Where Do Security Policies Come From?

Dinei Florêncio and Cormac Herley
Microsoft Research
One Microsoft Way
Redmond, WA, USA
dinei@microsoft.com, cormac@microsoft.com

ABSTRACT

We examine the password policies of 75 different websites. Our goal is understand the enormous diversity of requirements: some will accept simple six-character passwords, while others impose rules of great complexity on their users. We compare different features of the sites to find which characteristics are correlated with stronger policies. Our results are surprising: greater security demands do not appear to be a factor. The size of the site, the number of users, the value of the assets protected and the frequency of attacks show no correlation with strength. In fact we find the reverse: some of the largest, most attacked sites with greatest assets allow relatively weak passwords. Instead, we find that those sites that accept advertising, purchase sponsored links and where the user has a choice show strong inverse correlation with strength.

We conclude that the sites with the most restrictive password policies do not have greater security concerns, they are simply better insulated from the consequences of poor usability. Online retailers and sites that sell advertising must compete vigorously for users and traffic. In contrast to government and university sites, poor usability is a luxury they cannot afford. This in turn suggests that much of the extra strength demanded by the more restrictive policies is superfluous: it causes considerable inconvenience for negligible security improvement.

1. INTRODUCTION

Passwords remain the dominant means of authentication to web sites. Different sites have different policies: some insist on very complex passwords, while some allow relatively weak ones. Complexity increases the resistance of the password to brute-force attacks, but reduces usability. Our goal in this paper is to understand why there is so much diversity of requirements. That is, what causes some sites to require very restrictive policies when others clearly manage with less?

We perform a study of password policies at 75 different sites. These include top, high and medium traffic sites, universities, banks, brokerages and government sites. The policies range from single-character unrestricted passwords, to 12-character passwords that must include upper, and lowercase, digits and special characters. The sites also span an enormous range in terms of traffic, number of user accounts and value of resources protected. Each of these policies gives us a data point on the tradeoff between security and usability as decided by different people. We examine several of the factors that might influence the need for greater security to see if there is correlation with enforced password strength.

Our results are somewhat surprising. We find that none of the factors that might require greater security seems a factor. The size of the site, the number of user accounts, the value of the resources protected, and the frequency of non-strength related attacks all correlate very poorly with the strength required by the site. Some of the largest, highest value and most attacked sites on the Internet such as Paypal, Amazon and Fidelity Investments allow relatively weak passwords. We also examine several factors unrelated to security. We find that sites that accept advertising, purchase adwords, have a revenue opportunity per login, or where the user has choice, tend to have less restrictive policies. For example, we find median password policy strength of 31 bits for banks and 19.9 bits for `.com` sites, but 43.7 and 47.6 for `.edu` and `.gov` sites respectively. Our analysis suggests that strong-policy sites do not have greater security needs. Rather, it appears that they are better insulated from the consequences of imposing poor usability decisions on their users. For commercial retailers like Amazon, and advertising supported sites like Facebook, every login event is a revenue opportunity. Anything that interferes with usability affects the business directly. At government sites and universities every lo-

Symposium on Usable Privacy and Security (SOUPS) 2010, July 14–16, 2010, Redmond, WA USA
.

gin event is, at best, neutral, or, at worst, a cost. The consequences of poor usability decisions are less direct. That simple difference in incentives turns out to be a better predictor of password policy than any security requirement. This in turn suggests that some of stronger policies are needlessly complex: they cause considerable inconvenience for negligible security improvement. Why do password policies matter? In 2010 there are approximately 1.7 billion Internet users [1]; in 1990 there were fewer than 3 million [2]. Thus, there are probably about 10 billion password protected accounts in use today, and this number is growing rapidly. Thus, in the tradeoff between security and usability, erring on the side of unnecessarily strong policies causes an enormous usability burden and consumes cognitive effort that could be better spent elsewhere.

2. METHODOLOGY

We have gathered the password policies from the 75 sites listed in Table 7. Our means of selecting sites is as follows. We've chosen sites in several different categories: top, high, and medium traffic sites, banks and brokerages, universities, and government sites. The top, high and medium traffic sites are drawn from particular ranges on the traffic site `www.quantcast.com`. The banks are the top commercial banks and brokerages in the US as ranked by the Federal Financial Institutions Examinations Council FFIEC) [17]. The universities selected are the ten largest universities in the US by student enrolment. The government sites are the ten highest traffic rank sites with top-level domain `.gov`. We have also added a few other categories of sites to check particular hypotheses, or for comparison interest. We added the top ten ranked Computer Science departments.

To understand password policies we have wherever possible opened an actual account. We indicate in the "Acct" column of Table 7 whether we have set up an account or not. When we have, this means that we have verified the minimum allowable password strength. On the site `www.facebook.com`, for example, we have verified that a 6-digit PIN is acceptable by setting the password for an account controlled by one of the authors.

For some of the sites, we have been unable to set up accounts. For these we rely on published password policies. Searching for these policies was done manually using an Internet search engine. When this is the case we give a hyper-link to the source of the policy information in the electronic version of this paper. Some sites (particularly universities and government sites) can have many different computer systems. For example, students at the business school may use an entirely different system with entirely different policies from undergraduates. Thus, some of the university policies we indicate may be merely department policies. In the case of the government site `ca.gov`, for example, we found two distinct published password policies, which we list in Table 1. While these vary in detail, they conform to the general clustering we find in Section 3.2. In these cases, we document the first published policy that we found. Thus, while there is no guarantee that we find the only password policy in force at a university, there should be no bias in the policies indicated in Table 7.

Site	Len	Char. Sets	Strength
cdph.ca.gov	8	3 (3 out of U, L, N, S)	47.6
cps.ca.gov	9	2 (UorL + NorS)	46.6

Table 1: CA.GOV policies found.

2.1 Measuring Password Strength

Strength is intended to measure the resistance of a password to brute-force attacks. Measuring the strength of an individual password is non-trivial. Obviously length and composition (*i.e.*, number of different character sets) are good and increase the strength. Obviously, dictionary membership, repeated characters and consecutive sequences (*e.g.*, "abcedf" or "asdfgh") are bad and reduce it. Password-strength meters usually use a combination of length and composition to gauge strength; some check against a basic dictionary to flag the most common passwords. However, there does not appear to be a universally agreed-upon means to measure strength.

Measuring the strength of a policy is different. The intent of a policy is, presumably, to force that all users employ minimally strong passwords. To measure the minimum strength of the policy we use $N_{min} \log_2 C_{min}$ where C_{min} is the cardinality of the minimum character set required, and N_{min} is the minimum length. For example, the strength of a policy that requires 6-character passwords that allows digits would be $6 \times \log_2 10 \approx 19.9$ bits. A policy that requires 8 character upper, lower case letters and digits would be $8 \times \log_2(26+26+10) \approx 47.6$ bits. Table 2 gives examples of several different policies and their strength under this measure. This is clearly an approximate measure of strength, and arguments could be made whether this represents a true measure of the difficulty of attacking passwords that conform to the policy. Burr *et al.*[42] estimate that user chosen passwords have far less entropy than a randomly chosen password. The six character lowercase password "hwlbzu" is probably far more secure than the eight character "Pa$$w0rd" even though the first belongs to a 28-bit policy, and the second to a 52-bit policy. However this measure captures the strength of passwords that minimally conform to the policy. Since users appear to gravitate toward the weakest passwords allowed by the policy of a site [18] this probably gives a representative picture of the burden.

