the customer can inform the institution to get their money refunded. The benefits of opting into this system are obvious to the customer, who can easily monitor their account.

What you are and biometrics

Biometric authentication consists of determining the identity of a person based on a set of recognisable and verifiable data, which are specific to a person or biologically unique to individuals. Biometrics-based user authentication systems are becoming increasingly popular, compared to traditional authentication methods that are based on secrets (something users know). The problem with traditional methods is that they cannot discriminate between an impostor who fraudulently gains access and a legitimate user. Additionally, biometrics-based authentication methods are more convenient because users do not have to memorise passwords and/or follow policies around them such as password complexity and password rotation.

In spite of the many advantages of biometrics-based authentication methods, they also come with multiple disadvantages ranging from security risks (eg synthetic fingerprints, face, etc.) to implementation challenges (eg large amount of data for face recognition). Therefore, in case of using such methods, we should find the trade-off between usability and security. There are two categories of biometric authentication mechanisms: 1) physiological; and 2) behavioural.

Physiological biometrics

Physiologic biometrics rely on the physical measurements of the human body. Examples include fingerprint-based authentication, face recognition, eye iris and hand form. The recognition systems based on physiological features have high accuracy compared to other methods. In addition, physiological characteristics of the human body generally do not change over time. This increases the reliability of physiological-based authentication methods.

There are exceptional cases, however, where physiological features can be affected and have an impact on the accuracy of these methods. For example, fingerprints of people working in chemical industry and iris of diabetes patients can degrade over time, making these methods less reliable for impacted subgroups.[53]

Behavioural biometrics

Behavioural-based biometric authentication is based on the differences in how distinct users perform a task. Examples include speech recognition, signature dynamics (eg handwriting) and typing patterns. In comparison to physiological-based methods, this category has a lower level of accuracy and reliability. The variance in behaviours may not be adequate, affecting accuracy. Simultaneously, behavioural characteristics may change over time and thus limit reliability. For example, a person's voice may change over time due to aging or emotions.[54]

There are several authentication solutions enabled by biometric measurements; however, different measurements do not have the same level of reliability. Researchers at IBM conducted a study to measure the usability of authentication using face, voice, gesture, face/voice, gesture/voice and password. They found that each biometric modality has unique strengths and weaknesses. For example, face and voice were fast but universally usable; voice was the least reliable among all; password and gesture had the least rate of enrol failure, while voice and face had the highest. The same study

also shows that the combined method had the least performance among all.[55]

Challenges

Some of the challenges of biometric-based authentication solutions include implementation complications, risks around some of the methods and usability of these biometric authentication methods.

1. *Data volume*: Unlike secret-based authentication methods such as passwords, biometric-based authentication methods often require more storage space to record biometric characteristics. For example, in the case of password-based authentication, username and password combined can be a length of 30–60 characters, whereas biometric objects can occupy significantly more space. Additionally, applying hashing as a solution to this can raise other issues such as the need for more computational resources;
2. *Inaccurate matching*: In password-based authentication methods, a unique hash of the password is stored on the back-end systems, so that when users enter their passwords, the authentication systems simply compares the hash of the entered password and the stored hash on the remote server and if they match the user is authenticated. In biometrics-based authentication, however, user entries (eg fingerprint scans) may differ over time and scanners may not accurately capture data. Therefore, this increases the rate of authentication failure;
3. *Data breaches*: The uniqueness of biometrics used for authentication can be an advantage and at the same time a disadvantage. It is an advantage because it can help to identify users more accurately (unique to individuals); however, unlike passwords that we can easily change in case of a data breach, the problem with biometrics is that they are unchangeable during lifetime. That means that if they are breached, the consequences are far greater than username and password breaches.

