Abstract

Users are often considered the weakest link in the security chain because of their natural propensity for choosing convenience over safe practice. One area with a vast amount of evidence related to poor user behaviour is that of password management. For example, when hackers gain unauthorised access to public websites, subsequent analysis generally confirms that compromised passwords are to blame. We have a pretty good idea of the extent to which careless behaviour impacts on the individual user's personal security. However, we don't fully understand the impact on the organisation as a whole when such laxity is aggregated across a large number of employees, nor do we know how best to intervene so as to improve the level of protection of critical systems. Current wisdom mandates the use of increasingly draconian policies to curb insecure behaviours but it is clear that this approach has limited effectiveness. Unfortunately, no one really understands how the individual directives contained in these policies impact on the security of the systems in an organisation. Sometimes a mandated tightening of policy can have unexpected side-effects which are not easily anticipated and may indeed prove entirely counterproductive. It would be very difficult to investigate these issues in a real-life environment so here we describe a simulation model, which seeks to replicate a typical organisation, with employee agents using a number of systems over an extended period. The model is configurable, allowing adjustment of particular input parameters in order to reflect different policy dictats so as to determine their impact on the security of the simulated organisation's IT infrastructure. This tool will support security specialists developing policies within their organisations by quantifying the longitudinal impacts of particular rules.

Keywords:
Passwords, Policies, Organisation, Security, Authentication

Introduction

1.1 “Good Practice” in information security states that at least the following password rules must be included in information security policies, and enforced within organisations FIPS (1985).

- Do not recycle passwords, use a different password on each system.
- Use strong passwords.
- Do not write passwords down.
- Do not share passwords: never tell anyone your password.

1.2 The thinking behind these rules is depicted in Figure 1. The term "weak password" in the figure refers to so-called common passwords (See Section 4.4): e.g. a password which is a variant of a user's own name (Brown et al. 2004) or the system name (Bishop & Klein 1995) or a variation of a previous password (reuse) (Riley 2006) etc. A very common practice is recycling (Inglesant & Sasse, 2010), the use of the same password on more than one system.

1.3 The arrows in the diagram indicate a conventionally assumed causal relationship. So, for example, as more passwords are written down, so more will leak (become known to others), and this will reduce overall system security. The potential for increased security incidents increases, thus escalating the vulnerability of the systems. The dashed line indicates a more tenuous link: this causative will only occur if passwords leak outside the organisation. The certain danger is internal since most recorded passwords will be easily accessed by people within the trusted interior. The four causatives are on the left of the diagram and one
can see the attraction of attempting to tighten security by forbidding them. The assumption is clearly that removal of these triggers will eliminate those factors that lead to poorer system security.

Figure 1. Commonly Deployed Coping Tactics to cope with Multiple Passwords

1.4 Unfortunately, the prevalence of these triggers shows that the practices in question will not be easy to eradicate. There is evidence that users engage in these potentially damaging user behaviours as *coping tactics* (Adams & Sasse, 1999), an inevitable result of people being unrealistically overloaded with passwords while having only limited human capability to cope with the memorial load.

1.5 When a prominent system such as Sony, is breached, researchers invariably report on the large percentage of "common" passwords chosen and the limited number of strong ones in use (See Section 4.4). This confirms that a large percentage of users choose weak passwords, even though they are probably aware that their accounts could easily be compromised as a consequence.

1.6 Some user behaviours have a wider impact than others. For example, a user may use the same password on all their systems at work and, if this password becomes known to another employee, there is an opportunity for much more damage than if no such recycling has been practised. Of course, if a user shares his or her credentials and another user logs in using these, there is no evidence that this masquerading has occurred. An audit might well conclude that a particular organisation's security is satisfactory without realising that people are using each other's credentials. Since non-repudiation is such a core concept of information security this user behaviour must be seen as constituting a threat of a fundamental nature.

