You Have Three Tries Before Lockout - Why Three?

K. Renaud, R. English, T. Wynne and F. Weber

School of Computing Science, University of Glasgow
e-mail: {karen.renaud, 2rosanne.english}@glasgow.ac.uk; 3t.o.wynne@gmail.com;
40802085W@student.gla.ac.uk

Abstract

It is considered good practice to lock users out if they enter the wrong password three times. This is applied almost universally by systems across the globe. Three tries was probably considered a good balance originally between allowing the legitimate user to make some genuine errors and foiling an attacker. This rule makes sense intuitively yet there is no empirical evidence that three tries is the most efficacious number. It is entirely possible that the number should not be three, but some other number, such as two, five or even seven. It is very hard to test this since attempts could be either a legitimate user attempting to recall his/her password, or an intruder trying to breach the account. If an attacker is allowed more attempts one could imagine the system's security being compromised. Here we argue for the use of a simulation engine to test the effects of such password-related security measures on the security of the entire eco-system. A simulation approach expedites no-risk empirical testing. We use a simulator called SimPass, which models both user password-related behaviour and potential password-based attacks from within and outside an organization. We provide evidence of the expected security impact of increasing the prevalence of password sharing. That is it will lead to increased use of others' credentials and a lack of accountability. We then test different settings for locking of accounts after a certain number of failed authentication attempts to determine a potentially optimal setting. We find that a three times lockout policy might well be too stringent and deserves further investigation.

Keywords

Simulation, Passwords, Security Policies

1. Introduction

Companies increasingly rely on digital systems to run their businesses effectively. The digitisation of the work environment requires that efforts be made to protect such records from unauthorised access. Every user has to be identified and such an identity verified for the system to grant access to protected information. Most organisations deploy passwords. Due to forgetting, insecure coping behaviours which compromise the potential security of the mechanism are employed (Adams & Sasse, 1999; Gehringer, 2002; Herley, 2009; Inglesant & Sasse, 2010). Organisations respond to these insecure behaviours in two ways:

1) *Policing the Human*: Organisations write and enforce security policies which include sections forbidding a variety of insecure password behaviours and providing guidelines for good password practice (Lubbe & Klopper, 2005).

2) *Implementing Technical Controls*: to strengthen the security of the system. An example is to force users to provide a password that satisfies strength metrics (Grainger, 2002) or forcing regular password changes.

Composing and enforcing information security policies can be challenging for organisations (Mandajuno & Sota, 2004; Posthumus and Von Solms, 2004). One can obtain a generic policy but these need to be tailored to the requirements of the organisation (Da Veiga & Eloff, 2007). Policies have to be reviewed on a regular basis (Williams, 2001), but the real impact and potential side effects of policy changes, made in good faith, are hard to pin down and often only emerge later. Thus changes could unwittingly have a detrimental effect on the security of an eco-system.

It is almost impossible to test the effects of policy rules in a real life environment since causatives and behaviours are so complicated and the effects often difficult to detect. Von Solms and Von Solms (2006) argue that one should not include any directive in a security policy that cannot be measured. Yet some of the most commonly included password-related directives have their roots in legacy practices. For example, password sharing is forbidden yet how can one measure the incidence of this activity? The user of a shared password leaves no trace since the access manifests as legitimate use. Finally, it is a brave security officer who adjusts well-established mechanisms. If a security breach occurs subsequently, fingers might well be pointed in his/her direction. Moreover, he or she will have no evidence to prove that the change did not cause the breach. Hence in security many play it safe, adopting approved mechanism in order to ensure their own job security.

An alternative to testing policy changes in the wild is the use of a simulation engine. Simulation is a well-established approach (Simon, 1996) whereby a software model is abstracted from knowledge garnered from a set of observed real systems and then run with a range of input parameters of interest. A validated model will be able to test predictions across a generalised subset of the parameter space where conditions are similar to the validation points. Simulations are helpful in two ways: they can explain (in the sense of identifying a unifying model) retrospectively what has already been observed; more importantly, they can give insight into the functioning of systems of the modelled type, in particular, exploring possible side-effects in previously unexplored regions of the parameter space. The findings produced by a simulation must be confirmed by means of observation in a real environment. However, it does present a no-risk mechanism for testing changes and can help to predict the possible riskiness thereof.