Continuous authentication

Users have different styles and ways of typing and pressing keys on the keyboards. Keystroke dynamics is biometric technique that aims to identify users through verifying the way of typing the credentials, habitual rhythm patterns and timing of keystrokes. In some implementations, this technique goes beyond a one-time authentication and monitors changes in the patterns to secure the session after its opening by detecting anomalies. This is called continuous authentication. There are multiple challenges that this technique faces. The performance of this solution can be affected by the user's feeling and mood at the moment of typing. Additionally, there are multiple security concerns such as zero-effort attacks in which attackers impersonate users with weak patterns. Studies show, however, that the usability of this solution is highly accepted by users.[56]

Risk-based or context-based authentication

Risk-based authentication (RBA) is a new way of authenticating users enabled by artificial intelligence (AI). In this solution, the system creates a behavioural model (profile) for each user. Examples of factors that are included in each behavioural model are the times they access their accounts, devices they use to login, resources or places they access on a corporate network, and even their clicking and typing patterns. Any deviation from the behavioural model signals the system to require users to prove their identity through an authentication method, biometric verification, or receiving approval from network administrators.[57,58]

RBA takes multiple factors into consideration when determining whether a user's identity is authentic or not. Examples

include geo-location, IP, identification time, device identifiers (device profiling), behaviour profile (eg user activities, resources they access, behaviour patterns, etc.). In other words, in RBA solutions, in addition to focusing on what users know and have, they also focus on what users do. The RBA system assigns a risk level to each user after analysing users' profiles. Depending on the risk level, the system takes different paths to verify each user's identity. For example, users with low risk are authenticated transparently while the high-risk ones may be required to provide an additional proof of identity.

Due to the high accuracy and reliability, multiple enterprises and corporations such as RSA and Mastercard have been implementing this solution. RSA has incorporated an RBA component into its authenticator (RSA Authentication Manager 8.0 and later). Risk-based authentication may suffer from the typical limitations of machine learning based systems, ie false positives and false negatives. The impact of these can be limited by appropriate feature selection and modelling. For example, for it might be useful to create separate behavioural models for employees who have access to different data with distinct levels of sensitivity. The behaviour of some employees is a lot more constrained and may be easier to predict than others. For example, some employees may work at off hours and weekends and may be more difficult to model than some with very specific work hours or work locations.

Re-authentication

User re-authentication seeks to guarantee that the current user is the authorised user (the user who authenticated themself at the beginning of the session). This technique is used to protect users against unauthorised access to an account, either through the initial authentication such as stolen passwords or simply by exploiting an open session before logging out of the session. Traditional authentication mechanisms ask users to periodically re-authenticate themselves — for example, requiring users to authenticate after a fixed number of hours or days.[59] Studies show, however, that frequent authentication can be disruptive to users, expensive and often ineffective. Such technique places the burden of information security on the end user and this may result in attacks such as authentication reply attacks. To reduce disruption, researchers recommend using behavioural-based (risk-based) solutions in which re-authentication is suggested in case of anomalous activities and behaviours on the end user side.[60,61]

CONCLUSION

Authentication is the security control that underlies a layered defence strategy grounded in zero trust. Passwords — the most common authentication technology — show that the challenge is not just technical but also one of usability, adoption, etc. In this paper we reviewed existing research on passwords both from an academic and practitioner perspective. We note that many ideas, such as password complexity policies that intuitively offer security improvement, may result in unintended compensating behaviours that instead harm security. Furthermore, even promising solutions must be carefully implemented to be effective — for example, passwords managers only reduce risk when used in combination with password generators. Similarly, password disallowed lists are useful only if they consider simple modifications to prohibited passwords.

Based on the review of existing research, we make the following three high-level recommendations for designers to consider when implementing password-based authentication:

1. *Do no harm*: The first principle for designing authentication should be to do no harm. Any decision to add friction to the process should be supported by evidence. It is easy to mistake a common

practice for a best practice. Well-intentioned efforts to improve password security may result in compensating behaviours from users that in turn increase risk. It is particularly important to deprecate known bad strategies such as password expiration;

2. *Minimise user effort*: Any design should aim to minimise user effort. For example, SSO may reduce the number of passwords that a user has to manage. This can be combined with password disallow lists, password managers and password generators to further reduce user effort;
3. *Adopt compensating controls*: No strategy can mitigate the underlying limitations of passwords as a technology. It is important, then, to combine passwords with compensating controls through additional authentication factors, such as 2FA. By employing additional strategies such as risk-based authentication, designers can ensure that this additional friction is only incurred when needed.