Some sites have positional restrictions on the characters. For example, `schwab.com` requires (see `https://`

Length	CharSets	Strength
6	N	19.9
6	LN	31.0
6	UL	34.2
6	ULNS	39.5
8	N	26.6
8	ULN	47.6
8	ULNS	52.7
10	N	33.2
10	L	47.0
10	ULNS	65.8

Table 2: Example password policies and their associated strengths. The symbols U, L, N, and S stand for upper, lower, numbers and special characters. For example "N" implies that digits alone are acceptable, while ULN indicates that both upper and lower-case along with digits are required.

`www.schwab.com/library/html/Privacy.html#Password`):
"Your password must be 6-8 characters long. It also must:

- Include both letters AND numbers.
- Include at least one number BETWEEN the first and last character.
- Contain no symbols (!, %, #, *etc.*)"

The second restriction (that there be a number between, *i.e.*, not at the beginning or end) increases the burden on the user but is not captured by our measure of strength. Since only 5 of 75 sites in Table 7 have positional restrictions we believe that ignoring this effect has minor influence on our analysis.

Thus, while the measure we use of policy strength is imperfect, it is adequate for our needs. Further this measure appears to preserve the ordering of policies in terms security and of burden on the user. We are primarily interested in the struggle between usability and security. Password strength gives a crude one dimensional measure of both of those things. This strength measure roughly preserves the ordering of difficulty of brute-forcing passwords. That is, sites with higher minimum strength have passwords that are harder to brute-force than those with lower. The security ordering is approximate; *i.e.*, differences of a few bits may not be meaningful. Strength also gives a measure of of usability that approximately preserves ordering: sites with lower strength have fewer restrictions and thus allow passwords that are easier for users to choose and remember. Again the the ordering is approximate. Our conclusions will not depend on small differences in strength.

2.2 What Threats Do Password Policies Address?

The purpose of password policies is to reduce certain attacks on user accounts. Principally these are:

- Online brute-force attacks
- Off-line attacks on the file of hashed passwords
- Password re-use across sites.

Strength policies, of course, have no influence on attacks such as phishing, keylogging, session hijacking *etc.*

2.2.1 *Online Brute-force Attacks*

A basic brute-force attack occurs when an attacker repeatedly tries many passwords for a single user account. It is standard practice to guard against this attack by locking an account, for example, after a threshold number of unsuccessful login attempts. For example, if an account is locked for 24 hours after three unsuccessful attempts, then even a 6-digit PIN can withstand 100 years of sustained attack [21]. More flexible lockout strategies, that render the attacker's job even harder, while inconveniencing legitimate users less, are also possible [38]. Thus a good lockout policy effectively makes direct brute forcing on a single account infeasible. The Denial of Service (DoS) vulnerability that it opens is a price that many large sites have decided they must accept or manage through back-end fraud detection.

A bulk-guessing attack occurs when the attacker distributes the guesses among many different accounts [31, 21]. Thus, rather than send one million password attempts against a single account (which will be blocked by the lockout policies) the attacker may send one attempt each against a million different accounts. This type of attack is much harder to address using lockout policies, since no account receives an unusual amount of traffic. However, it does require that the attacker know, or guess, a large collection of account usernames. It is only feasible against sites with enormous user bases [21]. It also ensures that the attack is un-targeted: the attacker has no control over which account he will break.

2.2.2 *Off-line Brute-force Attacks*

To authenticate a user at login requires verifying that the correct password has been entered. The best practice regarding the handling of passwords is to store, not the password itself, but a hashed version. By computing the hash of what the user enters the server can verify whether this matches the hashed password on file. Thus, there is no need for the server to retain a copy of user passwords, and it is regarded as bad practice to do so (although certain sites do [29]). Further, the password is generally salted with a per-account salt before hashing:

$$\text{Salted Hash} = \text{hash}(\text{password.salt}),$$

where "." denotes concatenation. The salt can be stored alongside the username and the salted hash. Thus the file of hashed passwords might have rows of the form:

$$[\text{username}, \text{hash}(\text{password.salt}), \text{salt}].$$

To authenticate a user the server need merely recalculate the salted hash and compare with the stored value.

Off-line attacks on the file of hashed passwords are a serious threat. A person who obtains the hashed passwords file might then deploy a tool such as JohnTheRipper [3] to crack the passwords. This attack is frequently cited as the disgruntled employee attack: someone who obtains a copy of the file might afterward attack the hashes of all user accounts at leisure, since no lockout policy limits the number of trials. The attacker can try passwords as quickly as his machine can calculate hashes.

Several protections are employed against this risk. First, the salt that is added before hashing prevents a rainbow attack [36], in which the attacker pre-computes the hashes of common passwords and strings. This ensures that even common passwords such as "abcdefg" will require significant effort to brute-force. Second, an iterated hash, which is designed to be slow to compute, can be used to slow down any brute-forcing attempt. For example, if the hashing algorithm is hash(password.salt) = SHA1^M(password.salt) then SHA1 is computed M times before producing the output. This introduces an M-fold delay into computing the hash. This slows the verification process for the user scarcely at all, but slows the off-line attacker down by a factor of M. Obviously, M is chosen so that acceptable delay is presented to the user.

Finally, and most importantly, the site must guard against access to the file. For any off-line attack to make sense the attacker must have read access to the hashed password file. In addition he should lack write access. If the attacker has write access he might as well write his own hash and effectively change the user's password; he can change it back after he has accessed the account to avoid arousing suspicion. For web accounts neither read nor write access will be available to an attacker who lacks administrative privileges. Thus it is purely an administrator, and more likely an ex-administrator, that is the main risk.

2.2.3 *Password Re-use Across Sites*

Some sites have policies that make password sharing difficult. For example, the third requirement in Schwab's policy above (which forbids symbols) ensures that no Schwab password could also be used at CMU (which requires them). It is possible that this is not accidental: perhaps some sites choose restrictive policies to discourage password re-use across sites? This would be most easily accomplished by truly capricious composition requirements. For example, one site requires that every password contain one of the two symbols '%' and '-'. Alternatively, positional requirements, such as "one number between the first and last characters" (second of Schwab's requirements) can have this effect. In our examination we found that such rules are rare. We found this only 5 cases out of 75 had positional requirements. While complex rules may make cross-sharing of passwords harder, and this is often suggested as a best practice, we found little to suggest that this is a primary goal of password policies.

3. THE DATA

The data we have gathered is presented in Table 7. The traffic rank, and the determination of whether the site accepts advertising are drawn from `www.quantcast.com` which tracks data for advertisers.

3.1 Diversity of Password Policies

In Figure 1 we plot strength *vs.* number of sites, showing the distribution of policies. We first observe that there is great diversity in policy strengths. There is little sign of an industry-standard or preferred policy. The diversity of strengths suggests that policy decisions are made more or less independently at different sites. Second, there is an enormous range of strengths. Certain sites have truly weak policies: Wikipedia for example allows single digit passwords, as do a few of the medium traffic sites. Since Wikipedia allows edits without logging in, passwords don't necessarily protect much by way of privilege or resources. Even if we ignore such sites, and restrict attention to those that involve email, commerce *etc*, there is a 30-bit range from low to high. Ranging from 20 bits at the low end (*e.g.*, Facebook, Live, Amazon *etc*) to 52+ bits at the high end (*e.g.*, Princeton, CMU, UsaJobs.gov) there is an enormous range. The weaker policies are far weaker than the strong. If passwords were randomly chosen, the weakest Amazon passwords would come into range after about of $2^{20} \approx 10^6$ attempts, and UsaJobs passwords after $2^{52} \approx 4.5 \times 10^{15}$ attempts. That is, there are nine orders of magnitude difference between how hard it is to brute-force an Amazon password and a UsaJobs.gov one. Do the security requirements of Amazon and Usajobs.gov really vary that much?