2. Testing Security Controls

Organisations cannot realistically experiment with the relaxation of password control techniques in case the security of the organisation's systems becomes compromised. A simulation environment offers the opportunity to experiment safely and gain insights into the potential side effects of a relaxation, or adjustment, of a policy or technique.

Simulation has been used to good effect in other contexts (Scalese & Issenberg, 2005; Vaughn, 1995; Vickery et al., 2000; Reinhart & Fitz, 2006). The common theme in these usages is that simulations provide a quasi-environment that attempts to mirror the real-life environment in all its essential features. This environment supports no-risk testing and experimentation and can support knowledge discovery.

Technical security has benefitted from the use of simulations (Lee et al., 2005) but we have not been able to find evidence of simulations being used to inform design of secure socio-technical systems. The SimPass simulation engine was developed to emulate the behaviour of agents (employees both malicious and non-malicious, and hackers) in an organisation with a number of systems which the agents attempt to access over a period of time using a username and password combination (Renaud & McKenzie, 2013). On commencing the simulation, a number of agents and systems are generated. Each SimPass entity has a number of configurable attributes informing their behaviours, such as whether the agent is dishonest, malicious, or likely to share passwords. Systems can either issue passwords or allow users to choose their own, for example. Depending on the settings, the agent reacts in different ways when asked to authenticate at random intervals. Well-established forgetting statistics (Ebbinghaus, 1964) are used to make an agent forget passwords. Agents will also engage in particular coping behaviours to deal with the load that passwords impose on them, such as writing them down, reusing, recycling, using weak common passwords, sharing and stealing passwords. Further detail of SimPass (including the default configurations as used in this work) can be found in Renaud & McKenzie (2013).

By configuring the SimPass engine, we can examine the impact of a particular behaviour or policy setting. The outputs we are interested in for the purpose of this paper are the percentage of “bad logins” (an agent using another agent’s credentials), and the number of lock-out events (caused when an agent forgets a password).

We will test two commonly utilized techniques/policies for enhancing security, one from each of the categories mentioned in the introduction. The first appears in most organizational security policies (forbidding password sharing). The second is almost the default policy in most password implementations (locking accounts if a person enters an incorrect password a certain number of times, usually 3).

- *Password Sharing:* Evidence of password sharing abounds (Adams & Sasse, 1999). Sharing is frowned upon by security aficionados (Mandajuno & Sota, 2004; Lubbe & Klopper, 2005). Moreover, banks increasingly implement policies that free them of any liability should the password or PIN for a bank account be disclosed (Murdoch et al., 2010).

  On the other hand sharing has the potential to reduce wasted time. Moreover, it seems reasonable for departments to have a common password since they share roles and responsibilities. Singh et al. (2007) explored password sharing in banking. They concluded that there are good reasons to share passwords, such as if someone is suffering from a disability that does not allow them to go shopping for themselves or when the access to an ATM is restricted. The reasoning behind barring of sharing could be the ability to link transactions to individual employees (achieving non-repudiation). Hence this is essentially a risk mitigation technique.

- *Three times lock out*: Most systems allow people to make three faulty authentication attempts before locking them out of the system, and requiring them to contact system support to be granted access again. This policy allows the user to make a limited number of errors but resists the efforts of hackers. There is very little in the literature questioning this default setting. Brostoff and Sasse (2003), however, did investigate the widely used "3 strikes and you're out" policy. They conclude by advocating the use of 10 login attempts instead of 3 to reduce the workload for systems administrators and help desks. What their study does not investigate is the security impact of such an increase. Since this, too, is a risk mitigation technique, we would have to show a negligible risk increase if this limit were to be relaxed.

In both these cases the spirit of the security control is to reduce the number of "**bad logins**". These are system accesses where person A uses person B's credentials to access the system. This can happen if B willingly shares his or her credentials with A, or where A obtains them fraudulently or manages to guess them.

Hypotheses

The hypotheses for the effects of sharing are as follows:

> **H1:** As password sharing increases the number of lockout events will decrease.

> **H2:** As password sharing increases, there will be a significant increase in the percentage of bad logins.