References

1. Inglesant, P. G. and Sasse, M. A. (April 2010), 'The true cost of unusable password policies: Password use in the wild', Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 383–392.
2. Ion, I., Reeder, R. and Consolvo, S. (July. 2015), '"... no one can hack my mind": Comparing Expert and Non-Expert Security Practices', 11th Symposium on Usable Privacy and Security, pp. 327–346.
3. Carr, M. and Shahandashti, S. F. (March 2020), 'Revisiting Security Vulnerabilities in Commercial Password Managers', Cornell University, available at https://arxiv.org/abs/2003.01985 (accessed 4th November, 2020).
4. Li, Z., He, W., Akhawe, D. and Song, D. (2014), 'The emperor's new password manager: Security analysis of web-based password managers', 23rd USENIX Security Symposium Security 14, pp. 465–479.
5. Singh, R., Agarwal, A., Singh, M., Nagpal, S. and Vatsa, M. (February 2020), 'On the robustness of face recognition algorithms against attacks and bias', available at https://arxiv.org/pdf/2002.02942.pdf (accessed 4th November, 2020).
6. Bontrager, P., Roy, A., Togelius, J., Memon, N. and Ross, A. (October 2018), 'Deepmasterprints: Generating masterprints for dictionary attacks via latent variable evolution', IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS), pp. 1–9.
7. Rosencrance, L. (August 2003), 'Survey: Insecure passwords can be costly for companies', ComputerWorld, available at https://www.computerworld.com/article/2571799/survey--insecure-passwords-can-be-costly-for-companies.html (accessed 4th November, 2020).
8. Password Random (2020), 'Most popular password list', available at https://www.passwordrandom.com/most-popular-passwords (accessed 4th November, 2020).
9. AlSabah, M., Oligeri, G. and Riley, R. (August 2018), 'Your culture is in your password: An analysis of a demographically-diverse password dataset', *Computers and Security*, Vol. 77, pp. 427–441.
10. Marquardson, J. (2012), 'Password policy effects on entropy and recall: Research in progress', *AMCIS*.
11. Florêncio, D. and Herley, C. (July 2020), 'Where do security policies come from?', Proceedings of the Sixth Symposium on Usable Privacy and Security, pp. 1–14.
12. Komanduri, S., Shay, R., Kelley, P. G., Mazurek, M. L., Bauer, L., Christin, N., Cranor, L. F. and Egelman, S. (May 2011), 'Of passwords and people: Measuring the effect of password-composition policies', Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 2595–2604.
13. Burr, W., Dodson, D. and Polk W. (June 2004), 'Electronic authentication guideline', National Institute of Standards and Technology.
14. Wash, R., Rader, E., Berman, R. and Wellmer, Z. (2016), 'Understanding password choices: How frequently entered passwords are re-used across websites', 12th Symposium on Usable Privacy and Security, pp. 175–188.
15. Zhang, Y., Monrose, F. and Reiter, M. K. (October 2010), 'The security of modern password expiration: An algorithmic framework and empirical analysis', Proceedings of the 17th ACM Conference on Computer and Communications Security, pp. 176–186.
16. Egelman, S., Sotirakopoulos, A., Muslukhov, I., Beznosov, K. and Herley, C. (April 2013), 'Does my password go up to eleven? The impact of password meters on password selection', Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 2379–2388.
17. Ur, B., Kelley, P. G., Komanduri, S., Lee, J., Maass, M., Mazurek, M. L., Passaro, T., Shay, R., Vidas, T., Bauer, L. and Christin, N. (2012), 'How does your password measure up? The effect of strength meters on password creation', presented as part of the 21st Security Symposium, pp. 65–80.
18. Forget, A., Chiasson, S., Van Oorschot, P. C. and Biddle, R. (July 2018), 'Improving text passwords through persuasion', Proceedings of the 4th Symposium on Usable Privacy and Security, pp. 1–12.