1.7 It is inherently difficult to measure the true impact of password coping tactics on the security of the systems of any particular organisation. Password policies must be evaluated within the context of the entire organisational structure and not based on the behaviour of a few individuals. Moreover, it is clear that the view depicted in Figure 1 is rather naïve because it does not consider why people are tempted to use these coping tactics. There is an implication that undesirable password practices can be attributed to innate human weakness whereas, in fact, the system within which the human functions clearly plays a key role in inducing these user behaviours; in consequence, merely focusing on the symptoms and trying to eliminate them without identifying underlying causes is bound to fail. Figure 2 expands one part of systems diagram slightly, showing the interaction of a few more factors that common sense suggests could play a role in leading to leaked passwords.

Figure 2. System Causatives and their Role in Leading to Undesirable Password Practices

1.8 For example, the response to a security breach is often a strengthening of password requirements (link from "security incident" to "required password strength" in the diagram). Tari et al. (2006) have shown that the more complex a password, the more easily it can be observed by another user (link from "required password strength" to "observability"). Complex passwords are also more easily forgotten, so users are probably more likely to write them down. Hence this effort to strengthen passwords may actually lead to more passwords being leaked.
1.9 Users do not operate in isolation: a number of factors may encourage them to behave in a particular way, and they often feel they have no option but to break rules which are very difficult or too constraining to obey. Hence great care needs to be taken to ensure that policy rules are indeed possible to follow, do not present users with an ethical dilemma and do not require too much effort (Inglesant & Sasse 2010).

1.10 For example, we know that when people work together in groups they will start to trust one another and share information. Sharing passwords is a natural extension in many scenarios and insisting that colleagues do not do so may present them with an ethical dilemma where the group sees itself as working together towards a common purpose. Denying a colleague the use of your password when he or she needs to make a contribution to this common purpose may be seen as unhelpful and so incur significant peer pressure (Renaud 2012).

1.11 Given the complex nature of the psychological and social factors at play, the typical degree of heterogeneity of the systems involved and the infeasibility of conducting live field studies in a real organisational environment, it is hardly ever feasible to gauge the impact of user password behaviours on the security of an organisation’s systems by experimental means alone. A viable alternative is the use of computer, a well-established approach (Simon 1969) 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. It must always be borne in mind that observation of a real system is the only sure way to establish the level of accuracy of simulation predictions and thus validate the model in question. However, as we cannot, in general, test predictions across the whole parameter space, we extrapolate that a validated model will be able to make correct predictions across a generalised subset of the parameter space where conditions are similar to the validation points. With such reservations in mind, 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, predicting user behaviour in previously unexplored regions of the parameter space.

1.12 It must be acknowledged, at the outset, that a simulation is only as good as the assumptions that are built into it and in what follows an effort has been made to ensure that the model rules and default input parameters are grounded in the literature wherever suitable studies exist. It is reasonable to assume that such a simulation can provide understanding and insight which is simply impossible to gain in a real-life setting. It can also help us to understand the wider effects of variations in individual human behaviours integrated across organisations, something almost impossible to gauge experimentally in the real world. As a result, it can assist us in understanding the often unexpected effects of particular policy directives. So, for example, if the organisation’s auditors require password changes more frequently than previously, this requirement can be “plugged into” the simulation, and the net effect can be charted.

1.13 This paper reports on an actual implementation of just such a simulator, SimPass, an engine which models user password usage within an organisation. In a SimPass organisation, users authenticate using usernames and passwords, on a variety of different systems. Over time they manifest the kinds of behaviours their real life counterparts engage in. The simulation also includes hacker agents and malicious agents, both attempting to breach user accounts. At the end of the simulation the engine provides summary data related to the security of the system after a period of time, reflecting the impact of particular agent behaviours.

1.14 The rest of the paper is structured as follows. Section 2 briefly describes the simulation model: the main entities abstracted from real-world scenarios for the purposes of the simulation model and the behaviour of the agents within the simulation. Section 3 explains how the engine was implemented. Section 4 explains how the system is configured in order to test the effects of different policy and overall system settings. Section 5 gives an example of how a particular policy change was tested using SimPass and Section 6 concludes.

