

The Password Game: negative externalities from weak password practices

Sören Preibusch and Joseph Bonneau

University of Cambridge, Computer Laboratory
Cambridge CB3 0FD, UK
sdp36 — jcb82 @c1.cam.ac.uk

Abstract The combination of username and password is widely used as a human authentication mechanism on the Web. Despite this universal adoption and despite their long tradition, password schemes exhibit a high number of security flaws which jeopardise the confidentiality and integrity of personal information. As Web users tend to reuse the same password for several sites, security negligence at any one site introduces a negative externality into the entire password ecosystem. We analyse this market inefficiency as the equilibrium between password deployment strategies at security-concerned Web sites and indifferent Web sites. The game-theoretic prediction is challenged by an empirical analysis. By a manual inspection of 150 public Web sites that offer free yet password-protected sign-up, complemented by an automated sampling of 2184 Web sites, we demonstrate that observed password practices follow the theory: Web sites that have little incentive to invest in security are indeed found to have weaker password schemes, thereby facilitating the compromise of other sites. We use the theoretical model to explore which technical and regulatory approaches could eliminate the empirically detected inefficiency in the market for password protection.

1 Password practices on the Web

Computer systems have traditionally authenticated users by prompting for a username and textual password. This practice has proliferated on the Web as users now create new login credentials on a monthly basis. More than 90% of the 100 most visited destinations on the Web and more than 80% of the top 1000 Web sites collect passwords from their users. Well-known weaknesses of password schemes were inherited by the Web, including weak user-chosen passwords and password-reuse across domains. These problems were exacerbated by self-enrolment on the Web and the proliferation of password accounts per user (estimated to be over 25 [9]).

Numerous password enhancement schemes have been proposed, including improved cryptographic protocols and alternative means for creating and entering passwords such as graphical passwords. The merit of such alternative authentication schemes remains contested as users continue to pick easily guessable secrets and fall for simple social engineering attacks. Automatic management tools beyond simple browser password caches have seen limited deployment [14].

In parallel, with the advent of global social networks as new identity providers, single sign-on schemes (SSO) have gained new interest after previous attempts such as Microsoft Passport have failed. Service providers have started to accept Facebook Connect for general-purpose Web authentication [2], whilst the non-proprietary OpenID [13] has gained an enthusiastic following but limited deployment.

In an earlier investigation, we set out to quantitatively assess the current state of password implementations on the Web [5]. In the first large-scale survey of current password practices, we surveyed 150 sites equally sampled from mainstream and less popular Web sites. We concluded that a large number of technical issues remain on the Web, with almost all sites studied demonstrating some security weaknesses and many sites implementing unique and inconsistent password security policies.

In this work, we present a fresh game-theoretic analysis into the economics behind these technical failures. Building on our previous identification of two key negative externalities in the password market, we model firms' decisions to invest in password security as a simple game between security-interested Web sites and security-indifferent Web sites. This partition is supported by the existing password landscape. We demonstrate that the empirical evidence about inefficient password practices matches the theoretical prediction from the game-theoretic model. We affirm the validity of our empirical evidence by analysing a second sample of public Web sites which is larger by an order of magnitude. We conclude that market failures provide a pertinent economic explanation as to why password security weaknesses are so commonly observed.

2 Incentives for password deployment

From a service provider's perspective, passwords enable authentication by ensuring that a specific identity can only be used by parties knowing the correct password. Passwords may also sustain real-world authorisation decisions if a username/password combination is tied to a pre-established offline identity, as in the case of online banking or corporate email facilities. Merchant Web sites, which often enable users to store payment details for future shopping, have direct incentives to secure password authentication, as they can be held liable for any purchases made using stolen password credentials.

In other sites, interaction with other users of the same site (e.g., social networking) or beyond (e.g., webmail) introduces the risk of harmful actions if authentication fails, such as fraud, spam, or blackmail which may damage the reputation of the operator of the Web site as well as harm users. Passwords are intended to make users accountable for their actions and eliminate plausible deniability. In some European jurisdictions at least, operators of interactive sites can be held responsible for their users' actions if they fail to take appropriate counter-measures. Banning users, locking accounts, and keeping evidence from electronic discovery for court trials would become impossible if access were not restricted. Some operators therefore have an intrinsic motivation to secure their

Web sites to protect their own assets, notwithstanding the negative business impact which careless practices can bring.¹

These incentives to secure password authentication do not hold for all Web sites using passwords, such as free news Web sites, which use passwords to provide access to a stored identity for customisation. From a security point of view, it is not clear why a service that is offered for free requires authentication. As long as users can create an unlimited number of free accounts, password authentication seems dispensable. It places no restrictions on the user other than the hassle of account creation—which is a rather weak throttling mechanism. The motivation for password deployment may be revenue rather than security. In combination with usernames, passwords act as persistent identifiers; they increase the accuracy of behavioural profiling. Password collection also initiates account creation during which further personal information is collected, including contact information such email address and socio-demographic details. This data may be used directly for targeted advertising or indirectly when aggregated to construct an audience profile, based on which advertising businesses book slots on the Web site. Password-protected accounts facilitate the extraction and monetarisation of customer data even when they have dubious security value.