3.2 Clustering

In spite of the diversity of strengths the policies listed in Table 7 are far from random. Some patterns are very evident. In Table 3 we show the median strength by category. There is clear clustering of policy strength by category. For example all of the high traffic sites have relatively weak policies, with a median of 19.9 bits. Banks and brokerages have a mixture ranging from weak to medium strength with median of 31.0. Universities and government sites, with a few excep-

Site	Median Policy Strength
Top Traffic	19.9
High Traffic	19.9
Medium Traffic	8.3
Financial	31.0
Large Universities	44.5
Top CS Depts	46.4
Government	47.6
All `.com`	19.9
All `.edu`	43.7
All `.gov`	47.6

Table 3: Median strengths of policies for various groupings.

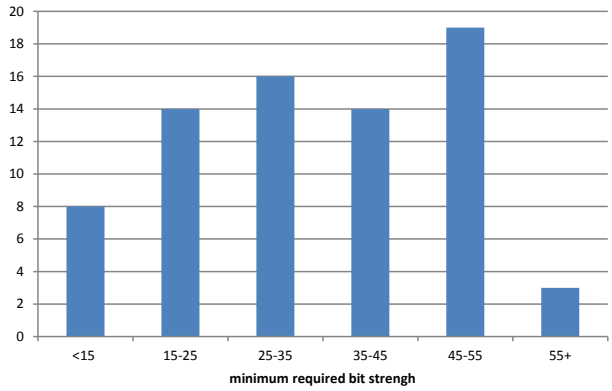

Figure 1: Histogram of policy strengths. Observe that policies span an enormous range: 55 bits is enormously stronger than 20 bits.

tions, have very strong policies with medians of 43.7 and 47.6 respectively. We divide the data of Figure 1 by top level domain and plot the policy strengths of the `.com`, `.edu` and `.gov` sites in Figure 2. A very clear pattern emerges: the `.com` sites are separated from the `.edu` and `.gov`. They have respective median policy strengths of 19.9, 43.7 and 47.6.

It is possible that not all sites make independent decisions. The clustering that is clear among the `.edu` and `.gov` sites might suggest that some sites may decide policies based on what their peers do, or on some guidelines. For example government and university sites may be under greater pressure to comply with the NIST [42] or DoD [8] password guidelines.

4. FACTORS THAT MIGHT INFLUENCE STRENGTH AND USABILITY

We now examine several factors that might explain the stronger policies of some sites. In each of the following sections we examine features of the sites that might force some sites to require more secure policies. All sites, of course, must guard against both online and off-line brute-force attacks. Some sites manage this with far lower strength policies than others. Are there reasons that make the stronger-policy sites more likely to be attacked? Or is the additional strength demanded by those sites superfluous?

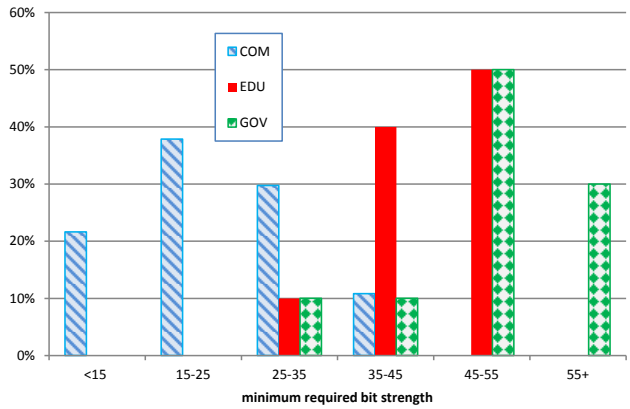

Figure 2: Histogram of policy strengths by first level domain. Observe that .coms tend to adopt significantly less stringent policies.

4.1 Are Password Policies Based on Observation and Evidence?

We should first consider whether policies are based on evidence. For example, might it be the case that sites have, by trial and error, reached the policy needed to protect their resources? For example, might those with stronger policies have seen greater attacks and learned the need for greater security? We now argue that this is not the case for several reasons.

1. Policies cannot be changed easily
2. Only when policies are too lax does a site get any evidence of brute-forcing
3. Best practices prevent gathering of data
4. Sites cannot necessarily distinguish brute-force from other attacks.

First, a trial and error approach to policy is hard. Tightening and loosening policy to explore the feasability space is not practical.

Second, it is not surprising that evidence does not appear to guide policy formation. Policies that are significantly too weak would be the best source of data: only by making the mistake of being too lax can we determine where significant breaches occur. However policies that are strong enough to repel brute-force attacks do nothing to tell us how much cushion the policy

provides. Such a policy gives us no data on whether the policy is far too strict, a little too strict or just right.

Third, the best practices for handling passwords makes gathering of such evidence hard. Best practice is to store not the password, but its salted hash (see Section 2.2.2). Thus, in general, the site has no information about the strength of user passwords. It has no means of determining, for example, whether users who report account hijacking have weaker passwords than average. To make this determination it would be necessary to store strength information. This would be very risky: if an attacker obtained strength measures in addition to the file of hashed passwords this would give him a road-map as to which to attack first.

Finally, it is hard to determine whether bulk guessing is responsible for hijackings. Here, since the attacker distributes his tries among many accounts, there will be little trace in the logfiles. One million accounts might each have a single unsuccessful login attempt, but it is exceedingly difficult to link this information with a successful login. Thus, when the owner of the hacked account complains or raises the alarm, it is by no means simple to determine whether they were a victim of phishing, keylogging or bulk-guessing. Thus, we reject the hypothesis that evidence of actual brute-force attacks forms policy.

4.2 What is The Size of the Service?

A factor that might generate the need for greater security is the size of the service. A potentially serious threat for web-sites is that of a bulk-guessing attack [26], explained in Section 2.2.1. This requires that the attacker can determine, or guess, the usernames of a large number of users [21]. So this attack works best against very large sites, and those where the username is known. For very large sites, that have tens or even hundreds of millions of users, it is safe for an attacker to assume that the username space is fairly well occupied.

Thus, if bulk guessing is a major threat we would expect to see some correlation between strength requirements and the number of user accounts at a site. If we take traffic as being correlated with number of users Table 7 makes clear that no such relation holds. In fact the weakest policies are found at the sites with highest rank. For clarity we pull the top five traffic, and top five universities and tabulate in Table 4. The largest, top traffic sites on the Internet have weaker, not stronger policies than those further down the list. Thus, an inverse relation appears to hold between traffic and the strength of the password policy that a site forces on its users.

Since traffic correlates only approximately with number of user accounts in Table 4 we also tabulate the number of users of various different sites. If bulk-guessing is a significant threat then we would expect to see larger sites force stronger password policies. Again, no such pattern emerges. Again, the reverse is the case: the larger the site the weaker the policy it forces on users. Thus, we reject the hypothesis that traffic or number of users explains the increase strength of the strong-policy sites.

Site	Users	Rank	Min. Strength
Facebook	400 million	2	19.9
Yahoo!	260 million	3	19.9
Live	260 million	8	19.9
Gmail	91 million	1	26.6
Twitter	76 million	31	19.9
Ohio State	51800	1811	41.4
Arizona State	51200	3288	47.6
U. of Florida	50900	1382	47.6
U. Minnesota	50400	919	35.7
U. Texas	49000	946	47.6

Table 4: Number of users at five top traffic sites, and five largest universities. University numbers are undergraduate enrollment, so may understate the true number of users by 50% or so to account for faculty, staff and graduate students.