19. Florencio, D. and Herley, C. (May 2007), 'A large-scale study of web password habits', Proceedings of the 16th International Conference on World Wide Web, pp. 657–666.
20. Summers, W. C. and Bosworth, E. (January 2004), 'Password policy: The good, the bad, and the ugly', Proceedings of the Winter International Symposium on Information and Communication Technologies, pp. 1–6.
21. Li, Z., He, W., Akhawe, D. and Song, D. (2014), 'The emperor's new password manager: Security analysis of web-based password managers', 23rd Security Symposium, pp. 465–479.
22. Seiler-Hwang, S., Arias-Cabarcos, P., Marín, A., Almenares, F., Díaz-Sánchez, D. and Becker, C. (November 2019), "'I don't see why I would ever want to use it": Analyzing the Usability of Popular Smartphone Password Managers', Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security, pp. 1937–1953.
23. Zhang, S. A., Pearman, S., Bauer, L. and Christin, N. (2019), 'Why people (don't) use password managers effectively', 15th Symposium on Usable Privacy and Security.
24. Fagan, M., Albayram, Y., Khan, M. M. and Buck, R. (December 2017), 'An investigation into users' considerations towards using password managers', *Human-centric Computing and Information Sciences*, Vol. 7, No. 1, p. 12.
25. Lyastani, S. G., Schilling, M., Fahl, S., Backes, M. and Bugiel, S. (September 2018), 'Better Managed than Memorized? Studying the Impact of Managers on Password Strength and Reuse', 27th Security Symposium, pp. 203–220.
26. Pashalidis, A. and Mitchell, C. J. (July 2003), 'A taxonomy of single sign-on systems', Australasian Conference on Information Security and Privacy, Springer, Berlin, Heidelberg, pp. 249–264.
27. Li, W. and Mitchell, C. J. (October 2014), 'Security issues in OAuth 2.0 SSO implementations', International Conference on Information Security, Springer, Cham, pp. 529–541.
28. Mainka, C., Mladenov, V., Schwenk, J. and Wich, T. (April 2017), 'SoK: Single sign-on security—an evaluation of openID connect', IEEE European Symposium on Security and Privacy, pp. 251–266.
29. Yue, C. (2013), 'The devil is phishing: Rethinking web single sign-on systems security', 6th Workshop on Large-Scale Exploits and Emergent Threats.
30. Chun, K. L. and Katuk, N. (September 2014), 'A usability study of social media credentials as a single-sign-on mechanism: Student access to online teaching materials', *Journal of Industrial and Intelligent Information*, Vol. 2, No. 3.
31. Linden, M. and Vilpola, I. (April 2005), 'An empirical study on the usability of logout in a single sign-on system', International Conference on Information Security Practice and Experience, Springer, Berlin, Heidelberg, pp. 243–254.
32. Freeman, B. (July 2008), 'Yahoo! OpenID: One key, many doors', OpenID User Experience Research.
33. Waters, S. (2012), 'Web-based single sign-on: An examination of security and usability', *Computer Science*.
34. Sun, S. T., Boshmaf, Y., Hawkey, K. and Beznosov, K. (September 2010), 'A billion keys, but few locks: The crisis of web single sign-on', Proceedings of the 2010 New Security Paradigms Workshop, pp. 61–72.
35. *Ibid.*, note 25.
36. Seitz, T., Hartmann, M., Pfab, J. and Souque, S. (May 2017), 'Do differences in password policies prevent password reuse?', Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems, pp. 2056–2063.
37. *Ibid.*, note 14.
38. Habib, H., Colnago, J., Melicher, W., Ur, B., Segreti, S., Bauer, L., Christin, N. and Cranor, L. (2017), 'Password creation in the presence of blacklists', Proceedings of Workshop on Usable Security (USEC), p. 50.
39. Wang, C., Jan, S. T., Hu, H., Bossart, D. and Wang, G. (March 2018), 'The next domino to fall: Empirical analysis of user passwords across online services', Proceedings of the 8th ACM Conference on Data and Application Security and Privacy, pp. 196–203.
40. Bulakh, V., Kaizer, A. J. and Gupta, M. (October 2017), 'All Your Accounts Are Belong to Us', International Conference on Security and Privacy in Communication Systems, Springer, Cham, pp. 245–269.
41. Florêncio, D., Herley, C. and Van Oorschot, P. C. (2014), 'An administrator's guide to internet password research', 28th Large Installation System Administration Conference (LISA14), pp. 44–61.
42. Bonneau, J., Bursztein, E., Caron, I., Jackson, R. and Williamson, M. (May 2018), 'Secrets, lies, and account recovery: Lessons from the use of personal knowledge questions at Google', Proceedings of the 24th International Conference on World Wide Web, pp. 141–150.
43. Raman, J., Umapathy, K. and Huang, H. (2018), 'Security and User Experience: A Holistic Model for Captcha Usability Issues', SAIS 2018 Proceedings, p. 27.
44. Golla, M., Schnitzler, T. and Dürmuth, M. (2016), "'Will Any Password Do?": Exploring Rate-Limiting on the Web', Way Workshop, available at https://wayworkshop.org/2018/papers/way2018-golla.pdf (accessed 4th November, 2020).
45. Suo, X., Zhu, Y. and Owen, G. S. (December 2005), 'Graphical passwords: A survey', 21st Annual Computer Security Applications Conference, p. 10.
46. Brostoff, S. and Sasse, M. A. (2000), 'Are Passfaces More Usable than Passwords? A Field Trial Investigation', in *People and Computers XIV — Usability or Else!*, Springer, London, pp. 405–424.
47. Uellenbeck, S., Dürmuth, M., Wolf, C. and Holz, T. (November 2013), 'Quantifying the security of graphical passwords: The case of android unlock patterns', Proceedings of the 2013 ACM SIGSAC Conference on Computer and Communications Security, pp. 161–172.
48. Bonneau, J. and Shutova, E. (February 2012), 'Linguistic properties of multi-word passphrases',