Simulation Model

2.1 Employees usually have an assigned position within a hierarchical organisational structure, as shown in Figure 3. Employees often work closely with other people in their “branch” of the structure, with infrequent interactions with other branches. They build relationships with their branch colleagues, and work towards a common purpose. In carrying out their duties they make use of one or more computer systems which can be internal or external, with respect to visibility to the outside world. Systems can be attacked by outside hackers and by internal malicious employees but the security of the organisation is also at risk from the ill- advised actions of well-intentioned employees.
2.2 The key entities which SimPass abstracts are employees and systems. We begin by briefly examining the nature of each in turn.

**Employees**

2.3 Employees work within a particular environment, and are subject to the pressures of that environment: their position within the hierarchy, the tasks they are paid to undertake, their workload, the culture of the organisation and the quality of the relationship with their colleagues (Figure 4). They are also constrained by various information security policies.
2.5 An organisation will typically have a number of software systems, executing on different hardware. Such systems are either visible to the outside world, or hidden behind a firewall. Some systems issue passwords while others allow employees to choose their own. Furthermore, systems can be configured to implement particular organisational rules, such as, for example password length, lockouts etc. SimPass’s model of a system housed within the organisational context is shown in Figure 6.

![SimPass model](http://jasss.soc.surrey.ac.uk/16/3/3.html)

### Model Agent Behaviour

3.1 SimPass is a multi-agent system which simulates the organisational Ecosystem as described in the previous section. A SimPass agent is described as a 4-tuple: 

\[
\text{Agent} = (\text{Sit}, \text{Act}, \text{Dat}, f_{\text{Agent}})
\]

where

- \(\text{Sit}\) is the set of situations the agent can be in,
- \(\text{Act}\) is the set of actions that the agent can perform,
- \(\text{Dat}\) is the set of possible value combinations of the agent's individual settings, and
- \(f_{\text{Agent}}\) is the agent's decision function, and can be expressed as follows:

\[
f_{\text{Agent}} : \text{Dat} \times \text{Sit} \rightarrow \text{Act}.
\]

There are four kinds of agents representing regular employees, system administrators, malicious employees and hackers (outsiders). These agents do not operate in isolation but interact with other agents regularly and in a variety of ways.

3.2 When the engine starts, systems and agents are created and configured with characteristics tailored as discussed in Section 4. The configuration settings are enumerated in the appendix. Regular agents, \(\text{Agent}_{\text{Regular}} = (\text{Sit}, \text{Act}, \text{Dat}, f_{\text{Regular}})\) are simply trying to do their jobs by using the various systems they have credentials to access.

- \(\text{Sit}\) constraints are contextual, depicting the situation an agent is in. For example, an agent wants to log in, or needs to enrol for a new system.
- \(\text{Act}\) is the set of actions that the agent can perform in a situation. For example, if an agent forgets a password, it can be locked out, try to get a shared password, or try to steal a password.
- \(\text{Dat}\) is the set of possible value combinations of the agent's internal settings, as shown in Figure 5.

3.3 The other agents have more specific natures.

- **Malicious agents** \(\text{Agent}_{\text{Malicious}} = (\text{Sit}, \text{Act}, \text{Dat}, f_{\text{Malicious}})\) will use their systems as usual, but sometimes they might try to try to use another agents' credentials perhaps because they have a grudge against their victims or because they wish to carry out fraudulent activity (attacks are randomly generated). If they decide to do this, they will use shared, stolen or guessed passwords to gain access to accounts with other employees' credentials. All such logins are termed "bad".
- **Hacker agents** \(\text{Agent}_{\text{Hacker}} = (\text{Sit}, \text{Act}, \text{Dat}, f_{\text{Hacker}})\) try to attack the system from outside: trying to guess credentials without the ability to find written records, and only being able to access visible systems. Hackers can try to breach systems using well-known default system passwords, which might not have been reset by less conscientious system
administrators. They could also try to gain access to password files, which they can try using brute force. Failing this, they can try a combination of user names and passwords to attempt to access the system. Successful hacker logins are also termed "bad".