Regardless of the motivation, there are costs associated with deploying password authentication which rise with the sophistication of password security mechanisms. These include at least the need for skilled programmers to implement and maintain the system and operating costs of password storage servers. Passwords also add communication overhead, which can be exacerbated by the need to transmitted encrypted data.

3 Negative externalities from password deployment

Password deployment by one company generates negative externalities in two ways: first, consumers are asked to remember yet another password; second due to password reuse each password server is an additional failure point for the entire system.

3.1 Too many passwords: tragedy of the commons

Given the low adoption of password management tools [14], mental storage capacity has the characteristics of a common good to password-collecting companies. A common good is a good from which no consumer can be excluded, as consumption is not or cannot be regulated and there is no cost for consuming the good. Unlike public goods, however, the quality of a common good decreases as more people consume it. There is an incentive to overuse the common good to the detriment of society as a whole, as one is not held accountable for the deterioration of the good's value. The resulting inefficiency is described as the “tragedy

¹ For example, a well-publicised password database compromise at the social gaming Web site RockYou brought both a storm of negative publicity and pressure from Facebook, a key business partner [15].

of the commons” in reference to pastures which were often over-grazed by the local community of farmers. Any given farmer could fully appropriate the benefits from sending another sheep to the commons, but the negative effects as the pasture depleted affected the entire community equally.

Humans have limited physiological capacity to remember random data. Passwords consume this capacity and as the total number of passwords remembered grows, the risk of forgetting typically increases. Indeed, the majority of surveyed users indicate that they do not use stronger passwords because they simply have too many to remember [11]. Yet companies can continue to ask consumers to register new passwords at no expense. Mental password storage capacity, therefore, exhibits characteristics of a common good: access to it is neither limited nor priced; with increased usage, its quality in terms of recall decreases. To cope with this password overload and maintain an acceptably low rate of forgotten passwords, consumers must lower their overall password quality. With per-password strength requirements in place, password reuse is a common coping strategy to lower the burden of remembering strong passwords.

3.2 Password reuse: negative externalities of insecurity

Password reuse enables each site to introduce a further negative externality on the password market. As companies have differing incentives to invest in password security, the resulting security of their password implementations inevitably differs. As the vast majority of consumers reuse the same password across multiple sites [9,10], passwords for high-security accounts such as banking and email are often held by less secure Web sites as well. The accumulation of high-security login credentials at Web sites with weaker incentives for security presents an attractive attack strategy: compromise a low-security site and attempt to use the login credentials at higher security sites. The feasibility of this attack has been strengthened by the trend towards all Web sites using email addresses as user identifiers, as observed at 87% of sites in our recent survey [5].

Indeed, within the past year strong evidence has emerged that this attack is not only feasible but is occurring in the wild. The January 2010 compromise of social gaming Web site RockYou! leaked over 32 million email/password pairs of which over 10% were claimed to be directly usable at PayPal [15]. The reuse of passwords from compromises at low-security sites is acknowledged as a common means of attack in documentation at both Windows Live [3] and Yahoo! [4]. The strongest evidence comes from Twitter, which in January 2010 forced millions of users to reset their passwords after detecting a large scale attack using a stolen list of credentials from a torrenting Web site [12].

Web sites with weak password security can thus create a real negative externality for more secure Web sites. As operators of low-security site do not bear the social costs of weak password practices, there is a tendency for under-investment in good password practices.

4 Password implementers: game-theoretical model

To understand the net incentive for companies to deploy password schemes and to invest into their security, we understand the Web as an arena where a password requirements game is played. We model password design choices as actions in a game with two players: one Web site operator with an intrinsic incentive to invest in security, the other without such an incentive, as detailed in Section 2. In the following, these two players will be labelled the security-sensitive Web site (\mathcal{W}_s) and the security-indifferent Web site (\mathcal{W}_i), respectively. The firms' payoffs in this game reflect how their own decisions to enforce passwords as well as those of the other firm affect their performance, risk exposure, and implementation costs. The resulting game-theoretic model will also serve to analyse the efficiency of the overall password practices on the Web.

4.1 Action spaces and timing

The action each Web site can take is to specify the set S of permissible passwords. When first using the site's service, users will need to create a password from within the set of permissible passwords. For simplicity, we assume there are weak passwords (wp) and strong passwords (sp). Although in reality password strength is a spectrum and little consensus exists as to how exactly two medium-strength passwords should be ranked in terms of security, the extremes can be identified quite easily.² Web sites \mathcal{W}_s and \mathcal{W}_i thus choose their action from the powerset of $\{wp, sp\}$: only strong passwords ($\{sp\}$), only weak passwords ($\{wp\}$), any password ($\{wp, sp\}$) or no password protection at all (\emptyset).