International Conference on Financial Cryptography and Data Security, Springer, Berlin, Heidelberg, pp. 1–12.
49. Shay, R., Kelley, P. G., Komanduri, S., Mazurek, M. L., Ur, B., Vidas, T., Bauer, L., Christin, N. and Cranor, L. F. (July 2012), 'Correct horse battery staple: Exploring the usability of system-assigned passphrases', Proceedings of the 8th Symposium on Usable Privacy and Security, pp. 1–20.
50. Rao, A., Jha, B. and Kini, G. (February 2013), 'Effect of grammar on security of long passwords', Proceedings of the 3rd ACM Conference on Data and Application Security and Privacy, pp. 317–324.
51. De Cristofaro, E., Du, H., Freudiger, J. and Norcie, G. (September 2013), 'A comparative usability study of two-factor authentication', Cornell University, available at https://arxiv.org/abs/1309.5344 (accessed 4th November, 2020).
52. Das, S., Dingman, A. and Camp, L. J. (February 2018), 'Why Johnny doesn't use two factor a two-phase usability study of the FIDO U2F security key', International Conference on Financial Cryptography and Data Security, Springer, Berlin, Heidelberg, pp. 160–179.
53. El-Sallam, A., Sohel, F. and Bennamoun, M. (June 2011), 'Robust pose invariant shape-based hand recognition', 6th IEEE Conference on Industrial Electronics and Applications, pp. 281–286.
54. Hoang, T., Nguyen, T. D., Luong, C., Do, S. and Choi, D. (June 2013), 'Adaptive Cross-Device Gait Recognition Using a Mobile Accelerometer', *Journal of Information Processing Systems*, Vol. 9, No. 2, p. 333.
55. Trewin, S., Swart, C., Koved, L., Martino, J., Singh, K. and Ben-David, S. (December 2012), 'Biometric authentication on a mobile device: A study of user effort, error and task disruption', Proceedings of the 28th Annual Computer Security Applications Conference, pp. 159–168.
56. Giot, R. and Rosenberger, C. (January 2012), 'A new soft biometric approach for keystroke dynamics based on gender recognition', *International Journal of Information Technology and Management*, Vol. 11, Nos. 1–2, pp. 35–49.
57. *Ibid.*, note 38.
58. Jagadeesan, H. and Hsiao, M. S. (September 2009), 'A novel approach to design of user re-authentication systems', IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems, pp. 1–6.
59. Shen, C., Cai. Z., Guan., X., Du, Y. and Maxion, R. A. (October 2012), 'User authentication through mouse dynamics', *IEEE Transactions on Information Forensics and Security*, Vol. 8, No. 1, pp. 16–30.
60. *Ibid.*, note 58.
61. *Ibid.*, note 59.