Do Differences in Password Policies Prevent Password Reuse?

Tobias Seitz
Manuel Hartmann
Jakob Pfab
LMU Munich
Munich, Germany
tobias.seitz@ifi.lmu.de
hartmannm@cip.ifi.lmu.de
j.pfab@campus.lmu.de

Samuel Souque
École Nationale Supérieure d'Informatique pour l'Industrie et l'Entreprise
Évry, France
samuel.souque@ensiie.fr

CHI'17 Extended Abstracts, May 06-11 2017, Denver, CO, USA.
ACM 978-1-4503-4656-6/17/05.
http://dx.doi.org/10.1145/3027063.3053100

Abstract
Password policies were originally designed to make users pick stronger passwords. However, research has shown that they often fail to achieve this goal. In a systematic audit of the top 100 web sites in Germany, we explore if diversity in current real-world password policies prevents password reuse. We found that this is not the case: we are the first to show that a single password could hypothetically fulfill 99% of the policies under consideration. This is especially problematic because password reuse exposes users to similar risks as weak passwords. We thus propose a new approach for policies that focuses on password reuse and respects other websites to determine if a password should be accepted. This re-design takes current user behavior into account and potentially boosts the usability and security of password-based authentication.

Author Keywords
passwords; password-composition policies; usable security; authentication

ACM Classification Keywords
K.6.5 [Authentication]: Security and Protection

Introduction
While many alternatives exist, passwords are still the standard method for authentication on the web and will remain

so for the foreseeable future [3]. Since user-name and password based authentication is easy to implement, and since no alternative can satisfactorily replace them yet, the number of websites offering this authentication method is ever-growing. However, this growth in user accounts comes at a high price in terms of usability and security for the users. To manage so many accounts, users develop a variety of coping strategies. Most commonly, users choose secrets they can easily remember to remain independent of external management tools like written notes or password management software [18].

While this behavior is convenient, it also puts users at risk. Sophisticated attackers can quickly guess common passwords and their modifications [1, 12, 20]. Service providers try to mitigate the issue by enforcing a list of requirements upon the users' passwords. Those password composition policies make users come up with new passwords in case their preferred option is disqualified. Consequently, users continue to create new passwords until they have a sufficient set to fulfill most policies [8]. Florêncio and Herley have shown that a user maintains about five to seven passwords [6]. According to their large scale study, this small set of passwords is used to log into many different websites, thus reuse was identified as another coping strategy.

Since it is virtually impossible to remember a strong unique password for every account, password reuse is reasonable from a usability perspective [7]. However, if users are unaware of the websites that share the same credentials, password reuse can put them through much more trouble than weak passwords alone. If the password database of one service leaks, the affected service can prompt the users to change the corresponding password to secure the account after the leak is detected. All other accounts sharing the same credentials, however, are still at risk [9]. To secure those, users will have to first identify them and manually change the passwords which is a tedious process that many are unwilling to undertake.

In our work, we focus on the effect of password policies on password reuse. The research community has not reached consensus on the "best" policy to recommend to service providers. In some cases, the recommendations are even contradictory (cf. [2] and [14]). At this point, there is a shortcoming on specific data about actual password policies in the wild.

To obtain an overview about password policies in practice, this paper investigates the policies of the top 100 German websites according to Alexa.com[1]. Afterwards, we evaluate their effectiveness against password reuse and found that it was easily possible to come up with passwords that could be shared among 99 percent of the tested sites. We conclude that the diversity in policies is currently unable to mitigate risks caused by password reuse. We contribute a methodology to test real-world policies and results from its application to a large number of websites. Moreover, we present an approach to help users minimize risks through password reuse. Our concept does not require to create a unique password for each account, but prevents sharing secrets across different website categories.

Related Work

A lot of research around passwords has been conducted in the past few decades. The community explained why people tend to use rather poor passwords and fail to create good ones [3], analyzed password reuse [5] and showed how users' behavior would have to change to become well-protected [13].

[1] http://www.alexa.com/topsites/countries/DE, as of 09/01/2016

Password (-part)	Test case
aaa	short
123456	digits
123456789a	alphanumeric
a123456789	digit positions
password	dictionary
password with space	spaces
Aa1! "#$%&'()*+,-./:	ASCII / complex
Aa1ÁÂÄÅÆÇ	extended ASCII
ÈÉÌÍÎÑÒÓÙÚÜÝ	
œåß´ç®ƒ√†®©∂∫¥	Unicode

Table 1: Excerpt from our password set (shortened for layout purposes). We identified 46 passwords that were suitable to audit the policy of a website.