The hypotheses for the effects of authentication attempt restrictions are as follows:

> **H3:** Allowing a larger number of authentication attempts before lock out reduces the number of users being locked out

> **H4:** Allowing a larger number of authentication attempts before lock out does not increase the number of bad logins

Hypotheses H1 is expected to be supported since if more people know a password, it is likely if one user forgets it they can ask their trusted colleague. We also anticipate H2 to be supported as a bad login is where a user B logs in as user A; if B is given a password it is assumed they will use it. We anticipate H3 will be supported as a user will have more attempts if they have forgotten their password, and H4 not to be supported as allowing more legitimate attempts could allow more illegitimate attempts (and hence successes).

3. Simulations

3.1. Effects of Sharing

To test the impact of sharing on system security we ran a number of simulations with a 0%, 25%, 50%, 75% and 100% sharing prevalence. (Independent variable: % sharing, dependent variables: number of lockout events & number of bad logins.). Actual sharing rates are hard to gauge but some studies suggest it in some contexts as many as two thirds of employees share passwords (Renaud, 2013).

We constructed histograms and Q-Q plots for both dependent variables at each level of sharing. Normality was demonstrated by approximate bell curves with no apparent skew in the histograms and an approximately straight line in the QQ plots. Figure 1 shows that increasing the prevalence of sharing does indeed reduce the number of accounts locked (ANOVA < 0.000; F=7649.9). This allows us to reject the null hypotheses for H1. There is also a statistically significant difference in terms of the number of bad logins (ANOVA < 0.000; F=2232.723) with increased sharing prevalence, the corresponding boxplot is shown in Figure. This allows us to reject the null hypothesis for H2 also.

The first two hypotheses are supported as expected. We will now test how increasing the number of tries before lockout would impact on system security and on end-users.

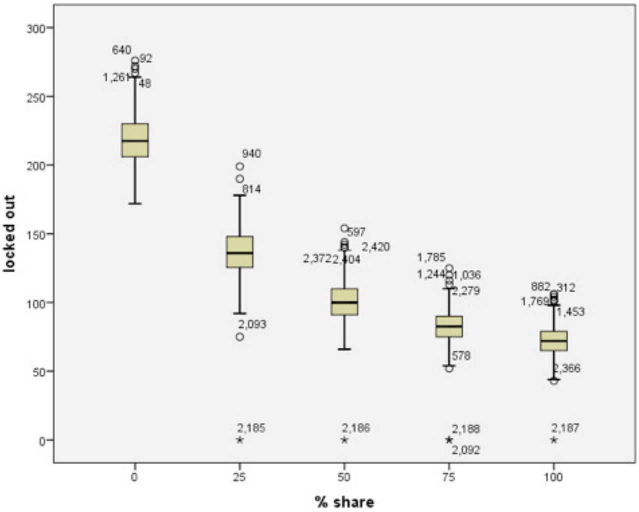

Figure 1: H1 Boxplot

3.2. Lockout after how many Tries?

We ran 500 simulations for each of 3, 5, 7, 9, 11 and 13 tries until lockout. (Independent variable: number of authentication attempts permitted before lock out; dependent variables: number of accounts locked for H3 and the number of bad logins for H4.)

Normality was established by examining the histograms and normal Q-Q plots for each level of attempts before lock out. The data follows the expected normal distribution well with the exception of one obvious outlier. The ANOVA results for H3 are presented in Table 1. Since the significance level is less than 0.05 we can deduce that there is a significant difference in the means of at least one group. The Levene statistic was calculated and the significance value was 0.173, meaning the homogeneity of variance is not violated and it is acceptable to use the Tukey HSD test to establish which groups have significant differences.