- **Sysadmin agents** $\text{Agent}_{\text{Sysadmin}}(\text{Sit}, \text{Act}, \text{Dat}, f_{\text{Sysadmin}})$ administer one or more of the systems in SimPass. These employees are generally considered to be honest: they do not engage in any of the harmful coping behaviours mentioned above. However a sysadmin may occasionally fail to attend to its duties by, for example, failing to patch a system for which it is responsible.

### 3.4 Agents

Agents are randomly assigned to an organisational management hierarchy thus ensuring that they work in groups, with a number of trusted "colleagues" whom they could on occasion ask to share passwords with them. Agents are initially given access to a limited number of systems. Each week thereafter a random number of agents will be required to enrol for additional systems, as is the case when one gains experience in a typical organisation.

**Regular Agents**

#### 3.5

To support the information security principles of non-repudiation and authorisation, agents are given their own credentials, username and password for each of the systems to which they have access, and for a given individual the number of such systems tends to increase the longer he or she is employed. Each agent thus "owns" a set of credentials for each system which they use. Passwords are selected based on the characteristics of both the agent and the system involved, as shown in Figure 7. For example, if an agent sometimes recycles it might well use one of its existing passwords rather than choosing a new one. New passwords are chosen from a representative password repository (as described in Section 4.4).

![Figure 7. Actions For a Regular Agent Enrolling for a New System](image)

#### 3.6

For each of the agent's systems a "next usage" is randomly chosen to be one of the following number of days: 1, 2, 3, 7, 14, 30, 60, or 90, reflecting daily, frequent, weekly, bi-weekly, monthly and three-monthly usage. The choice is weighted to favour frequent accesses more than infrequent ones. This is achieved by selecting one of the appropriate time intervals from a list according to probabilities determined by associated weights as follows: 10, 9, 9, 10, 5, 1, 1. Thus daily usage will appear 10 times more often than 3 monthly usage.

#### 3.7

As shown in Figure 8, agent A, when prompted to use its systems, will attempt to log in. If the password is remembered it will do so without incident. If not, a number of actions can result. It can accept that it is locked out, which means that it cannot complete its tasks for the day and must wait for a replacement password. If the work is urgent it could try to obtain credentials from another agent, B, either because the latter has willingly shared them, or because A has managed to obtain them dishonestly. The systems being logged into will never uncover this kind of activity, since it appears to be legitimate use by agent B and so the principle of non-repudiation is broken.
3.8 In summary, if an agent tries to log in but the password has been "forgotten", it has three options: request a new password and be locked out of the system; ask a fellow agent to share its password; or try to steal one. The tactic is chosen randomly but with probabilities depending on the agent's propensities. For example, only dishonest agents will steal passwords. See Figure 9.

3.9 If an agent shares or steals a password and logs into the system using it, that is termed a "bad" login. An agent who logs into the system using its own credentials executes a "good" login.

Malicious Agents

3.10 A malicious agent is an otherwise regular agent who, for whatever reason (e.g. revenge, fraud), tries to masquerade as another agent credentials to gain access to systems. The attacking agent will use the following tactics to log into the target's account on some system:

1. Check whether the target agent has, some time in the past, shared a password with the attacker. The agent will test to see whether it is still valid.
2. Check whether the target has recorded its password, and not secured it. Try to access the system using this password.
3. Opportunistically try the following strategies:
   a. a randomly chosen common alphanumeric password string from a list maintained by SimPass. Examples are 123456, password1, or qwerty;
   b. a variation of the username, eg John1; or
   c. a variation of the system name eg Amazon1.

