"Ten strikes and you're out": Increasing the number of login attempts can improve password usability

Sacha Brostoff
Department of Computer Science
UCL (University College London)
London, WC1E 6BT, UK
+44 20 7679 3039
s.brostoff@cs.ucl.ac.uk

M. Angela Sasse
Department of Computer Science
UCL (University College London)
London, WC1E 6BT, UK
+44 20 7679 7212
a.sasse@cs.ucl.ac.uk

ABSTRACT
Many users today are struggling to manage an increasing number of passwords. As a consequence, many organizations face an increasing demand on an expensive resource – the system administrators or help desks. This paper suggests that re-considering the "*3- strikes"* policy commonly applied to password login systems would be an immediate way of reducing this demand. We analyzed 10 weeks worth of system logs from a sample of 386 users, whose login attempts were not restricted in the usual manner. During that period, only 10% of login attempts failed. We predict that requests for password reminders could be reduced by up to 44% by increasing the number of strikes from 3 to ten.

Keywords
Computer security, passwords, security policies, human memory, user-costs, audit trails

INTRODUCTION
Many users are struggling to manage an increasing number of passwords [4] – the most commonly used authentication mechanism for computer security today. A policy commonly applied to password mechanisms is the "*three strikes"* policy, meaning a user's account is locked after 3 failed login attempts. Some organizations employ less drastic penalties - such as disabling logins to the account for a short period. This penalty represents at best an annoyance for legitimate users prevented from logging in, and can at worst result in a significant disruption to their work. After the account being suspended, users need to contact a system administrator or helpdesks to have their password *reset*-the user is given a new system password, and asked to substitute it with one of her own choice – but which conforms with a number of restrictions designed to ensure the password is secure. Re-setting a password therefore consumes considerable user and company resources: time and effort not spent on production tasks, which may involve a customer waiting to be dealt with. Furthermore, it creates a high mental cost for the individual user, to select and then memorise the new password. For a certain amount of time after the re-set, this new password has to "compete" with the now defunct one in the users' memory [4]. Furthermore, there is no perceived benefit for individual users in exchange for all the effort they have to expend. In the longer term, such policies can foster negative perceptions of computer security, which predisposes users to subvert security mechanisms [1]. Given the negative impact the policy has on users, and considering that it offers no additional protection against *cracking* (which most organizations perceive to be the biggest threat to password security), we feel it is time to question the validity of the policy. The security community cannot provide a rationale as to why 3 failed attempts is the right cut-off point. The few discussions of the policy that can be found [e.g. 5] are not based on empirical evidence. This paper introduces a methodology that can be used to inform such discussions, and presents the first results of its application.

LEAVE BLANK THE LAST 2.5 cm (1") OF THE LEFT COLUMN ON THE FIRST PAGE FOR THE COPYRIGHT NOTICE.

METHODOLOGY
The participants in the study were 386 undergraduate students enrolled on Computer Science courses at UCL. We examined system logs of their accounts on a Web-based coursework system [described in 2] over the duration of an academic term (10 weeks).

Participants logged to practice coursework questions and submit their answers to web-based multiple-choice assignments they had to complete for course credit. The participants were allowed to practice coursework questions as often as they desired, but were allowed to submit each coursework exam only once. System logs recorded every successful and unsuccessful login attempt, as well as use of the *password reminder* facility. The reminder facility e-mails users their passwords on request – a practice which harbors significant security risks, and is thus not viable in many organizational contexts. In most organizations, accounts are locked after failed attempts, and system administrators or system helpdesk have to re-store accounts for users. In many organizations, the cost of running these

helpdesks has rocketed with the proliferation of systems requiring password authentication [4].

RESULTS

Out of 386 participants, 34 (9%) required a reminder of their password over the 10-week period. Table 1 shows that the average failure rate for passwords was one login failing per 10 attempts (10%). Approximately 7% of these failed logins led to password reminder requests. This means that in organization where passwords are re-set through helpdesks, approximately 0.7% of login attempts can be expected to result in a helpdesk call.

Table 1–Login success and failure

	Number of				
Measure	Login attempts	Successful logins	Failed logins	Reminders requested	Login failure rate
Total	13305	12044	1261	87	N/a
Average	34.5	31	3.3	0.23	0.10
Min.	1	0	0	0	0
Max.	348	339	71	9	1
St.dev.	35.5	32	7.6	0.87	0.16
N. of people	386	386	386	386	386

Figure 1 shows the distribution of login failures: the light bars represent the login failures of participants who used the password reminder facility, and the dark bars represent the login failures of participants who did not need reminders. Figure 1 shows that participants who required password reminders suffered proportionately greater numbers of failed login attempts-for example 5% of this group had 14 failed logins.