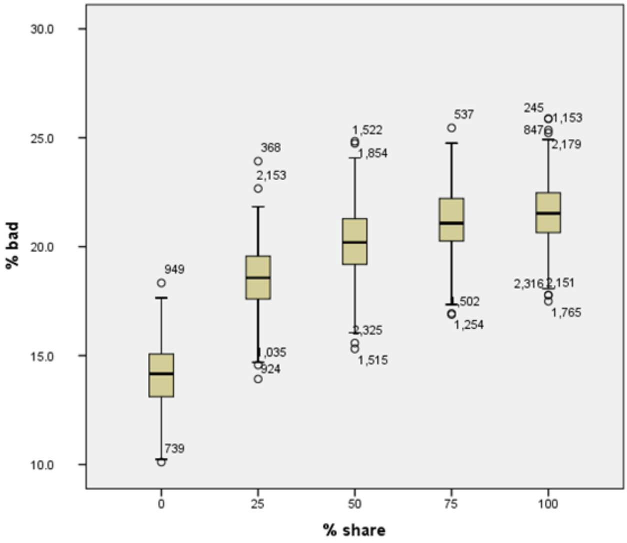

Figure 2: H2 Boxplot

The multiple comparisons tests using Tukey HSD showed significant differences in the group using a lock out rate of 3 compared to each of the other levels (5, 7, 9, 11, and 13). All other group comparisons showed no significant difference. The multiple comparisons table is presented in Table 1 and has been restricted to show only the significant groups. Examining the boxplot in Figure 3 we can see that the median number of locked out accounts is higher for 3 attempts than for each of the other levels of attempts. We can conclude from this that a value of three does not appear optimal and increasing this number to 5 provides a significantly lower number of accounts locked out but increasing this to a value higher than 5 appears to have no significant affect on the number of accounts locked out. The next step is to examine

H4 i.e. the impact of this increase on the number of bad logins to establish if there is an acceptable level of authentication attempts that do not result in a significant increase in the number of bad logins.

(I) Tries till lockout	(J) Tries till lockout	Mean Difference (I-J)	Std Error	Sig	95%	
					Lower Bound	Upper Bound
3	5	6.124**	1.001	0.000	3.27	8.98
	7	5.814*	1.001	0.000	2.96	8.67
	9	6.532*	1.001	0.000	3.68	9.39
	11	6.566*	1.001	0.000	3.71	9.42
	13	7.074*	1.001	0.000	4.22	9.93

Table 1: H3 Multiple Comparisons

The first step for testing H4 was to establish the normality of the data for the simulations before applying parametric tests. Satisfied with the normality of the data, it was possible to progress to the ANOVA, Levene, and Tukey HD tests. The number of bad logins was the independent variable and the number of authentication attempts before lock out was the dependent variable. The multiple comparisons were completed using Tukey HSD as the Levene statistic was 0.137, larger than the value of 0.005 required for non-violation of the homogeneity of variance assumption. As with H3, the only groups which resulted in a significant difference were at the level of 3 attempts compared to each other level (7, 9, 11, and 13) and in addition 5 attempts compared to 11 is also significantly different. All other group comparisons showed no significant difference. The resulting multiple comparisons for the significant groups are shown in Table 2.

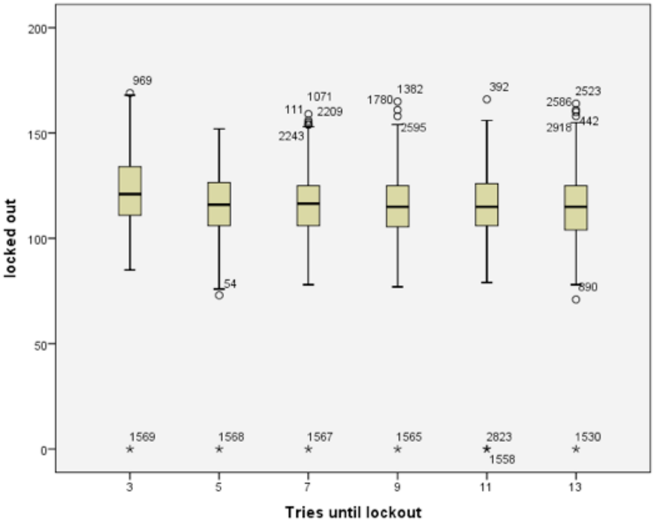

Figure 3: H3 Boxplot

(I) Tries till lockout	(J) Tries till lockout	Mean Difference (I-J)	Std Error	Sig	95%	
					Lower Bound	Upper Bound
3	5	.651*	.094	.000	.38	.92
	7	.699*	.094	.000	.43	.97
	9	.816*	.094	.000	.55	1.08
	11	.953*	.094	.000	.69	1.22
	13	.906*	.094	.000	.64	1.17
5	11	.302*	.094	.017	.03	.57

Table 2: H4 Multiple Comparisons

The corresponding boxplot is shown in Figure 4 where it can be seen that a level of three attempts before lock out results in a higher median percentage of bad logins for group 7, 9, 11, and 13 whilst 5 provides a comparable result. The conclusion we can draw from this is that a higher value than 3 (5, 7, 9, 11, and 13) for the number of attempts before being locked out provides a similar or lower median number of bad logins. Thus 3 attempts appear potentially less optimal in the lockout values which were examined here. A higher number of attempts before lockout could potentially reduce lockouts as well as reducing the percentage of bad logins. We will suggest an explanation for this non-intuitive result in the next section.