3.11 If the malicious agent manages to breach the target's account, it will try to use the same password on other systems, in the hope that the target agent recycles passwords. It might also choose to change the target agent's password, thus locking the victim out, and disrupting its ability to do its job (see Figure 10).
3.12 Hackers often target specific organisations, for different reasons. Newspaper stories of their exploits are easy to find (BBC 2011). Hackers will first try the default password of all visible systems (van Doorn 1992) in the hope that the system administrator will not have reset these (Workman 2008). Hackers will also try to get hold of the password file if it this is not secured properly. With access to this file, a brute force attack will be carried out to try to determine the passwords for the listed agent names. Correctly guessed combinations will be used to break into systems. Their tactics are depicted in Figure 11.

3.13 The next step will be to try username-password combinations. Since these attacks are usually conducted without personal knowledge of system users, a targeted approach, where the hacker guesses a password for a specific user based on personal knowledge, is not usually feasible. SimPass hacker agents use a generic approach, simply trying various username and password combinations to see whether they can gain access:

- a randomly chosen common password from a list maintained by SimPass. Examples are 123456, password1, or qwerty;
- a variation of the username, eg John1;
- a variation of the system name, eg Amazon1

3.14 If the hacker agent breaches an account it might also change the victim’s password, thus locking it out, and disrupting its ability to carry out its tasks.

3.15 Systems admin agents are responsible for one or more systems into which they will randomly log in. If they are trained, they will change their systems’ admin passwords, and keep the systems patched and the password files secured. Untrained systems administrators might well neglect these responsibilities and make it more likely that a hacker can breach the systems.

Summary

3.16 Table 1 summarises the internal settings of the different agents (Dat)
Simulation Engine

4.1 SimPass is implemented in Java, as a multi-threaded application which generates discrete events at regular intervals, as shown in Figure 12. On startup it will read a configuration file, and initialise the simulation as follows:

1. An object is created to represent each agent and a thread launched for each (regular, malicious, sysadmin and hacker).
2. A system object is created for each "system" in the simulation.
3. A time manager object ensures that the simulation runs for as many days as specified. As each "day" starts, the time manager prompts each agent to log into those systems scheduled for use on that day. The time manager advances the day counter when all agents have concluded their day's tasks.
   a. Every time the agent logs in, the system randomly generates a "next use" until the simulation ends. When the "next use" day is chosen the system will decide whether the agent will forget the password or not, based on the literature on memorability (Section 4.1.1).
   b. The same mechanism applies to malicious and hacker agents who will, at random intervals, carry out attacks.
4. Agents interact with one another as they carry out their daily tasks.
   a. Sharing passwords with colleagues.
   b. Trying to find other agents' recorded passwords ie. stealing them.
   c. Observing each other typing in passwords. Whereas theft is goal-directed and will happen as a result of an agent's either having lost their own password or maliciously wanting to breach someone's account, observation can happen casually. Agents will not always "remember" an observed password: this, too, is randomised.
5. Agents and systems log all their activities to individual log files. At the end of the simulation a summary of all activity is printed to a summary log file to support further analysis.
6. SimPass keeps a tally of particular events in the system to support quantification of overall system security, as shown in Table 2.

---

Table 1: Agent Internal Configurations

<table>
<thead>
<tr>
<th></th>
<th>Dishonest</th>
<th>Malicious</th>
<th>Forgetful</th>
<th>Sharing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regular</td>
<td>Yes/No</td>
<td>No</td>
<td>Yes</td>
<td>Yes/No</td>
</tr>
<tr>
<td>Malicious</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes/No</td>
</tr>
<tr>
<td>System Administrators</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Hackers</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

The following section explains how the engine has been implemented.
Table 2: Events of Interest in the Simulation

<table>
<thead>
<tr>
<th>Event</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good login</td>
<td>A login where the agent uses his own credentials</td>
</tr>
<tr>
<td>Bad login</td>
<td>A login where the agent uses someone else's credentials (shared, observed, guessed or stolen)</td>
</tr>
<tr>
<td>Lockout</td>
<td>When a user has had to request a password reset</td>
</tr>
<tr>
<td>Stolen Password</td>
<td>A password which has been observed, and recorded, by another agent or where a written record of a password is discovered by another agent.</td>
</tr>
<tr>
<td>Shared Password</td>
<td>A password which the agent has willingly allowed someone else to use</td>
</tr>
</tbody>
</table>