4.3 Is the Username Public?

The bulk-guessing attacker must either know or be able to guess the usernames of a large number of users. The attack requires that he distribute the guesses among many accounts, and thereby evade both lockout and fraud detection. In Section 4.2 we examined size as one attribute that aids the bulk-guessing attacker. However, for sites where the username is public, there is also little difficulty obtaining the list. Email accounts, for example, aren't private; they are, by nature, public. For some sites, *e.g.*, email accounts, the username is visible to an attacker or can be determined. For example, many companies have email accounts for employees of the form: `firstname.lastname@company.com`. This makes bulk guessing against the email portal simple if a list of employees can be obtained. Thus, we might expect that if the username is public a stronger password policy must be imposed on users.

For email providers, social networking and auction sites we consider username to be public. Thus for a majority of the top traffic sites username is public. It used to be common practice for banks and brokerages to use either Social Security Number (SSN) or account number as the username. For example, some banks originally gave existing customers online access using their SSN as username and ATM card PIN as password. This eased the way toward getting many customers online quickly without the need for expensive in-person bank visits, or phone support. Most banks now ap-

Site	Assets	Min. Strength
Bank of America	$2.2 trillion	41.0
Chase	$2.0 trillion	36.2
Citibank	$1.8 trillion	31.0
Fidelity	$1.4 trillion	19.9
WellsFargo	$1.2 trillion	31.0
Vanguard	$1.0 trillion	26.6
Paypal	$290 billion*	26.6

Table 5: Value of assets and password strength. Except where noted the assets data comes from the FFIEC [17]. Fidelity and Vanguard assets from their press sites. *For Paypal we list their annual transaction volume, since they do not manage assets.

pear to offer the option of using a chosen username, and some mandate changing away from SSN. Account numbers are printed on checks and cannot be considered private. SSN is marginally private information. In fact, Acquisti and Gross [9], show that SSNs are in certain cases predictable from entirely public data. While difficult to generalize, for financial institutions the username is public, even though many are making effort to end this practice. For universities in many cases the username is also an email address, and is thus public.

University usernames tend to be public, but so also are those of top traffic sites and email providers. Thus, while there is no reverse correlation we reject the hypothesis that having a public username drives the requirement of policy strength. If the largest email providers, such as hotmail, Yahoo! and Gmail can manage with weak policies it doesn't appear that visibility of the username makes bulk guessing sufficiently bad to warrant increased strength.

4.4 What is the Value of the Resources Protected?

A very obvious possible determinant of security requirements is the value of the resources protected. Greater security is probably warranted for financial accounts than social networking ones. For the financial sites we tabulate the assets under management in Table 5. It is difficult to compare assets across the site categories selected. However, it is hard to argue that value of assets are responsible for strong policies at UsaJobs.gov when we compare with Fidelity or Paypal. Thus, the sites in Table 5 provide counter-examples to the hypothesis that value of assets might be the determinant in requiring stronger policies.

4.5 What is the Extractable Value of the Resources Protected?

In most cases of cybercrime it is not the password the attacker wants, but money. Monetizing a hijacked account can itself be a difficult process. In fact there are numerous accounts that stolen credentials are offered for sale on underground markets for fractions of their apparent face value [39, 30, 25, 24]. The amount of money that can be extracted from an account is not necessarily related to the net assets. If there is a correlation between value of resources and strength of policy it it more likely to be extractable assets that will predict the need for more stringent policies. The greater the extractable value of an account to an attacker, the greater we would expect the security required of users to be.

Fortunately, we have a means of estimating which sites attackers value most. Password brute-forcing is merely one means of account hijacking. There are many other attacks on account credentials, among which phishing is one of the most popular. In seeing which sites are most targeted by phishers we get an indication of which accounts are most valuable to them. We tabulate the number of distinct phishing attacks targeting sites on our list in Table 6. The data comes from Avira's 2009 study of the subject [12]. As can be seen, Payal, Chase and eBay dominate the list. Interestingly, brokerages with large assets under management, like Fidelity, Vanguard and Schwab don't even make the list. Presumably it is a great deal easier to get money from a hijacked Paypal account than a Fidelity one.

Paypal is clearly the favorite target of phishers. Thus, it's attractiveness to attackers is not in doubt. It does not seem plausible that Paypal is targeted (relative to other sites) a great deal by more phishers than by bruteforcers. Thus Table 6 offers a crude guide to extractable assets. Paypal, Chase and eBay all have high extractable value and yet have relatively weak policies. Thus it does not appear that higher extractable value explains the difference between strong and weak policy sites.

4.6 Who Lives with the Consequences of a Breach?

When a free web-mail account is compromised it is largely the user who bears the direct consequences. While there are support costs, and loss of reputation, the resources protected behind many free sites belongs to the user. This situation may be different at other sites. We investigate the hypothesis that web-sites insist on greater strength when they bear the cost of a breach.

For financial institutions, ironically, the institution has most to lose. This is the case, at least in the US, since losses due to unauthorized transfers are governed by Regulation E of the Federal Reserve [4]. This covers all transfers except by check and credit card, and limits the user's liability to $50 if the loss is reported within two days of discovery. Some of the institutions go beyond this. For example, Wells Fargo, in their online security guarantee states [6] "We guarantee that

Site	Phishing	Min. Strength
Paypal	32205	27
Chase	25901	27
eBay	18738	31
Bank of America	4540	41
IRS	3712	47
Citibank	2265	31
Facebook	2217	20
Gmail	761	27
Yahoo!	761	20
WellsFargo	541	31

Table 6: Number of phishing sites attacking various sites in 2009. Observe that the ordering is very different from the listing of financial sites by assets in Figure 5. Paypal is the favorite target of phishers, while Fidelity, which has $1.4 trillion under management doesn't even feature. The phishing data comes from Avira [12].

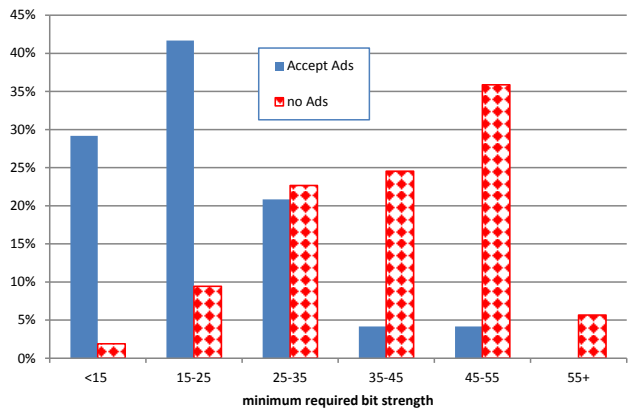

Figure 3: Histogram of policy strengths by sites that accept advertising and those that do not. The median strength when the site accepts ads is 19.9 bits and 41.4 when it does not.

you will be covered for 100% of funds removed from your Wells Fargo accounts in the unlikely event that someone you haven't authorized removes those funds through our Online Services." Similarly, Fidelity's Customer Protection Guarantee reads [5] "We will reimburse your Fidelity account for any losses due to unauthorized activity." Thus banks and brokerages provide counter-examples to the hypothesis that this explains the difference between strong and weak policy sites.