For simplicity, the Web sites choose their password schemes simultaneously. Still, the following analysis and conclusions would not differ much if the password requirements game is understood as a market entry game, where a newly created Web site and an incumbent Web site have to respond to the other player's past or pre-empted decisions. The timing of the game effectively determines how users can reuse their passwords, since only the password created with incumbent at an earlier point in time could be reused with the entrant at a later point in time. This directedness of the externality effect is removed in a simultaneous game. We believe sequential moves are contrary to the observation that, first, users' ability to change their passwords removes the relevance of timing, and second, the effects of password leakage do not depend on whether it was used originally or reused at the leaking site.

² Conflicting estimates of expected password strength given a set of requirements are found in NIST's Electronic Authentication Guideline [7] and the BSI IT-Grundschutz Methodology Catalogues [6], although both agree that phrases such as `pass` or `password` are weak.

4.2 Payoffs

The pay-offs in the password requirements game are determined by the summed positive and negative effects of password implementations, including interaction effects. Quantification of each of these seven effects is difficult. The rationale is as follows, expressed in terms of costs and benefits; a summary is provided in Table 1.

protection benefit. The security-concerned player receives a strictly greater payoff from requiring some password than from requiring no password at all.

For a security-concerned Web site, the protection effect mitigates liability and convinces users they interact securely with the site.

data collection benefit. For a security-indifferent Web site, requiring any form of user accounts triggers the inflow of personal information that can be repurposed, as detailed in Section 2.

fortification benefit. Following a similar argument, the payoff for \mathcal{W}_s decreases if weak passwords are accepted ($S_s = \{wp\}$ or $S_s = \{wp, sp\}$), since these provide less protection against fraudulent access. The benefit from strengthening password protection applies to \mathcal{W}_s only.

development costs. Any password scheme will require resources to be developed and/or deployed. Password implementation costs are nil if no passwords are collected, they are low if passwords are collected but no restrictions are enforced.

strength assessment costs. More restrictive password schemes require technical skill; higher implementation costs make payoffs decrease. Implementation costs do not differ in the required strength (only weak versus only strong), since only allowing strong passwords is equivalent to filtering out weak passwords.

sharpness benefit. If \mathcal{W}_s is the only player who enforces passwords (isolated deployment), it can make people care more about passwords and there is no obtunding effect. The sharpness effect eliminates the tragedy of the commons.

negative externality costs. The payoff for \mathcal{W}_s is severely affected if passwords created with \mathcal{W}_i are also permissible at \mathcal{W}_s and vice versa, as detailed in Section 3.2. We note that the positive sharpness effect will not be observed together with the negative externality.

The negative externality effect on payoffs captures the interaction of password practices: an indifferent company \mathcal{W}_i as described above will take inadequate measures to protect user account details, because the site has no incentive to do so. The interest an indifferent site has in the data is mainly their suitability for advertising or resale. Leaking that data through security holes corresponds to giving the data away for free, which would still not be a direct problem for the indifferent Web site. Subject to compatible password requirements, users may reuse passwords created with \mathcal{W}_i for \mathcal{W}_s ($S_i \subseteq S_s$) or vice versa ($S_s \subseteq S_i$). The negative externality of bi-directional reuse thus dissipates when sets of permissible passwords overlap ($S_i \cap S_s \neq \emptyset$).

payoff component	applies to	if own strategy	if opponent's strategy
<u>d</u> development –	$\mathcal{W}_s, \mathcal{W}_i$	$\{wp\}, \{sp\}, \{wp,sp\}$	<i>any</i>
<u>s</u> strength assessment –	$\mathcal{W}_s, \mathcal{W}_i$	$\{wp\}, \{sp\}$	<i>any</i>
c data collection +	\mathcal{W}_i	$\{wp\}, \{sp\}, \{wp,sp\}$	<i>any</i>
p protection +	\mathcal{W}_s	$\{wp\}, \{sp\}, \{wp,sp\}$	<i>any</i>
f fortification +	\mathcal{W}_s	$\{sp\}$	<i>any</i>
i sharpness +	\mathcal{W}_s	$\{wp\}, \{sp\}, \{wp,sp\}$	$\{\}$
<u>r</u> neg. externality –	\mathcal{W}_s	$\{wp\}, \{sp\}, \{wp,sp\}$	<i>any overlapping</i>

Table 1. Summary of costs (–) and benefits (+) summing up to the payoffs in the password game. Summands negative in value are underlined. The last two columns enumerate the player's own and the opponent's strategies for which the respective effect applies.

4.3 Security-indifferent Web sites \mathcal{W}_i

For the security-indifferent Web site, player \mathcal{W}_i , strategies $\{wp\}$ and $\{sp\}$ are dominated by $\{wp,sp\}$, because $d + c - |s| \leq d + c$. A security-indifferent Web site either implements no password scheme at all ($S_i = \{\}$), or a scheme that accepts weak and strong passwords alike ($S_i = \{wp,sp\}$). The latter will be beneficial if the opportunity costs of missing out on collecting personal information from Web users prevail over deployment costs of a password scheme.