Simulation Model Settings

Agent Characteristics

5.1 Various coping tactics are commonly used by people working with IT systems (Figure 13) and agents are designed to reflect these user behaviours. Each agent will deploy some, all or none of these mechanisms and will be subject to limitations reflecting those of average humans. Agent characteristics and how they help select default values for the input parameters to SimPass are discussed in the following subsections. Needless to say, each parameter can be varied from the selected default if desired, in order to explore the corresponding dimension of the input space.

Forgetting

5.2 A number of researchers have investigated forgetting rates. Florencio & Herley (2007) reported that 4.28% of regularly used passwords are forgotten. Bunnell (1997) reports on 27% forgetting rates after 2 weeks. Zviran and Haga (1993) reported on a 75% forgetting rate after 3 months. This was confirmed by Beedenbender (1990) who reported 72.8% forgetting after 3 months.

5.3 Some surveys have asked users to report on how many passwords forgotten after a month (Brown 2004), (Tamil et al. 2007; Elcomsoft Proactive Software 2009; Campbell & Bryant 2004). If the numbers of respondents are tallied, it becomes clear that 30% of passwords are forgotten after a month of non use. This is confirmed by the study carried out by Theusinger and Huber (2000) and by Brown et al. (2004), who found that 32% and 31% of passwords were forgotten by system users within a month.

5.4 As discussed in the psychological literature Ebbinghaus (1885), these figures are a good fit to a parabola with the formula: 
\[ y = (-0.002)x^2 + 0.96x + 3.04 \]
The forgetting rates used in SimPass, shown in Table 3, reflect this relationship.

Table 3: Forgetting Rates

<table>
<thead>
<tr>
<th>Intervals</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>7</th>
<th>14</th>
<th>30</th>
<th>60</th>
<th>90</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forgetting %</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>10</td>
<td>16</td>
<td>30</td>
<td>53</td>
<td>73</td>
</tr>
</tbody>
</table>

5.5 In SimPass these forgetting rates will be tailored depending on how many times a specific password has been used in the past. A
frequently used password is less likely to be forgotten than an infrequently used password so the system will factor in previous use when deciding whether or not a password will be forgotten. Figure 14 shows the forgetting rates in a typical simulation run. The upper line depicts the values given in the above table, and the line below those generated by SimPass itself.

![Figure 14. Forgetting Rates in a Typical Simulation (lower line=Simulation, upper-line based on forgetting rates)](http://jasss.soc.surrey.ac.uk/16/3/3.html)

**Sharing of Passwords**

5.6 Password sharing is strictly forbidden by most organisations yet the reality is that it is widely practiced. A number of password use surveys report on the prevalence of sharing (or at least reports on those who will admit to sharing). Martinson (2005) reports 38.1%, Bryant & Campbell (2006) reports 42%, Stanton et al. (2005) reports 34%, Tamil et al. (2007) reports 33.9% and Campbell and Bryant (2004) report 40%. However, Hoonakker, Borneo and Carayon (2009) report that only 5% of respondents admitted to sharing passwords. It is possible that sharing is organisation-specific, but there may be other unreported factors at play here.

5.7 Since the majority of surveys report that close to a third of respondents share, and since most organisations find it difficult to accept any sharing at all, the default sharing percentage will be set to 33% but, as with all these values, this can be overridden in any SimPass simulation to explore the effect of different scenarios.