4.7 Is Advertising Accepted?

We have so far examined the trends that might push strength policies upward. For example, number of users and value of assets might tend to increase the attack energy and push policies in the direction of greater password strength. In the preceding sections we examined several of these forces and found no strong reason to explain the difference between those with stronger policies and weak. Now we examine several factors that might tend to push strength policies down. That is, all sites desire security, which exerts upward pressure on strength policies. If there were no cost to this then all sites would choose very complex password policies. However, sites also desire usability for their users, which exerts downward pressure. The more usable a site the more users are attracted to it. Attracting, and keeping users is an imperative for many web businesses. Traffic translates into revenue for sites that are advertising supported. We now examine whether there is inverse correlation between accepting advertising and password strength.

In Table 7 we tabulate whether a site accepts third-party advertising or not. As can be seen, the majority of top traffic sites are advertising supported. Banks, universities and government sites are not. In Figure 3 we show the histograms of sites that accept and do not accept advertising. The difference in histograms shows the stark contrast between the policies these two types of sites. The median strength for those that do is 19.9 bits, while it is 41.4 for those that do not.

This suggests a partial explanation of the question that has vexed us. The large `.com` sites live or die by the traffic they generate. The more users login and use their service the more traffic and revenue they generate. For example, Facebook, of course, wants as many users as possible. In addition, it wants them logging in as often as possible. Compromised user experience leads to less usage. Strong passwords diminish the user experience in that they are harder to remember. Forgetting a password, and going through the password reset procedure is inconvenient. Thus there is a powerful economic incentive for advertising supported sites to make passwords as usable as possible. Thus for sites that accept advertising there is a force opposing those that push for greater strength.

Advertising is one way in which web-sites generate revenue. For many sites this is far from being the dominant source of revenue however. Retailers such as Amazon, clearly have a revenue opportunity every time a user logs in. Brokerages, such as Fidelity, Schwab and Vanguard also have a revenue opportunity at each login. Every time a stock, bond, or mutual fund is bought or sold they make a commission, even if there is no advertising. As a for-profit university, where a large portion of student interaction is online, the University of Phoenix also has a revenue opportunity per login, even though it does not accept advertisements. Thus, even among those that do not accept advertising, several of those with less restrictive policies have less restrictive policies.

4.8 Does the Site Advertise?

An even more direct measure of the desire to attract traffic is whether the site itself advertises. This is evidence that it spends money to attract users. We now examine whether a site buys the Google adwords that correspond to its name. For example, when searching for "Fidelity" the first link is a sponsored one pointing to `www.fidelity.com`, indicating that Fidelity has bought this adword. Adwords are decided by auction, thus Fidelity has bid (and is paying) more than any other site was willing to pay to have their site appear in a privileged position in response to that query. Sponsored links appear either above or to the right of the ranked links returned. In the second to last column of Table 7 we tabulate whether we found sponsored links in response to Google queries that were paid for by, and pointed to the site. To ensure our result is unbiased we searched only for the name of the institution, both with and without spaces between words. Thus, for `overstock.com` we searched both for "overstock" and "over stock." Finding a sponsored link in this way certainly tells us that the site purchases adwords, whereas failure to find does not mean that no adwords are purchased. Some conclusions emerge. First, the top and high traffic sites generally do not buy sponsored links. These sites are large enough that they are ranked as the first returned link for a query "facebook" or "ebay." It makes little sense to pay Google for a sponsored link if the site itself is the first returned page. If we ignore the top and high traffic sites the median policy strength is 28.8 bits for those tat purchase adwords, and 41.4 for those that do not.

This feature is most interesting in the case of financial and government institutions and universities. The financial institutions, with the sole exception of JP Morgan Chase return a sponsored link. For the universities the reverse is the case: only University of Central Florida and University of Phoenix purchase sponsored links that point to their site. Not only does it place sponsored links for the query "University of Phoenix" but several other queries such as "University," "College," "Degree" all produce links sponsored by `phoenix.edu`. None of the `.gov` sites purchase sponsored links. Thus willingness to pay to attract traffic correlates well with less stringent policies.

4.9 Does the User have a Choice?

Just as there is diversity in the services offered by the sites, there is diversity in the nature of the users' relation with the site. With sites such as Facebook, Amazon, and Yahoo! the relation is entirely online, while with others it is the online portion of an interaction that takes place primarily in the off-line world, That is, a Facebook user opens and manages his account online; he never speaks to Facebook on the phone, and never visits a physical premises. With a university, on the other hand, a student's main contact is off-line.

This distinction is important as it indicates how much choice the user has at the time the online account is created (*i.e.*, when a password that conforms to policy is being chosen). For purely online accounts the user still has considerable choice at the time of account creation. Rather than open an Amazon account he can choose any other online retailer. Rather than choose Gmail, he can choose Yahoo! or hotmail for a webmail account. This is not the case with universities or government sites. A student at Ohio State, or most other universities, is already a student when he sets up an account. The web site is a monopoly provider of particular online services to the student body. Going to a different provider, or even choosing not to bother, is not an option. The University of Phoenix appears to be the only example of the universities studied where the user has choice at the time of account creation. We tabulate whether the user has a choice in the last column of Table 7.

At `.gov` sites, again users have no choice. There is only one Social Security Administration, one Internal Revenue Service, and one office of the Census. Figure 4 shows the histogram of strengths for the cases where the user does, and does not have choice. At a majority of the financial sites the user has no choice; *i.e.*, the relationship with the bank is probably already established prior to opening the web account. Paypal is an exception, since, for most users, is it an exclusively online relationship. In Figure 4 we display the histogram of sites where the user does and does not have a choice. The median strength for those where the user does is 19.9 bits, while it is 41.5 where the user does not. When the user has a choice at the time of account creation the site must compete for the the account. The large gap in median policy strength between these two cases suggests that sites that compete actively for users and traffic believe that restrictive policies can reduce traffic.

5. DISCUSSION

5.1 Security Demands, Usability Demands and Equilibrium

In our examination of security requirements (Sections 4.1 - 4.6) we failed to find any positive correlation between increased security demands and password strength. Those sites with more restrictive policies do not appear to have greater security concerns. In our examination of other factors (Sections 4.7 - 4.9) we did find that those sites which accept advertisements, purchase adwords and where the user has choice, appear to have less restrictive policies. These factors have in common that they indicate that the site competes for users and traffic; anything that affects usability has a negative impact.

This suggests that policy is determined, not by the

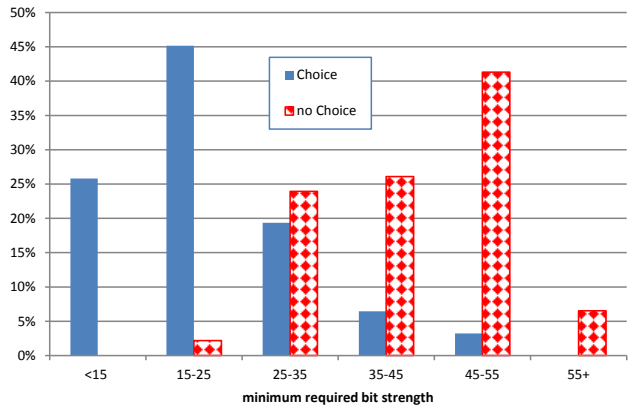

Figure 4: Histogram of policy strengths by sites where the user does and does not have choice at the time of account creation. The median strength when the user has a choice is 19.9 bits, and 41.5 when he does not.

security demands of the site, but by an equilibrium reached between the competing demands of security and usability. Security exerts an upward pressure, while usability exerts a downward pressure. Most of the sites we have examined have considerable security requirements. It is not plausible, for example, that sites like Amazon, Paypal and Fidelity persist with policies that do not allow them to protect user accounts from brute-force attacks. The security demands of their businesses are at least as great as any of the sites we have examined; and yet they manage to meet them with relatively unrestrictive policies. Thus, it does not appear to be security requirements that explain the diversity of password policies, but the different degrees to which sites face the consequences of poor usability. At Amazon, Paypal and Facebook the consequences of poor usability are great. Everything is optimized to make account creation and login as simple as possible. Any sub-optimality in either leads to lost revenue. The voices that argue for more restrictive policies meet vigorous push back. At government and university sites, by contrast, every login event is either a matter of indifference or a cost, and the direct consequences of poor usability are small. The data confirms that, at these sites, voices that argue for more restrictive policies have an easier task.