Ignoring future maintenance, password deployment can be thought of as an investment for \mathcal{W}_i : initial costs of d , result in accruing inflow of personal information. This data needs to be monetised to amortise the investment. c would therefore be the expected net present value from turning collected data into profits.

Deploying no password scheme will become the profit-maximising strategy if, first, \mathcal{W}_i is unable to monetise the personal information it collects through own or third-party use, or second, if personal information can be acquired without requiring passwords.

4.4 Security-concerned Web sites \mathcal{W}_s

Given that \mathcal{W}_i will play $\{\}$ or $\{wp,sp\}$, the security-concerned Web site \mathcal{W}_s will always prefer $\{sp\}$ over $\{wp\}$ as it provides extra fortification f . This preference would only be affected by the relative impact of r , the negative externality from password reuse, which does not make a difference except when \mathcal{W}_i would play $\{sp\}$ —a dominated strategy. Quite intuitively, allowing weak passwords only is a dominated strategy for Web sites sensitive to securing their systems.

As its opponent plays neither $\{sp\}$ nor $\{wp\}$, the difference in payoff for \mathcal{W}_s between $\{sp\}$ and $\{wp,sp\}$ lies in $s + f$, that is the extra profit from strong passwords versus the costs to identify and to enforce them. An argument could even be made that this sum actually is a negligible quantity: if the effort in

identifying strong passwords is expertise rather than the technical implementation, the difficulty lies in determining what non-trivial characteristics a strong password should have. Then, if f materialises in the form of reduced liability for the Web site when strong passwords are used, this is a result from a regulation prescribing “strong passwords” which would also hint at how to make technical systems compatible with the strength requirements.

4.5 Equilibria

In the light of the foregoing argument, two kinds of market equilibria are expected. First, the security-indifferent Web site accepts any password, and the security-concerned Web site requires strong passwords and potentially allows weak passwords as well, depending on the relative importance of the fortification effect compared to costs of assessing and enforcing password strength. This group of equilibria is inefficient due to negative externalities from password reuse. The second group of equilibria is reached if the security-indifferent Web site renounces password collection. Again, the concerned Web site may enforce strong passwords or accept weak and strong passwords. Collectively, this group of equilibria is socially better than the first (Pareto-superior): because $S_i = \{\}$ instead of $S_i = \{wp, sp\}$ requires $c + d \leq 0$ by definition, and in addition it holds that $i \geq 0 \geq r$, the social costs are lower if no password scheme is put in place at indifferent Web sites.

We now assume the likely case that benefits of protection are high for the security-concerned Web site (p is very high and f outweighs s). We also assume the indifferent Web site is similarly keen on collecting passwords ($c+d$ is positive). The password game then exhibits a single Nash equilibrium. This equilibrium is in dominant strategies, $(S_s, S_i) = (\{sp\}, \{wp, sp\})$. There is a unique combination of strategies that maximises social welfare, the sum of payoffs: $(\{sp\}, \{wp\})$; the payoff for the security-sensitive site is maximised in $(\{sp\}, \{\})$.

The Nash equilibrium is found, when the security-sensitive site enforces strong passwords, and also the indifferent site requires passwords but any password—weak or strong—will do there. In the Nash equilibrium, \mathcal{W}_i will receive a high payoff, since personal information (which translates to a revenue stream for the indifferent site) is collected, but little development effort is necessary. As the indifferent site will accept any password, weak or strong ($\{wp, sp\}$), practically no validity checking needs to be performed on the user-chosen passwords, eliminating the development costs s . The security-sensitive firm suffers from the negative externalities r this behaviour dissipates: due to password reuse, passwords from the security-sensitive site spill over to the indifferent site, where they are then subject to inadequate protection.

The security-sensitive site would get a maximum payoff if no indifferent site implemented any passwords ($S_i = \{\}$). Both the fortification effect and absence of negative externalities make this market outcome attractive for \mathcal{W}_s . Its only costs are the implementation effort for making only strong passwords acceptable. However, an indifferent site will not renounce password schemes since they act