Part of the research focused on the analysis of real-world passwords and their weaknesses. Weir et al. used millions of passwords from password database breaches to estimate their strength [20]. They observed that leaked passwords would not even withstand online attacks, in which attackers are usually limited in the amount of guesses they can make at a time [3]. Bonneau et al. also discussed this and put the predictability of user chosen passwords to the test [1]. The passwords were created under real-world circumstances, i.e. with real-world policies.

Inglesant and Sasse question the merit of common password policies [11]. They observed that users struggle with them and discuss the disadvantages of ignoring HCI design principles when administrators impose unusable password policies [11]. Supporting their argument, it was shown numerous times that stricter policies challenge the users a lot, often without the desired security benefits [13, 15, 16].

Wang and Wang evaluated the effectiveness of 50 policies against online guessing [19]. They found a large spectrum of policies at high-traffic sites in the US and China, and again showed that those policies often fail to make users create reasonably strong passwords. In contrast to their work, we focus on examining password reuse and give concrete recommendations for a policy re-design.

Method: Testing Password Policies

To scrutinize real world password policies, we examined the top 100 sites on the web for Germany, according to the Alexa ranking in May 2016. Our process to test the password policies consisted of two stages. This is required because, the information about a website's password policy is often concealed and not easily accessible. Thus, we require a suitable set of passwords to see whether a website accepts them or not.

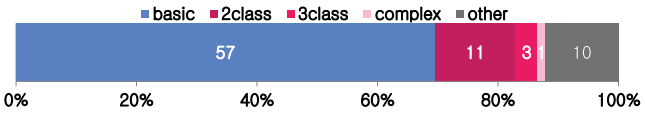

Figure 1: Distribution of complexity classes. Most of the examined web sites (65%) require only a minimum length. 'Other' encapsulates all sites with additional, sometimes unique restrictions and requirements.

Stage 1: Identification of Passwords

We started out with a set of 15 passwords that we constructed from guidelines and suggestions from the literature (e.g. [17]. We tried to create accounts wherever possible with this list of passwords. In case the registration failed, we identified the source and tried to modify the password to fulfill the policy. We added the modified password to our list.

Moreover, some sites specifically indicated black- and white-listed symbols or rules like maximum length. We added new passwords to our list that violated the rules. This allows us to later check if they would also violate the policies of other web sites. In this first stage, a set of **46 passwords** was identified (cf. Table 1)

Stage 2: Policy Evaluation with Test Set

We proceeded by re-checking account creation with the password armory to get in-depth results about every examined password policy. We structured the collected data by the following criteria:

- minimal and maximal password lengths
- mandatory, allowed, or forbidden character classes
- complexity class proposed by Shay et al. [16]
- pro-active dictionary and common passwords check
- additional comments about special password rules

Limitations

Our findings about policies are limited by three critical aspects. First, we could not create accounts on sites of banks, telecommunication or pay-tv providers, because they require offline registration. However those only accounted for a small number of web sites. Second, the longest password in our test set consisted of 246 characters from a variety of classes. If services accepted it, we conclude that there was no length restriction, but the length restriction might just be larger. However, we expect only expert users will use such long credentials and we can assume that those handle reuse more appropriately. Last, we did not include emoji in the test set, because they caused problems with masked password fields. They are usually encoded as two characters and distort fulfillment of length and character set requirements.

Results

It was possible for us to create accounts on 83 out of 100 sites on-line. The 17 missing sites required offline registration or do not offer public registration at all. We found large consistencies, so that it was possible to come up with a password that fulfills 82 (~98.88%) of the tested policies.

Complexity Requirements

A well-established terminology to describe policies was proposed by Shay et al. They characterize the resulting password by describing the number of character classes or complexity. For instance, the *3class12* policy requires at least *three* different character *classes* in passwords of minimum *length twelve*. A *comp8*-policy demands at least eight characters, a lowercase and an uppercase letter, a digit, a symbol, and muss pass a dictionary check – thus the result is considered a *complex* password. A *2word16*-password must be at least 16 characters long and consist of at least two words separated by a non-letter sequence.

Our research showed that all of these password complexity requirements are actually used in the wild albeit not consistently. The highest complexity in terms of character diversification was a comp6 policy implemented on lidl.de – a German super market franchise. Also, 69.5% of websites use a basic policy that has length as a sole requirement (see Figure 1). The second largest group demanded at least 2 different character classes (13.3%). The 'other' label sums up the remaining complexity classes. They cover special cases of password rules, which do not necessarily prevent password reuse. For instance, paypal.com prevents users from including a character three times in a row, while this is okay for most other sites. Bahn.de requires three *different* characters, and disallows using the first name, last name or user name.