5.2 How Strong do Passwords Need to be?

How strong a password needs to be seems to depend on whether we must protect against online or off-line attack. In turn, this question seems to reduce to whether we can prevent the file of hashed passwords from leaking to an attacker or not (*i.e.*, whether we deal with the attack of Section 2.2.1 or Section 2.2.2). If the file can be protected, then we need worry about online brute-force attacks only. The examples of Amazon, Paypal and Fidelity prove that sensible lockout policies and fraud detection ensure that this can be done at relatively modest strengths. If the file of hashed passwords cannot be protected, then greater strength gives some protection against attacks on the hashed passwords. Thus stronger policies protect, not against online attacks on user accounts, but against failure to protect the hashed password file.

This is interesting since it suggests that sites with stronger policies do not offer better protection against online attacks, they merely shift some of the burden of protecting against off-line attacks to the user. The cost of relatively weak password policies does not appear to be increased success of brute-forcing. Rather, it is that these sites must invest greater effort to ensure that the file of hashed passwords never leaks. The benefit is that they offer a more usable experience to their customers. When sites enforce very restrictive policies it does not appear that they see brute-forcing less. The benefit is that they enjoy some cushion in the event that the hashed passwords ever leak. The cost for this cushioning is borne by their users.

Our conclusion on password strength is informed by data. Some of the most attacked sites on the web manage with passwords of length 6 or 8. Several require two character sets; *e.g.*, lowercase and digits, or lower and upper case. After this, explicitly forbidding common passwords such as "abcdef" appears a better approach than imposing additional complexity. Looking at Table 7, insistence upon special characters in the password appears to be the exclusive preserve of those insulated from the effects of poor usability. Equally, (again from Table 7) the practice of forcing regular password changes, which Spafford [16] suggests "has little or no end impact on improving security" is mostly enforced by university and government sites.

5.3 Policies Do not Need to Tighten With Time

Increasing amounts of cybercrime, identity theft and phishing are often invoked as reasons for increasingly stringent password requirements. We argue that this view is incorrect: there is no reason why password policies in 2010 need be any stronger than they were in 2000. Moore's law and reductions in the cost of computation have no influence whatever on online brute-force attacks. Advances in cracking software, faster hardware, or more hardware, do not make the online attacker's job easier. He is limited by the lockout policy, which limits his attempts per unit time and fraud detection. It is worth noting that improvements in off-line brute-force attacks can also be limited. If M is chosen to generate a fixed delay per hash computed, this can be increased as machine speeds improve. A 10× improvement in compute ability can be accommodated by replacing $\mathrm{SHA1}^{M}()$ with $\mathrm{SHA1}^{10M}()$.

There have been a number of breaches involving passwords recently. Twitter was the subject of large online brute-force attack [7]. Failure to lock accounts after several attempts allowed compromise of several user accounts. Recently Twitter announced a requirement that users strengthen passwords. Rather than increase from the current bit-strength of 19.9, they explicitly rule out the 370 most common passwords. RockYou also had a recent attack. Their site had a SQL injection vulnerability and an attacker gained access to (and posted online) 32 million passwords that were kept in the clear. Rockyou also announced changes to password policies: instead of an unconstrained 5-character password users must chose 8-character passwords with with at least two of upper, lower case, numbers and special characters. Thus, an attack unrelated to password strength caused a tightening of strength policy. As Zwicky points out [15] "the strength of peoples' passwords at RockYou was totally irrelevant." The Imperva analysis [28] suggests that a brute-forcing strategy against RockYou would have yielded a significant fraction of accounts. Yet, the need for stronger user passwords is a strange conclusion to draw from this episode: we do not know if any of the accounts were brute-forced, but we do know that 100% of them were compromised. The RockYou user who chose a 10-character complex password suffered exactly the same fate as the one who chose "abcdef."

6. RELATED WORK

The literature on passwords and alternative means of authentication is vast. There has been a growing literature documenting that users are overwhelmed with password policies and the difficulty of choosing, remembering and maintaining many different accounts. Adams and Sasse [11] show that choosing and remembering strong passwords is a challenge for many users. Zurko and Simon [35] is an early example calling for security policies that pay attention to the burden placed on users. Norman probably speaks for many when he speaks of his frustration with the Northwestern University password policies [14]: "Because when security gets in the way, sensible, well-meaning, dedicated people develop hacks and workarounds to defeat it." Incidentally, the Northwestern policy that Norman cites (listed in Table 7) is among the least restrictive university policies.

In earlier work we document that users often choose weak passwords and re-use them liberally [18]. In studying the behavior of half a million users we discovered that users generally gravitate toward the weakest passwords allowed by policy, that they have on average 25 passwords each, and re-use each password across 6.5 sites. Gaw and Felten [41] also study password habits in a user study of undergraduates. They find lower numbers of accounts and re-use rates, but did find that both increased steadily with time.

St. Clair *et al.*examine the question of whether passwords are facing exhaustion [32]. Numerous alternatives to text passwords have been proposed. These include graphical passwords [27], and one time passwords [22, 20]. Florêncio *et al.*[21] suggest that password strength for web accounts is not as important as frequently assumed. They argue that when there is only an online brute-force attack adequate lockout policies make brute-forcing infeasible. In one of the most closely related works Mannan and van Oorschot [33] examine usability in online banking. They study policies beyond passwords, and find that compliance is in some cases almost impossible. Herley *et al.*[13] examines the state of passwords and why better progress has not been made toward stronger authentication methods. Beautement *et al.*[10] suggest that users have a finite budget for dealing with security policies, and that increasing complexity in one area must be matched by reductions elsewhere. Herley [23] suggests that users behave rationally in ignoring recommendation to choose stronger passwords and other security advice. The recommendations place considerable burden on them, and deliver little reduction in risk. Sasse *et al.*[34] investigate strategies to enhance password strength and security, while reducing the burden on users. Very recently, Bonneau and Preibush [29] performed a study of how passwords are handled at 150 different web-sites. Theirs is the only other work we know of that attempts to gather and interpret such a large collection of policies and practices. Inglesant and Sasse [37] examine password practices in several organizations and suggest that security managers systematically underestimate the cost that stringent policies impose. If, as this paper suggests, those policies are unnecessarily stringent this implies that much of this cost is wasted. Several other authors have recently suggested that our practices on security matters may be outdated and in need of revision. Bellovin [40] suggests that "security by checklist" is producing perverse outcomes.