$\frac{\mathcal{W}_s}{\mathcal{W}_i}$	$\{\}$	$\{wp\}$	$\{sp\}$	$\{wp,sp\}$
$\{\}$	$\frac{0}{0}$	$\frac{0}{c+\underline{d}+\underline{s}}$	$\frac{0}{c+\underline{d}+\underline{s}}$	$\frac{0}{c+\underline{d}}$
$\{wp\}$	$\frac{p+\underline{d}+\underline{s}+i}{0}$	$\frac{p+\underline{d}+\underline{s}+r}{c+\underline{d}+\underline{s}}$	$\frac{p+\underline{d}+\underline{s}}{c+\underline{d}+\underline{s}}$	$\frac{p+\underline{d}+\underline{s}+r}{c+\underline{d}}$
$\{sp\}$	$\frac{p+\underline{d}+\underline{s}+f+i}{0}$	$\frac{p+\underline{d}+\underline{s}+f}{c+\underline{d}+\underline{s}}$	$\frac{p+\underline{d}+\underline{s}+f+r}{c+\underline{d}+\underline{s}}$	$\frac{p+\underline{d}+\underline{s}+f+r}{c+\underline{d}}$
$\{wp,sp\}$	$\frac{p+\underline{d}+i}{0}$	$\frac{p+\underline{d}+r}{c+\underline{d}+\underline{s}}$	$\frac{p+\underline{d}+r}{c+\underline{d}+\underline{s}}$	$\frac{p+\underline{d}+r}{c+\underline{d}}$

Table 2. Algebraic expression of payoffs for the row player (\mathcal{W}_s , above the line) and the column-player (\mathcal{W}_i , below the line). Letters denote the cost/benefit effects from Table 1.

as a means to collect email addresses, for instance, and thereby realise positive payoffs c .

Comparing the Nash equilibrium with the social optimum is instructive. Both are Pareto-optimal; not both companies can improve their payoffs at once, compared to these strategy combinations. Maximum differentiation in the password requirements space maximise social welfare: the indifferent firm accepts only weak passwords and the security-sensitive firm accepts only strong passwords.³ The reason this equilibrium is not realised is the disincentive for the indifferent site to invest in restricting permissible passwords. The combined payoff in the social optimum is higher by $|s - r|$ than in the Nash equilibrium.

Interestingly, if the security-sensitive site would subsidise the indifferent site to help it in developing desirable password schemes which only accept weak passwords, the social optimum could be achieved. \mathcal{W}_s would need to transfer an amount between $|r|$ and $|s|$ to \mathcal{W}_i . Alternatively, \mathcal{W}_s could lower the development costs for \mathcal{W}_i , to make this site prefer $S_i = \{wp\}$ over $S_i = \{wp,sp\}$. Whilst this approach is equivalent to a subsidisation, it provides a technical rather than a transfer solution to the market failure. Indeed, the security-concerned Web site has already invested s in a password-strength assessment technology. These alternatives for regulation are discussed in greater detail in Section 6.

³ The reversed maximum differentiation yields a lower overall payoff since the fortification effect f does not apply to \mathcal{W}_i .

5 Password implementers: empirical evidence

Analysis of the password practices at 150 public Web sites in the areas of electronic commerce, identity services such as emailing and social networking, and content services such as news, reveals market-wide technical and organisational shortcomings [5]. While problems are widespread, there is evidence that weak password practices are more common in the news industry and lower-tier Web sites. This observed dichotomy is confirmed in a larger sample of 2184 automatically analysed public Web sites.

5.1 Notes on the datasets

Password practices in the wild are quantitatively assessed using two datasets, called the “Password Thicket” dataset and the “BugMeNot / Alexa” dataset respectively.

The *Password Thicket* dataset comprises 150 manually surveyed Web sites, equally sampled from mainstream and less popular Web sites [5]. The sample is organised by industry into three equally sized groups: identity sites (which use passwords to protect a user’s identity for interacting with other users, notably webmail and social networking), electronic commerce sites (designed for purchasing goods with no interaction with other users), as well as news and content sites (using passwords to customise site layout or limiting access to account holders only). Our methodology to assess the quality of password implementations was based on manual inspection of the security measures during enrolment (account creation), login, password reset, and password recovery. The details of our experimental setup and our main technical findings are reported elsewhere [5]. This dataset is characterised by the depth of investigation for each Web site in the sample.

The “BugMeNot / Alexa” (*BMN*) dataset comprises 2184 Web sites listed on www.bugmenot.com. BugMeNot is a major credential-swapping site “created as a mechanism to quickly bypass the login of Web sites that require compulsory registration and/or the collection of personal/demographic information”, a practice considered a “pointless” exercise [1]. Users of BugMeNot can upload new username/password combinations for a given Web site. Fellow visitors of BugMeNot can retrieve these credentials for free. Web site operators can request to be removed from BugMeNot, and it is possible to check which sites have taken this step. BugMeNot explicitly bans pay-per-view sites, community sites for interaction amongst members, sites with a fraud risk due to banking/commerce details stored with the user accounts. When browsing the BugMeNot repository, sites which were blocked can be told apart from sites for which no passwords have been submitted yet.

We crawled BugMeNot in a two-step process. First, the 3172 most popular Web sites were retrieved from Alexa Top Sites, an Amazon Web service. Popularity is measured by the proprietary Alexa rank to which traffic rank is a paramount ingredient. Popularity in the USA is used rather than the international ranking to avoid having localised versions of the same site showing up

twice (e.g., google.de on position 15 and google.com.hk on position 16). For each Web site in this top list, the Alexa Web Information Service was queried for the date the site went online, for the median load time, for whether it contains adult content, and for the first three categories this site pertains to.