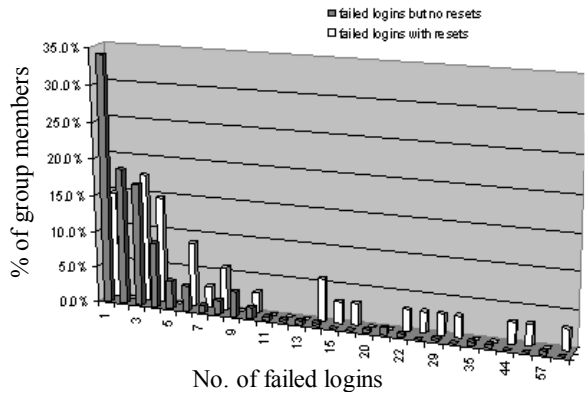

Figure 1-Distribution of numbers of login failures, for users who did/did not require password reminders.

Table 2 shows descriptive statistics about the relative numbers of failed login attempts for the two groups. N=236 because the other 150 participants did not have any login failures, and so could not be included in the table. Participants in the table who used password reminders experienced on average three times as many failed login attempts as participants who did not. This difference is highly significant ($F_{1,234}$=28.7, p<.001, eta squared = .109 , power = 1).

Participants who required password reminders experienced an average of 6.9 failed login attempts per reminder, with a standard deviation of 7.5, and an average number of 2 calls to the helpdesk.

Table 2 Failed login attempts, for participants who did/did not require reminders

Group	Mean failed logins	Median failed logins	Min failed logins	Max failed logins	Std. Dev. failed logins	N. of people
No reminders group	4.1	2	1	57	6.5	202
Reminders group	12.7	6	1	71	16.3	34
Total	5.4	3	1	71	9.1	236

To perform our analysis, we have assumed that the users in gave up in trying to login after having all their failed login attempts in one session. This assumption has certain consequences that will be discussed in a later section: *problems and risks and for the methodology*.

DISCUSSION

Why not allow users 10 attempts?

There are 3 immediate benefits from increasing the number of attempts users are allowed.

1. It reduces the demand on an expensive resource - system administrators or help desks.
2. Not having to change a password reduces the mental load on users, and
3. reduces the time taken away from, and interference caused with, users' production tasks.

We suggest that 2. and 3. lead to a fourth benefit:

4. prevents an erosion of the respect with which users hold security, and improve compliance with important security roles.

Organizations need to weigh these benefits against the increased vulnerability to internal attackers trying to guess another users' password.

Benefits 2. & 3. only apply to the group of users that recovered from login failure, since the other group was doomed to forget their password anyway. However, the fourth benefit applies to both groups, who may perceive that they've been unnecessarily forced to change their password by three strikes will. We will deal with these separately below.

User costs

For this issue we must look at the distribution of failed logins among people who did not later require helpdesk support (Figure 1). The existing norm of *three strikes* is predicted to work for 107 of the 202 users in the group who recovered and succeeded to login (53% of this group, or 28% of all participants). Though 107 would be passed, the remaining 95 users in this group (47% of this group, or 25% of all users) would be penalised by a three strikes rule because they took more than three attempts to successfully login. The proposed norm of *ten strikes* would be enough to accommodate the vast majority of users who recovered from login failure (187 users, 93% of this group, 48% of all participants), penalizing the remaining 15 people in the group.

Assuming that these figures are correct, what would be the impact on helpdesk use? With no restrictions (equivalent to an ***unlimited strikes*** policy) we observed a **baseline of 87** helpdesk requests-which were due to the group who could not recover from their failed logins. If we moved to a ***ten strikes*** policy we would expect a further 15 helpdesk requests (see previous paragraph), making a total of **102 requests**. However, if we moved to a ***three strikes*** policy we would expect to see 95 requests above the baseline (see previous paragraph), leading to a total of **182 requests**. Thus, moving from 3 to ten strikes could slash these predicted password-related helpdesk calls from 182 to 102 - a 44% reduction. Given the cost involved in running helpdesks, this represents a significant saving.

User perceptions

To this issue, we must also look at the distribution of people who experienced login failures that *did* result in helpdesk requests (light bars in Figure 1). 11 people out of 34 in this group would have survived a *three strikes* policy, leaving 23 people struck out. Extending the number of strikes to ten would double the number of people accommodated to 23, leaving only a third of this group (11 people) prematurely forced to call a helpdesk.