7. CONCLUSION

Where do security policies come from? Our online and off-line lives are full of examples of security policies that restrict our behavior. We run anti-virus and choose strong passwords. We remove our shoes and laptops and restrict ourselves to 3 oz. quantities of liquids and gels. While most of us understand and accept that there is a tradeoff between security and convenience, how and by whom is this tradeoff decided? Few would argue with getting a lot more security for a little inconvenience. But, if the decision-making process is obscure how can we be sure we're not getting lots of inconvenience for little improvement in security? When the US Transportation Security Administration decided to impose a rule forbidding passengers to leave their seats or have anything on their lap in the last one hour of flight the outcry was immediate: "the people who run America's

airport security apparatus appear to have gone insane" (the Economist Dec. 27, 2009). Absent such absurdities it is hard to tell whether security policies have the convenience-security tradeoff just right, or whether they are overshooting greatly and imposing considerable inconvenience for marginal benefit.

Our conclusions suggest that, at least in the case of passwords, exactly such an overshoot occurs. Some of the largest and most attacked sites on the web allow 6 character PINS or lowercase passwords. By contrast, government and university sites generally have far stronger (and far less usable) policies. The reason we suggest lies not in greater security requirements, but in greater insulation from the consequences of poor usability. Most organizations have security professionals who demand stronger policies, but only some have usability imperatives strong enough to push back. When the voices that advocate for usability are absent or weak, security measures become needlessly restrictive. The watchers must be watched, not merely to ensure that they do not steal or cheat, but also to ensure that they do not decide to make their job a little easier at the cost of great inconvenience to everyone else.

8. REFERENCES

[1] http://www.internetworldstats.com.

[2] http://www.worldmapper.org/display.php?selected=336.

[3] http://www.openwall.com/john/.

[4] Regulation E of the Federal Reserve Board. http://ecfr.gpoaccess.gov/cgi/t/text/text-idx?c=ecfr&sid=0283a311c8b13f29f284816d4dc5aeb7&rgn=div9&view=text&node=12:2.0.1.1.6.0.3.19.14&idno=12.

[5] The Fidelity Customer Protection Guarantee. http://personal.fidelity.com/accounts/services/findanswer/content/security.shtml.cvsr?refpr=custopq11.

[6] Wells Fargo: Online Security Guarantee. https://www.wellsfargo.com/privacy_security/online/guarantee.

[7] Wired: Weak Password Brings 'Happiness' to Twitter Hacker. http://blog.wired.com/27bstroke6/2009/01/professed-twitt.html.

[8] Department of Defense Password Management Guideline. Technical Report CSC-STD-002-85, U.S. Dept. of Defense, Computer Security Center, 1985.

[9] A. Acquisti and R. Gross. Predicting Social Security Numbers from Public Data. *Proc. Natl. Acad. Science*, 2009.

[10] A. Beautement, M.A. Sasse and M. Wonham. The Compliance Budget: Managing Security Behaviour in Organisations. *NSPW*, 2008.

[11] A. Adams and M. A. Sasse. Users Are Not the Enemy. *Commun. ACM*, 42(12), 1999.

[12] Avira TechBlog. The Most Phished Brands of 2009. http://techblog.avira.com/2009/12/19/the-most-phished-brands-of-2009/en/.

[13] C. Herley, P.C. van Oorschot and A.S. Patrick. Passwords: If We're So Smart Why Are We Still Using Them? *Proc. Financial Crypto 2009*.

[14] D.A. Norman. The Way I See It: When security gets in the way. *Interactions*, 16(6):60–63, 2009.

[15] E. Zwicky. Brute Force and Ignorance. *;login*, April 2010.

[16] E.H. Spafford. Security Myths and Passwords. *http://www.cerias.purdue.edu/site/blog/post/password-change-myths/*.

[17] Federal Financial Institutions Examination Council. Top 50 Bank Holding Companies 2009. http://www.ffiec.gov/nicpubweb/nicweb/Top50form.aspx.

[18] D. Florêncio and C. Herley. A Large-Scale Study of Web Password Habits. *WWW 2007, Banff.*

[19] D. Florêncio and C. Herley. Stopping Phishing Attacks Even when the Victims Ignore Warnings. *MSR Tech. Report*, 2005.

[20] D. Florêncio and C. Herley. KLASSP: Entering Passwords on a Spyware Infected Machine. *ACSAC*, 2006.

[21] D. Florêncio, C. Herley, and B. Coskun. Do Strong Web Passwords Accomplish Anything? *Proc. Usenix Hot Topics in Security*, 2007.

[22] N. Haller. The S/KEY One-Time Password System. *Proc. ISOC Symposium on Network and Distributed System Security*, 1994.

[23] C. Herley. So Long, And No Thanks for the Externalities: The Rational Rejection of Security Advice by Users. *NSPW 2009, Oxford.*

[24] C. Herley and D. Florêncio. A Profitless Endeavor: Phishing as Tragedy of the Commons. *NSPW 2008, Lake Tahoe, CA.*

[25] C. Herley and D. Florêncio. Nobody Sells Gold for the Price of Silver: Dishonesty, Uncertainty and the Underground Economy. *WEIS 2009, London.*

[26] K. Hole, V. Moen, and T. Tjostheim. Case Study: Online Banking Security. In *IEEE Security and Privacy*, pages 14–20, 2006.

[27] I. Jermyn and A. Mayer and F. Monrose and M.K. Reiter and A.D. Rubin. The Design and Analysis of Graphical Passwords. In *Usenix Security*, 1999.

[28] Imperva. Consumer Password Worst Practices.

[29] J. Bonneau and S. Preibusch. The Password Thicket: technical and Market Failures in Human Authentication on the Web. *WEIS*, 2010.

[30] J. Franklin and V. Paxson and A. Perrig and S. Savage. An Inquiry into the Nature and Causes of the Wealth of Internet Miscreants. *Proc. CCS*, 2007.

[31] K. J. Hole and V. Moen and T. Tjostheim. Case Study: Online banking Security. *IEEE Security & Privacy Magazine*, 2006.

[32] L. St. Clair and L. Johansen and W. Enck and M. Pirretti and P. Traynor and P. McDaniel and T. Jaeger. Password Exhaustion: Predicting the End of Password Usefulness. In *Proc. of 2nd Intl Conf. on Information Systems Security (ICISS)*, 2006.

[33] M. Mannan and P.C. van Oorschot. Security and Usability: The Gap in Real-World Online Banking. *NSPW*, 2007.

[34] M.A. Sasse, S. Brostoff and D. Weirich. Transforming the "weakest link": a human-computer interaction approach to usable and effective security. In *BT Technology Journal*, 2001.

[35] M.E. Zurko and R. T. Simon. User-Centered Security. *NSPW*, 1996.

[36] P. Oechslin. Making a faster crytanalytical time-memory trade-off. *Advances in Cryptology - CRYPTO 2003*, 2003.

[37] P. Inglesant and M. A. Sasse. The True Cost of Unusable Password Policies: Password use in the Wild. *CHI*, 2010.

[38] P.C. van Oorschot, S. Stubblebine. On Countering Online Dictionary Attacks with Login Histories and Humans-in-the-Loop. *ACM TISSEC vol.9 issue 3*, 2006.

[39] R. Thomas and J. Martin. The Underground Economy: Priceless. *Usenix ;login:*, 2006.

[40] S. Bellovin. Security by Checklist. *IEEE Security & Privacy Mag.*, 2008.

[41] S. Gaw and E.W. Felten. Password Management Strategies for Online Accounts. *Proc. SOUPS.*

[42] W. E. Burr, D. F. Dodson W. T. Polk. Electronic Authentication Guideline. In *NIST Special Publication 800-63*, 2006. http://csrc.nist.gov/publications/nistpubs/800-63/SP800-63V1_0_2.pdf.