A Web site’s category correspond to the classification in the Open Directory Project. These categories are used for coarse, keyword-based binary identification of news/weather/magazine Web sites. Whilst classification was automatic, one of the authors and another non-expert but skilled rater also classified a random subset of 388 of the sites (12%). These raters agreed on the classification at an intraclass correlation coefficient 0.64 (model 2, single measure). Although agreement with the automated classification was at an ICC of 0.49 and 0.42 respectively, the programmatic rating was judged acceptable, since inspection of diverging ratings indicated that manual classification resulted in a broader set of news sites (i.e. the automated rating was typically stricter). In summary, 274 sites are classified as news sites, 2882 as non-news sites; for 16 sites, the category information was missing and no classification was performed.

In a second step, each of the top Web sites was matched mechanically against BugMeNot. The listing status was recorded as one of ‘ok’ (there are accounts listed for this sites), ‘blocked’ (the Web site is barred from BugMeNot), or ‘missing’ (the Web site is not listed on BugMeNot). Missing Web sites may simply have never been registered by the community and never been blocked by the site operators, or may not collect passwords at all. For the sites in our survey, the majority of missing Web sites appear to not collect passwords. In total, 988 of the top sites were not listed in BugMeNot; the remaining 2184 sites making up the *BMN* dataset are divided into 531 ‘blocked’ Web sites and 1653 ‘ok’ Web sites. In summary, the *BMN* dataset is characterised by the breadth of investigation and focuses on the most popular destinations on the Web.

5.2 Practices at content sites

Content sites are the prototypical example of security-indifferent Web sites that use passwords as a trigger for harvesting profile information (Section 2). They are significantly more likely to collect personal information at the time when a new password is created. All but 2% collect email addresses on a mandatory basis and they are more likely to verify these with very high significance.

Password carelessness is more prevalent at news sites with high significance. This is manifested in very significantly lower adoption of TLS to protect passwords and sending the cleartext passwords in a welcome email or “forgotten password” responses. Also, news sites place fewer limits on guessing passwords, so they could be abused as password oracles in brute-force attacks against more security-sensitive sites. Users of news sites are rarely given password advice or hints on how to make a password more secure, for instance by including digits or via a graphical password strength indicator.

5.3 Practices at sites with merchant facilities

Merchant facilities are not exclusive to e-commerce sites but may be used at other Web sites as well. Sites which store users' payment details perform significantly better on overall password security and in several key measures, including TLS deployment and notification to users about password reset events. Although the lack of encrypted data transmission is widespread, the dichotomy between content sites and e-commerce sites is very strongly significant and TLS deployment is strongly correlated with merchant facilities. It holds that sites which process and store fraud-prone details such as payment information take more care in handling passwords securely and offer more advanced security features overall.

Critically, we observe that sites with merchant facilities are also significantly more likely to impose minimum password lengths and to blacklist common passwords. If sites with merchant facilities are interpreted to represent the class of security-interested Web sites, this directly supports the predictions of the password game.

5.4 Prevention of credential sharing

Web site operators' eagerness to stop the sharing of credentials amongst their users can be interpreted as an indication of whether or not password compromise is deemed a serious risk. In merging BugMeNot data with Alexa traffic rank data, we observe that very-high traffic Web sites are much more likely to block listing of their credential of BugMeNot (Figure 1; 28% blocked in upper half versus 20% in lower half, highly significant).

After the top 50 Web sites, there is a small and gradual decrease in the rate of blocking. However, blocking is still quite common among the lower ranked sites, at above 25% in the band of the 500th to 1000th most popular sites, and 17% around rank 3000. Password collection remains at a high level of above 80% through the entire range up to the 1000th rank and is still at 69% for up to the 3000th rank. Thus, across a range of Web sites BugMeNot blocking is a useful indicator of Web sites' real security motivations for collecting passwords.

We see a pronounced trend towards news Web sites blocking BugMeNot sharing significantly less often. We observed 18% of non-news sites to block BugMeNot sharing, while only 8% of news Web sites did so. If we restrict ourselves to only Web sites which exist in BugMeNot's database the divide is even stronger, with 26% of non-news Web sites blocking BugMeNot sharing and only 9% of news Web sites doing so. Both results are highly statistically significant ($p < 0.0001$ using a two-tailed G -test). Thus, the BugMeNot data provides strong evidence that news Web sites do have lower security concerns than other password-collecting Web sites.

In summary, a Web site's password practices improve with popularity, industry-specific leadership, and growth. Password deployments are prevalent across popularity strata, but the most popular sites are also most likely to exhibit password care in taking measures to prevent password sharing amongst users.