The figures are much more dramatic when we add in the users from the previous section who would have been all right if given enough chances, and so who had a valid complaint. With *three strikes* the number being unfairly curtailed is 118 (23+95) out of 386 (31%). With *ten strikes* it is only 26 (11+15) (7%).

The risks of 10 attempts

By increasing the number of strikes, you increase the chance that an internal adversary may successfully guess the password and gain access to another users' account. Moving from three strikes to ten approximately triples this risk. However, if an organization promotes strong password content policies – which are needed to counteract the external threat of password *cracking* - then the actual risk will still be very small. Moreover, Viega and McGraw [4] suggest that there be another strike counter operating in conjunction with the first, recording the total number of strikes rather than the number of strikes in a session. After a suitably small total number of strikes is reached, such as 200, then additional security procedures are started. This would help to reduce any negative impact of moving from three strikes to ten.

Limitations of the study

In this study, we have equated the number of password reminders with login problems that require helpdesk support. This in effect overestimates the number of login failures experienced before requiring helpdesk support, and the number of attempts users have to be allowed to reduce helpdesk load is probably lower than 10.

We also have to consider that participants in this experiment had no password restrictions or policies placed upon them. Users facing a set of policies governing their passwords may behave differently. For example, users under a *three strikes* policy who need to contact a help desk are likely to use password prompts or caches (passwords kept in their diary, PDA, post-it notes, etc.) to avoid calling the help desk – especially since some organizations, in the face of rising help desk cost, have taken to reprimanding users who draw on the resource too often. These behaviors circumvent other common security policies, creating significant security risks [1].

Even though we did not require them to do so, participants in this study tended to choose cryptographically strong passwords, which are difficult to remember [5]. Many organizations enforce cryptographically strong passwords, so the rate of failure should be comparable.

Strengths of the research methodology

Another argument of this paper is that we need more data on performance of security mechanims in everyday use in order to make good design decisions. System logs – such as the ones used in this study – are a valuable tool to determine performance. Data collection is relatively simple. Specialized apparatus need not be necessary, but the authentication mechanism needs to be configured to record every authentication event. Analysis is relatively simple as well. Data preparation consists of counting the number of events of each type for each user. Event types are quickly, accurately and reliably distinguishable, as they are logged with different labels. Data analysis can be achieved using ubiquitous tools such as spreadsheets that offer pivot tables. This approach is therefore simple, applicable in many real world contexts, and with few resources, and can even be conducted by non-usability experts - such as system administrators - who have easy access to the data.

FURTHER WORK

We believe that the following work will provide valuable knowledge for improving password mechanisms:

Studying the effect of security policies-the present approach relies upon participants whose password use is not restricted by security policies. We identified a potential risk to the validity of our results that participants might behave differently if they were subject to common security policies, including three strikes. This study should be repeated to see the effects of these policies, including: requirement for strong password content, password expiry, multiple passwords, combinations of the above, and combinations of the above with unsynchronized password expiry. As the design of the study becomes more complex, we can become much more confident that it reflects the situation of real users.

Studying the effect of corporate populations-this is useful in two different ways: it enables us to make better predictions from student populations, and it can remove the necessity of complicated experimental design to achieve the results asked for in the previous recommendation for further work

Creating and disseminating do-it-yourself research tools - it should be possible to create research tools that dramatically reduce the workload for systems administrators who want to do this work themselves. The tools would be constructed for popular computing platforms including instructions on the configuration changes necessary to collect the data, and analysis templates into which systems administrators can drop the raw data, and have much of the analysis work done for them.

REFERENCES

1. Adams, A. & Sasse, M. A. Users are not the enemy. *Communications of the ACM*, Vol. 42, No. 12. December, 1999.
2. Brostoff, S. Improving password system effectiveness *Department of Computer Science*, UCL, London, in preparation.
3 Brostoff, S. and Sasse, A., Are Passfaces more usable than passwords? A field trial investigation. in *People and Computers XIV - Usability or Else*, (Sunderland, UK, 2000), Springer, 405-424.
4. Sasse, A., Brostoff, S. and Weirich, D. Transforming the 'weakest link' — a human-computer interaction approach to usable and effective security. *BT Technology Journal*, *19* (3). 122-131.
5. Viega, J. and McGraw, G. *Building Secure Software: How to Avoid Security Problems the Right Way*. Addison Wesley, 2001.
6. Zviran, M. and Haga, W.J. A comparison of password techniques for multilevel authentication mechanisms. *The Computer Journal*, *36* (3). 227-237.