Site	Traffic Rank	Acct.	Min. Len.	Char Sets	Min. Strength	Exp. (days)	Posn. Restrc?	Accepts Ads?[7]	Places Ads?[8]	User Choice?[9]
Top Traffic Sites[1]										
Google[2]	1	Y	8	1	26.6	N	N	Y	Y	Y
Facebook	2	Y	6	1	19.9	N	N	Y	N	Y
Yahoo!	3	Y	6	1	19.9	N	N	Y	N	Y
Youtube	5	Y	6	1	19.9	N	N	Y	N	Y
AOL	6	Y	8	1	26.6	N	N	Y	N	Y
Live[3]	8	Y	6	1	19.9	N	N	Y	N	Y
Wikipedia	9	Y	1	1	3.3	N	N	N	N	Y
eBay	10	Y	6	2	31.0	N	N	Y	Y	Y
Amazon	11	Y	6	1	19.9	N	N	Y	Y	Y
ask	12	Y	6	1	19.9	N	N	Y	Y	Y
weather	13	Y	6	1	19.9	N	N	Y	N	Y
answers	15	Y	1	1	3.3	N	N	Y	N	Y
Myspace	16	Y	6	2	31.0	N	N	Y	N	Y
Craigslist	17	Y	6	1	19.9	N	N	N[10]	N	Y
adobe	20	Y	6	1	19.9	N	N	N	Y	Y
High Traffic Sites[1]										
nih.gov	101	N	8	3	53.6	60	N	N	N	N
capitalone.com	102	Y	8	2	41.4	N	N	N	Y	N
rockyou.com	103	N	8	2	41.4	N	N	Y	N	Y
typepad.com	106	Y	6	1	19.9	N	N	Y	Y	Y
overstock.com	107	Y	5	1	16.6	N	N	N	Y	Y
latimes.com	108	Y	6	1	19.9	N	N	Y	N	Y
intuit.com	109	Y	6	1	19.9	N	N	Y	N	Y
cbssports.com	110	Y	4	1	13.3	N	N	Y	N	Y
Medium Traffic Sites[1]										
wowwiki.com	1001	Y	1	1	3.3	N	N	Y	N	Y
virginia.edu	1002	N	6	2	36.2		Y	N	N	N
pgatour.com	1003	Y	1	1	3.3	N	N	Y	N	Y
hollywood.com	1004	Y	1	1	3.3	N	N	Y	N	Y
mit.edu	1006	N	6	2	31.0	N	N	N	N	N
okcupid.com	1007	Y	4	1	13.3	N	N	Y	N	Y
istockphoto.com	1008	Y	5	2	25.8	N	N	N	Y	Y
highschoolsports.net	1010	Y	1	1	3.3	N	N	Y	N	Y
Banks and Brokerages										
Fidelity	224	Y	6	1	19.9	N	N	N	Y	N
Vanguard	629	Y	8	1	26.6	N	N	N	Y	N
Schwab	2266	N	6	2	31.0	N	Y	N	Y	N
WellsFargo	80	Y	6	2	31.0	N	N	N	Y	N
BoA	48	Y	8	2	41.4	N	N	N	Y	N
JP Morgan Chase	2186	N	7	2	36.2	N	N	N	N	N
Citibank	316	Y	6	2	31.0	N	N	N	Y	N
PayPal	29	Y	8	1	26.6	N	N	Y	Y	Y
US Bank	316	N	8	1	26.6	N	N	N	Y	N
Large Universities[4]										
Ohio State U	1811	N	8	2	41.4	365	N	N	N	N
Arizona State U	3288	N	8	3	47.6	180	N	Y	N	N
U. of Florida	1382	N	8	3	47.6		N	N	N	N
U. of Minn.	919	N	6	3	35.7	N	N	N	N	N
U. of Texas	946	N	8	3	47.6	N	N	N	N	N
U. Central Florida	6313	N	8	3	47.6	N	N	N	Y	N

Continued on Next Page...

Site	Traffic Rank	Acct.	Min. Len.	Char Sets	Min. Strength	Exp. (days)	Posn. Restrc?	Accepts Ads?[7]	Places Ads?[8]	User Choice?[9]
Michigan State U	1174	N	8	3	47.6	N	N	N	N	N
Texas A& M	1418	N	6	3	35.7	183	N	N	N	N
U South Florida	2364	N	6	3	35.7	183	N	N	N	N
Penn. State U	977	N	8	2	41.4	183	N	N	N	N
Univ top CS Depts[5]										
MIT	1006	N	6	2	31.0	N	N	N	N	N
Stanford	858	N	8	3	47.6	180	N	N	N	N
UC Berkeley	905	N	8	2	41.4	N	N	N	N	N
CMU	3651	N	8	4	52.0	365	N	N	N	N
UIUC	3384	N	8	1	26.1	365	N	N	N	N
Cornell	955	N	7	3	41.7	183	N	N	N	N
Princeton	1879	N	8	4	52.7	N	N	N	N	N
U. of Washington	1032	N	8	2	45.6	N	N	N	N	N
Georgia Tech.	4687	N	8	3	47.6	N	N	N	N	N
U. of Texas	946	N	8	3	47.6	N	N	N	N	N
Government[1]										
irs.gov	63	N	8	3	47.6	90	N	N	N	N
usps.com [6]	68	Y	8	3	47.6	N	N	N	N	N
nih.gov	101	N	8	3	47.6	60	N	N	N	N
ca.gov	124	N	8	3	47.6	N	N	N	N	N
ed.gov	141	Y	8	1	26.6	N	N	N	N	N
noaa.gov	199	N	12	3	77.1	60	Y	N	N	N
weather.gov	228	N	12	3	77.1	180	N	N	N	N
census.gov	246	N	8	3	47.6	N	Y	N	N	N
ssa.gov	276	N	7	2	36.2	N	N	N	N	N
nasa.gov	342	N	12	4	79.0	N	N	N	N	N
Other sites										
U. of Phoenix	873	Y	7	2	36.2	N	N	N	Y	Y
Columbia	1350	N	6	2	31.0	N	N	N	N	N
Northwestern	4457	N	6	2	31.0	548	Y	N	N	N
VA	558	Y	8	4	52.7	N	N	N	N	N
USAJobs	590	Y	8	4	52.7	N	N	N	N	Y
TreasuryDirect	2421	Y	8	3	47.6	N	N	N	N	N
Twitter	31	Y	6	1	19.9	N	N	N	N	Y

Table 7: The Sites Examined.

[1] Traffic info from QuantCast.com. We investigated password policies for sites 1-20, 100-110, 1000-1010, and for top 10 government sites. We did not find policies for sites # 18 (about.com), # 104 (lowermybills.com), #105 (wheatherbug.com), # 1005 (taboolasyndication.com), and #1009 (inklineglobal.com).

[2] Google Account is also used on the site Blogger.com (# 14 in traffic).

[3] LiveID is used in four of the top 20 sites: MSN (# 4), Microsoft (# 7), Live (# 8), and Bing (# 19).

[4] Top 10 US universities by 2006 enrollment.

[5] Top CS Depts as per U.S.News.

[6] usps.com handles the redirected traffic from usps.gov.

[7] Advertising info from QuantCast.com.

[8]Does it purchase the AdWords for the name of the institution?

[9] Does the user tpically have a relationship with the institution even before first login to the site?

[10] Craigslist does, of course, accepts ads, but it does not accept paid advertising.