Password Length

Most sites require a minimal length of six to eight characters ($M = 6.3, SD = 1.9$, see Figure 2). 43 sites (51.8%) allowed passwords without a fixed maximum length, while the maximum length for the remaining 40 sites was 43 characters on average ($SD = 32$). It is also noteworthy that six websites (7%) allow passwords with a minimal length of one character, e.g. heise.de or chip.de that have an expectedly technical audience. It was surprising that also a website with high global traffic – wikipedia.org – implements a minimal length requirement of one character [2].

The upper bounds of minimum length and the lower bounds of maximum length rules do not intersect, which means that there is a 'golden reusable password length'. 10 websites disallow passwords with up to 20 characters. The maximal password length for ikea.com (10 characters) is only one character longer than the minimal required length of yahoo.com (9 characters). In other words, a single password

[2] re-checked on January 10th 2017

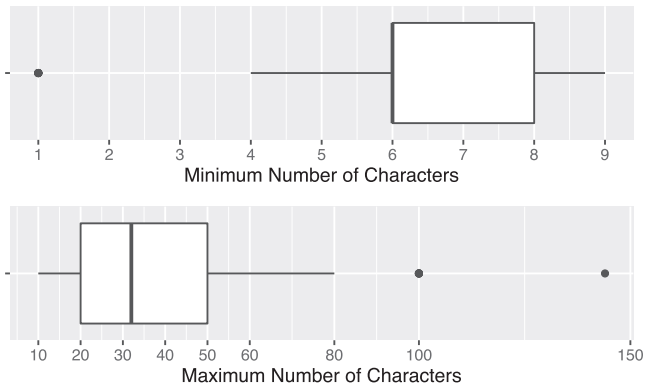

Figure 2: Distribution of password length rules. We excluded maximum lengths beyond 245 characters.

with nine or ten characters could be reused across all sites, if length were the only requirement. Thus, neither complexity nor length requirements prevent password *reuse*.

Character Sets

Consequently, only the diverse range of character sets can be a limiting factor regarding non-modified reuse of a single password. All tested sites allow lowercase letters, uppercase letters, and digits. However, non-alphanumeric characters have a large variety of limitations. As shown in Table 2, some websites like live.com allow only a very small subset of non-alphanumeric characters. Some websites allow ASCII-printable characters only (google.com), while others accepted extended ASCII characters or even Unicode characters in passwords.

Among websites that do not reveal their accepted character set, some seem to keep an intransparent blacklist of specific forbidden symbols and characters. Spaces and line-breaks are often banned. For instance, netflix.com prevents the use of the tilde character – '~' (Table 2, bottom). An outstanding case is lidl.de that gives a list of allowed characters which must be included at least once in the password to pass pro-active checks. Pornhub whitelists only two symbols that are prevented on most other sites. This leads to mutually exclusive policies, which ultimately prevent a 100% successful reuse rate.

Web site	Explicitly whitelisted symbols
ebay.de	!@#$%^*-_+=
gmx.net	!#$%&()*+,-./:;<=>?@[\]^_{|}~§ÄäÖöÜüß
t-online.de	!#$%&()*+,-./<=>?@[\]_{|}~
live.com	@#$%^*-_+=
mobile.de	!$%&?-_+#
pornhub.com	/_
chefkoch.de	äöüÄÖÜß, _, -, !?&.
zeit.de	äöüÄÖÜß ,.!?:;#&* ()_+=/<>-
lidl.de	@#$%^&+=.:-!?
Web site	**Explicitly blacklisted symbols**
spiegel.de	'space' 'new line'
welt.de	'space'
netflix.com	~

Table 2: We can see that many sites are mutually exclusive regarding the usage of certain symbols only by looking at a small subset of the tested sites. This list is not comprehensive, but gives a rough picture of current practice. The full data set is available on-line at https://github.com/mimuc/password-policy-dataset

Discussion

There were a number of instances where the policies seem either odd or counterintuitive, because they prevent stronger passwords by disallowing longer passwords containing a more diverse range of characters. Still, some services may benefit from introducing length or character restrictions in