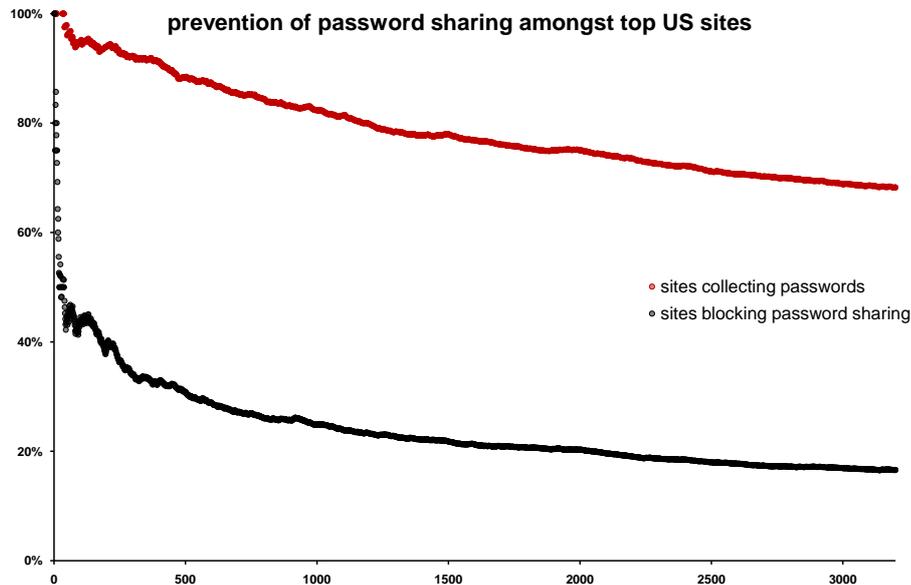


Figure 1. Proportion of sites collecting passwords (upper curve) and amongst these, the proportion of sites blocking password sharing (lower curve). Ratios given for top k US sites with k up to 3200. Bumps are artefacts of the increasing window size for the arithmetic mean.

5.5 Limitations

Despite these consistent results, we acknowledge there are limitations inherent to our empirical and theoretical methodologies.

First, the game-theoretic model may not have captured all intricacies of the players' payoffs. A notable simplification is the symmetry in the cost structures. It seems convincing that security-concerned Web sites exhibit a structural difference compared to other Web sites which would lower their costs for developing standard or sophisticated password schemes. Yet, the current model would then be an underestimation of the resulting inefficiency on the market. A more fundamental question seems to be whether password deployment decisions can be modelled as a game which has the underlying assumption of rational players.

Second, insofar as we base our analysis of data retrieved from Alexa and BugMeNot, we are limited by our resources in fully ascertaining their accuracy. All popularity ranks are based on Web usage behaviour from a sample of users who are not representative of the entire online population. We nonetheless deem the rankings reliable. Further, the BugMeNot database suffers from quality problems inherent to a crowd-sourced endeavour. Although we did not systematically probe the accuracy of the BugMeNot repository, we have tried to reduce recording errors by a full manual inspection of suspicious cases.

6 Observed market inefficiency and pathways for regulation

The ubiquity of password schemes on the Web contrasts with the variety of security flaws found in implementations across industries and Web site popularity levels. Failure to provide adequate password security cannot be attributed to technical problems alone; mis-aligned incentives on the market explain why some groups of Web sites score significantly better on password security.

In the early 1960s, the former US security advisor McGeorge Bundy noted: “if we guard our toothbrushes and diamonds with equal zeal, we will lose fewer toothbrushes and more diamonds”. In 2010, the Windows Live help pages echo: “Access to your financials and email account is a critical thing; therefore your passwords for these accounts should be un-guessable, even by a computer. If it is to post a response on a gamer forum, perhaps it doesn’t need to be so complex.” [3]

We have presented empirical evidence that password-equipped Web sites fall in two broad categories: on the one hand, there are those sites which have a self-interest to invest in security, on the other hand, some sites only use passwords as a trigger to collect personal information from their users for secondary purposes. Password carelessness is significantly more prevalent amongst the latter and encompasses phenomena such as sending out passwords in cleartext, allowing unlimited password guessing, lack of secure data transmission, and failure to fight sharing of credentials.

The observed lack in technical safeguards can be explained by a lack of business interest to invest in password security. At content sites, passwords do not serve a genuine authentication purpose but rather vest information profile creation with credibility. The resulting dichotomy in the market corresponds to the prediction of a game-theoretical model where a security-indifferent and a security-sensitive Web site operators make password deployment choices. Their differing motivations lead to differing optimal choices as to what combination of weak and strong passwords should be accepted, bearing in mind the required implementation efforts as well as increased levels of protection.

The password game exhibits a unique Nash equilibrium where security-indifferent Web sites will accept any password regardless of its security but security-sensitive Web sites allow strong passwords only. This outcome can also be observed empirically on the Web. It is not socially optimal, since the overlap in permissible passwords creates room for password reuse. Overuse of consumers’ mental ability to remember passwords is similar to the tragedy of the commons and makes consumers use the same passwords across sites, which differ in security practices. Careless handling of strong passwords at a security-indifferent site dissipate a negative externality on security-sensitive Web sites, where leaked or fraudulently acquired credentials can be reused by an attacker. As a result, the market allocation of password strength is inefficient.