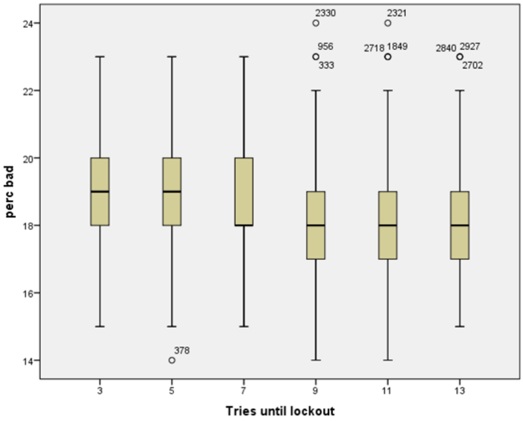

Figure 4: H4 Boxplot

4. Discussion

SimPass produced results that supported hypotheses H1 and H2. This was as expected since it logically follows that allowing sharing means less forgetting and the increase in the percentage of bad logins is a consequence of sharing. For H3 we provided evidence that increasing the number of permitted authentication attempts before lock out reduced the number of lock out events, as expected. Specifically, we found significant evidence that the mean number of locked accounts decreased significantly for attempt values of 5, 7, 9, 11, and 13 when compared to a value of 3 attempts before lock out. This contributes evidence to the hypothesis that whilst increasing the number of attempts helps to reduce the number of locked accounts, there may be a limit to the impact. Thus, selecting a value such as 5 may provide a sufficient significant reduction in the number of accounts locked out.

For H4 we provided evidence that there were a higher percentage of bad logins with 3 permitted attempts before lock out than with a greater number. This was a somewhat unexpected finding since the motivation behind locking out is that it deters potential intruders from carrying out attacks and continuing to try passwords until the correct password is provided. The flip side of this coin, however, is that legitimate users may borrow and steal passwords if they are locked out. This, too, increases the number of bad logins.

The data shows that a larger number of attempts did not necessarily have a significant impact on the percentage of bad logins. A higher value than 3 (5, 7, 9, 11, and 13) for the number of attempts before being locked out actually provides a lower mean percentage of bad logins, because bad logins include use of other employees' credentials both with and without their knowledge. Hence it seems that the number of attempts before lockout could be increased to five without compromising the security of the system, this making it easier for end users without increasing risk significantly.

The simulation suggests that system security, *per se*, would not be compromised, which would be a positive outcome for legitimate users. The reality is that most computer users have 5-6 distinct passwords, and allowing them a few more attempts might help them to fix on the one they used for the system in question. If the number of allowed tries was increased to 5, this would allow hackers two more attempts. When a lockout policy is implemented hackers will often start off with the most commonly used passwords, so if the relaxation in number of tries were accompanied by a strength requirement it might make the effects of the extra two tries negligible.

5. Conclusion

In this paper we use a simulation engine, SimPass, to test potential information security control mechanisms, specifically with respect to passwords. We tested the impact of password sharing and locking users out of their accounts after too many wrong password attempts. The engine produced predictable results in the first instance, demonstrating the negative effects of sharing on system security. In the second case, the simulation showed that the best number of tries to allow before lockout is five, not the de facto three so commonly used in industry.

It is undeniably challenging to carry out this kind of study in industry. There is a level of risk involved in increasing the number of tries before lockout which organisations are understandably reluctant to embrace. However, this will have to be done in order to validate these findings. What SimPass does do is to suggest potentially viable changes and give some indication as to the impact thereof.

6. Acknowledgements

The authors thank Theresa Omoronyia for help with the statistical analysis.