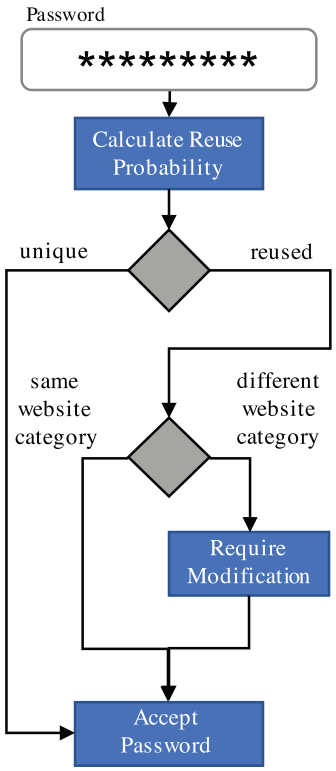

Figure 3: Flow chart of sign-up with dynamic password policy. The central idea is to check if the password is likely used on websites of different categories (e.g. email vs. online shopping). In such a case, the policy requires any kind of modification to the first password choice.

terms of usability. For example, Netflix might disallow the tilde character because on-screen keyboards on TV sets do not necessarily bear this character. On pure web-oriented pages without evident non-browser usage, however, such rules limit the usage of randomly generated passwords. Wikipedia likely allows single-character passwords because no personal information is required during sign-up.

Despite the limitations, we found that policies fail to prevent password reuse, and this is understandable because they were likely never designed with this intention. However, a password that is between 9 and 10 characters long, that is not an English word and that contains only digits, lower- and uppercase ASCII characters is valid on 99% of the sites in our set (for example `DenCHI2017`). These "reuse-requirements" go beyond the NIST guidelines [4]. Thus, if such a password is reused, online attacks against it would likely fail, assuming a common threshold of 1 million guesses [21]. However, such passwords become a great risk if they are reused across the board. Users may be "lucky" and their preferred password fits just these requirements. Then it is unlikely that they are forced to modify it when creating new accounts. Consequently the likelihood of compromising multiple accounts at once increases with every registration. We thus propose a different approach to password policies that can help overcome this risk.

Proposal: Dynamic Password Policies

We propose to adjust the password policy if a system detects a password that could be widely used (see Figure 3). For example, from our audit we know that "*opensesame!*" would be valid on 84% of the websites under consideration. The high percentage indicates that the password could have been already reused too often. The dynamic policy then decides in a next step if the user needs to modify their first-choice password.

To make the policy less restrictive and more usable, the policy takes the website category into consideration. For example, many users keep password portfolios with semantic cues, e.g. one password for social networking, a different password for shopping and another one for one-time accounts. If the policy on a shopping site detects that the entered password is valid on all social networking sites, it requires a modification. The prediction model for reuse probability should also be informed by additional factors. For instance, the policy can evaluate typing patterns during password selection. These may give additional hints of reuse, e.g. if the user types the password with the same speed as the user-name, this may indicate that they have typed the password often in the past and classify it as reused.

For a real-word implementation of our concept, up-to-date information on the password policies of external sites is necessary. By now, we can already implement such dynamic policies based on our manual evaluation. However, an automated solution is preferable. To achieve this, a site's password policy could be communicated using a markup language as recently proposed by Horsch et al. [10].

Conclusion and Future Work

We presented an audit of current password composition policies on the most visited web sites in Germany. To the best of our knowledge, we were the first to show that it is easily possible to find passwords that are accepted by 99% of these sites. This illustrates that the differences in current policies do not prevent password reuse. However, helping users to come up with smarter reuse strategies is an important topic to reduce both burden and security risks. Our future work will focus on password policies designed around password reuse. A first step in this direction is the proposed dynamic policy concept which we are currently finalizing and evaluating in terms of usability and security.

References

[1] Joseph Bonneau. 2012. The Science of Guessing: Analyzing an Anonymized Corpus of 70 Million Passwords. In *Proceedings - IEEE Symposium on Security and Privacy*. IEEE Comput. Soc, 538–552. DOI: http://dx.doi.org/10.1109/SP.2012.49

[2] Joseph Bonneau. 2016. Deep Dive: EFF's New Wordlists for Random Passphrases. (July 2016). https://www.eff.org/deeplinks/2016/07/new-wordlists-random-passphrases

[3] Joseph Bonneau, Cormac Herley, Paul C. Van Oorschot, and Frank Stajano. 2015. Passwords and the Evolution of Imperfect Authentication. *Commun. ACM* 58, 7 (2015), 78–87. DOI: http://dx.doi.org/10.1145/2699390

[4] William E. Burr, Donna F. Dodson, Elaine M. Newton, Ray A. Perlner, and W. Timothy Polk. 2011. *SP 800-63-1. Electronic Authentication Guideline*. Technical Report December. National Institute of Standards & Technology, Gaithersburg, MD, United States.