Due to the negative externalities from password reuse, Web sites for which security is critical find themselves exposed to threats created by not accounted for by security-indifferent sites. This calls for regulation. Effectively controlling

password deployments and their strength by a global, supra-national Internet authority remains unrealistic. Certification could become mandatory for new deployments, which are then watermarked so that any alterations break the certificate. Obviously, exactly those Web sites currently exhibiting weak password practices would have little incentive to participate.

Password deployment needs to be priced to internalise the externalities on other password implementers and consumers. Pricing need not be direct, although a pay-per-stored-password is the most direct solution and could be realised through charged-for database storage as a service (similar to the emailing facility offered by Amazon Web services for instance). Requiring yearly, printed data statements sent to users would involve costs that reduce the incentive to collect passwords without genuine authentication purposes [8]. These approaches could be complemented and supported by targeted legislation. A purely legal approach may increase the costs of leaking passwords by making the weaker-security Web site liable for account breaches at higher-security Web sites. Regulation could further reduce the ability for Web sites to monetise personal information collected via password schemes and thus make their deployment less attractive. Security-sensitive sites could also be given the opportunity to sue security-indifferent Web sites for unfair competition, creating a dynamic similar to the one that made German Web sites observe the requirement of an imprint with contact details. Abiding by a password scheme would not preclude content Web sites to tailor their services on a per-account basis; federated identities such as OpenID provide a viable and mutually beneficial alternative [13]. Insofar as password collection is merely a trigger to ask users for personal details, programmatic access to people repositories such as Facebook Connect makes this anchor redundant and may reduce the proliferation of password collection.

On the technical side, password kits could lower the costs for password-implementers to secure their systems. For security-indifferent Web sites, it should be cheaper to use an out-of-the-box password solution with known, weak security than implementing its own, arbitrarily secure system. Security-sensitive Web sites would have an incentive to sponsor the development and provision of such support tools, as demonstrated in Section 4. They would also have the beneficial side-effect of unifying the password experience for consumers and could potentially be branded. Open questions as to how such a password kit should look precisely and what password standards should be incorporated are important research challenges. Our limited consensus of what parameters constitute an appropriate password scheme calls for further, large-scale experimental research.

Acknowledgements

Katarzyna Krol provided valuable help in data collection and enhancement.

References

1. BugMeNot, Feb 2010.
2. Facebook Connect. 2010. <http://www.facebook.com/advertising/?connect>.
3. Windows Live Solution Center: Creating a strong password for your e-mail account. <http://windowslivehelp.com/solution.aspx?solutionid=3ca67154-2ee7-4da4-8b95-f8aef17a71bc>, September 2010.
4. Yahoo! Password Help. <http://help.yahoo.com/l/us/yahoo/abuse/password/faq.html>, September 2010.
5. Joseph Bonneau and Sören Preibusch. The password thicket: technical and market failures in human authentication on the web. In *The Ninth Workshop on the Economics of Information Security (WEIS 2010)*, 2010.
6. Bundesamt für Sicherheit in der Informationstechnik (BSI) (Federal Office for Information Security). *IT-Grundschatz Catalogues*. 2005.
7. William E. Burr, Donna F. Dodson, and W. Timothy Polk. Electronic Authentication Guideline. *NIST Special Publication 800-63*, April 2006.
8. Chaos Computer Club (CCC). Datenbrief. <http://www.ccc.de/datenbrief>, January 2010.
9. Dinei Florêncio and Cormac Herley. A large-scale study of web password habits. In *WWW '07: Proceedings of the 16th international conference on World Wide Web*, pages 657–666, New York, NY, USA, 2007. ACM.
10. Shirley Gaw and Edward W. Felten. Password Management Strategies for Online Accounts. In *SOUPS '06: Proceedings of the Second Symposium on Usable Privacy and Security*, pages 44–55, New York, NY, USA, 2006. ACM.
11. Gilbert Notoatmodjo and Clark Thomborson. Passwords and Perceptions. In Ljiljana Brankovic and Willy Susilo, editors, *Seventh Australasian Information Security Conference (AISC 2009)*, volume 98 of *CRPIT*, pages 71–78, Wellington, New Zealand, 2009. ACS.
12. Brian Prince. Twitter Details Phishing Attacks Behind Password Reset. *eWeek*, January 2010.
13. David Recordon and Drummond Reed. OpenID 2.0: a platform for user-centric identity management. In *DIM '06: Proceedings of the second ACM workshop on Digital identity management*, pages 11–16, New York, NY, USA, 2006. ACM.
14. Shannon Riley. Password Security: What Users Know and What They Actually Do. *Usability News*, 8(1), 2006.
15. Ashlee Vance. If Your Password Is 123456, Just Make It HackMe. *The New York Times*, January 2010.