7. References

Adams, A. and Sasse, M.A. (1999) 'Users are not the enemy', Communications of the ACM, vol. 42, no. 12, pp. 40–46.

Brostoff, S. and Sasse, M. (2003) 'Ten strikes and you're out: Increasing the number of login attempts can improve password usability', in CHI 2003 Workshop on Human-Computer Interaction and Security Systems, Ft. Lauderdale, Florida.

Da Veiga, A. and Eloff, J. (2007) 'An information security governance frame- work', Information Systems Management, vol. 24, no. 4, pp. 361–372.

Ebbinghaus, H. (1964) 'Memory: A contribution to experimental psychology', H. A. Ruger & C. E. Bussenius, Trans. (Ed.). New York: Dover. (Original work published 1885)

Gehringer, E. F. (2002) 'Choosing passwords: security and human factors', In (ISTAS'02). International Symposium on Technology and Society. IEEE, pp. 369–373.

Herley, C. (2009) 'So long, and no thanks for the externalities: the rational rejection of security advice by users', in Proceedings of the 2009 workshop on New security paradigms workshop. ACM, pp. 133–144.

Inglesant, P. and Sasse, M. (2010) 'The true cost of unusable password policies: password use in the wild', in Proceedings of the 28th International conference on Human factors in computing systems. ACM, pp. 383–392.

Lee, J.-S. Kim, D. S. Park, J. S. & Chi, S.-D. (2005) 'Design of intelligent security management system using simulation-based analysis', in Proceedings of the 18th Australian Joint conference on Advances in Artificial Intelligence, ser. AI'05. Berlin, Heidelberg: Springer-Verlag, pp. 766–775.

Lubbe, S. & Klopper, R. (2005) 'The problem with passwords', Alternation. Themes in Management and Informatics, vol. 12, no. 2, pp. 53–78.

Mandujano, S. & Soto, R. (2004) 'Deterring password sharing: User authentication via fuzzy c-means clustering applied to keystroke biometric data', in ENC 2004. Proceedings of the Fifth Mexican International Conference. IEEE, pp. 181–187.

Murdoch, S., Drimer, S., Anderson, R. & Bond, M. (2010) 'Chip and PIN is broken', in Security and Privacy (SP), 2010 IEEE Symposium on. IEEE, pp. 433–446.

Posthumus, S and Von Solms, R. (2004) 'A framework for the governance of information security', Computers & Security, vol. 23, no. 8, pp. 638–646.

Reinhart, C. & Fitz, A. (2006) 'Findings from a survey on the current use of daylight simulations in building design', Energy and Buildings, vol. 38, no. 7, pp. 824–835.

Renaud, K. & Mackenzie, L. (2013) 'SimPass: Quantifying the impact of password behaviours and policy directives on an organisation's systems', Journal of Artificial Societies and Simulation, vol. 16, no 3.

Renaud, K. (2012) 'Blaming Noncompliance Is Too Convenient. What Really Causes Information Breaches?', IEEE Security & Privacy, May, vol. 10, no 3, pp. 57–63.

Scalese, R. J. & Issenberg, S. B. (2005) 'Effective use of simulations for the teaching and acquisition of veterinary professional and clinical skills', Journal of Veterinary Medical Education, vol. 32, no. 4, pp. 461.

Simon, H. A. (1996) The sciences of the artificial. MIT press, Cambridge, Massachussets.

Singh, S., Cabraal, A., Demosthenous, C., Astbrink, G. & Furlong, M. (2007) 'Password sharing: implications for security design based on social practice', in Proceedings of the SIGCHI conference on Human factors in computing systems. ACM, pp. 895–904.

Vaughn, R. G. (1995) 'Use of simulations in a first-year civil procedure class', J. Legal Educ., vol. 45, pp. 480.

Vickery, P., Skerlj, P., Steckley, A. & Twisdale, L. (2000) 'Hurricane wind field model for use in hurricane simulations', Journal of Structural Engineering, vol. 126, no. 10, pp. 1203–1221.

Von Solms, R. & Von Solms, S. (2006) 'Information security governance: A model based on the direct–control cycle', Computers & Security, vol. 25, no. 6, pp. 408–412.

Williams, P. (2001) 'Information security governance', Information security technical report, vol. 6, no. 3, pp. 60–70.