[5] Anupam Das, Joseph Bonneau, Matthew Caesar, Nikita Borisov, and Xf Wang. 2014. The Tangled Web of Password Reuse. February (2014), 23–26. http://www.jbonneau.com/doc/DBCBW14-NDSS-tangled

[6] Dinei Florêncio and Cormac Herley. 2007. A Large-Scale Study of Web Password Habits. In *Proceedings of the 16th international conference on World Wide Web (WWW '07)*. ACM, 657–665. DOI: http://dx.doi.org/10.1145/1242572.1242661

[7] Dinei Florêncio, Cormac Herley, and Paul C. Van Oorschot. 2014. Password Portfolios and the Finite-Effort User: Sustainably Managing Large Numbers of Accounts. In *Proceedings of USENIX Security Symposium*. USENIX Association, San Diego, CA, USA, 575–590. https://www.usenix.org/system/files/conference/usenixsecurity14/sec14-paper-florencio.pdf

[8] Shirley Gaw and Edward W. Felten. 2006. Password management strategies for online accounts. In *Proceedings of the second symposium on Usable privacy and security (SOUPS '06)*. ACM, New York, NY, USA, 44–55. DOI: http://dx.doi.org/10.1145/1143120.1143127

[9] Cormac Herley and Wolter Pieters. 2015. If you were attacked, you'd be sorry: Counterfactuals as security arguments. In *Proceedings of the 2015 New Security Paradigms Workshop*. ACM, 112–123.

[10] Moritz Horsch, Mario Schlipf, Stefen Haas, Johannes Braun, and Johannes Buchmann. 2016. Password Policy Markup Language. In *Proceedings of Open Identify Summit*. Gesellschaft für Informatik, Rome, Italy, 135–147.

[11] Philip Inglesant and Martina Angela Sasse. 2010. The True Cost of Unusable Password Policies: Password Use in the Wild. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. 383–392. DOI: http://dx.doi.org/10.1145/1753326.1753384

[12] Michelle L Mazurek, Saranga Komanduri, Timothy Vidas, Lujo Bauer, Nicolas Christin, Lorrie Faith Cranor, Patrick Gage Kelley, Richard Shay, and Blase Ur. 2013. Measuring password guessability for an entire university. In *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security*. ACM, 173–186.

[13] William Melicher, Darya Kurilova, Sean M Segreti, Pranshu Kalvani, Richard Shay, Blase Ur, Lujo Bauer, Nicolas Christin, Lorrie Faith Cranor, and Michelle L Mazurek. 2016. Usability and security of text passwords on mobile devices. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 527–539.

[14] Bruce Schneier. 2014. Choosing Secure Passwords. (2014). https://www.schneier.com/blog/archives/2014/03/choosing_secure_1.html

[15] Richard Shay. 2015. *Creating Usable Policies for Stronger Passwords with MTurk*. Dissertation. Carnegie Mellon University.

[16] Richard Shay, Saranga Komanduri, Adam L Durity, Phillip Seyoung Huh, Michelle L Mazurek, Sean M Segreti, Blase Ur, Lujo Bauer, Nicolas Christin, and Lorrie Faith Cranor. 2014. Can long passwords be secure and usable?. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*. ACM, 2927–2936.

[17] Richard Shay, Saranga Komanduri, Adam L Durity, Phillip Seyoung Huh, Michelle L Mazurek, Sean M Segreti, Blase Ur, Lujo Bauer, Nicolas Christin, and Lorrie Faith Cranor. 2016. Designing Password Policies for Strength and Usability. *ACM Transactions on Information and System Security (TISSEC)* 18, 4 (2016), 13.

[18] Elizabeth Stobert and Robert Biddle. 2014. The Password Life Cycle: User Behaviour in Managing Passwords. In *Proceedings of the 10th Symposium On Usable Privacy and Security (SOUPS '14)*. ACM, New York, NY, USA, 243–255.

[19] Ding Wang and Ping Wang. 2015. The emperor's new password creation policies. In *European Symposium on Research in Computer Security*. Springer, 456–477.

[20] Matt Weir, Sudhir Aggarwal, Michael Collins, and Henry Stern. 2010. Testing metrics for password creation policies by attacking large sets of revealed passwords. In *Proceedings of the 17th ACM conference on Computer and communications security*. ACM, 162–175.

[21] Daniel Lowe Wheeler. 2016. zxcvbn: Low-budget password strength estimation. In *25th USENIX Security Symposium (USENIX Security 16). Austin, TX: USENIX Association*.