**Stealing of Passwords**

5.8 Using a stolen password is undeniably dishonest. How likely is it that an employee will do this? A quick look at dishonest behaviour in other settings is enlightening. Karstedt and Farrall (2006) found that 65% of people, given sufficient motivation, would behave dishonestly. Von Lohman (2004) reports on studies of P2P music sharing. Whereas 88% of the respondents in the study believed that this sharing was wrong, 56% still admitted downloading music illegally. Wilkes (1978) carried out a study into dishonest customer behaviour and found that for some offences between 70 and 80 percent of customers would offend. This study admittedly reports on customer behaviour, whereas SimPass is modelling employee behaviour. How honest can we expect employees to be? Wilson (2009) reports on a study by CyberArk, who surveyed 600 workers in New York and London. A surprising 48% said that they would steal their company's data if they were fired. Perhaps being fired constitutes sufficient motivation, but what about everyday behaviour? Wilkes (1978) cites a study by Tatham (1974) which reports that 50% of employees had admitted to stealing from their employers. Boye and Jones (1997) presented details of a study of restaurant employees which showed that 60% of respondents had stolen from their employers. There is some agreement that some organisations have more of a culture of dishonesty than others (Kidwell & Kochanowski 2005; Johnson & Philips 2003) and hence the dishonesty prevalence can be configured. The default prevalence of dishonesty in SimPass is 65%, taken from Karstedt and Farrall (2006).

5.9 Given the fact that someone is dishonest, does that necessarily mean that he/she will engage in stealing passwords? According to Cressey (1973), three elements must be present for a person to engage in dishonest behaviour: motivation, rationalisation and
5.10 Certainly there is evidence that people do indeed steal passwords (Kidwell & Kochanowski 2005; Forbath 2005), but there is no hard evidence in the literature which quantifies the extent of the problem. In the absence of evidence we argue that, human nature being what it is, the literature mentioned in the previous paragraphs is a reasonable predictor of whether people will rationalise dishonest behaviour or not, in this case stealing and using someone else's password.

5.11 Reflecting their human counterparts, SimPass agents are also categorised as "dishonest" or "honest"; the latter will never steal a password, the former, given sufficient motivation, will rationalise the use of another's password. Agents can steal passwords that other agents have written down or recorded in some other unsecured way (the prevalence of this is discussed in Section 4.1.8). The former may, in addition to deliberately stealing a password, observe another agent entering a password and record it for later use. The observation rate for a simple password is set at 1% and for a complex password it is set to 2%. Moreover, an agent will only observe and record a password if it has sufficient motivation. The motivation in this case is that an agent has previously forgotten a password or passwords and therefore has a reason to want to guard against this eventuality in the future.

5.12 Brown et al. (2004) reported that 45% of users used a variant of their own name as their password. This was confirmed by Harada and Kuroki (1996) who found a prevalence of 42%. SimPass uses a default of 45%.

5.13 Some people try to link their password to the system it is being used on, so as to increase their chances of remembering it. So, for example, they could use Amazon1 as their password for the Amazon website. Bishop and Klein (1995) reported that 11% of users employed of this tactic. Interestingly, however, Schneier[1] carried out an analysis of MySpace passwords and found the corresponding figure to be only 0.11%, which suggests that this prevalence varies across user populations. SimPass uses a default of 11% for this setting.

5.14 Many users remember a few passwords and then use them across a number of systems. This coping tactic is probably the most common. Being able to predict the true prevalence of this coping technique is difficult, due to the different percentages reported by different studies. In order to use a realistic percentage a tally was made of all respondents who admitted to this practice from the studies reported by: (Hoonakker et al. 2009; Campbell & Bruyant 2004; Riley 2006; Zviran & Haga 1993; Tamil et al. 2007; Martinson 2005; Brown et al. 2004). 1592 of the total of 2966 respondents admitted to recycling passwords comprising 54% of those surveyed. Adams and Sasse (1999) reported a 50% prevalence and Summers and Bosworth (2004) report 55%. An analysis of actual leaked passwords from multiple systems show a recycling prevalence of 92% (Hunt 2011), and a survey reported by SecurityWeek (2010) reports that 75% of people recycled passwords. This suggests that many fewer people are prepared to admit to this practice than actually engage in it. SimPass uses a default of 54% based on the above composite tally.

5.15 Florencio and Herley (2007) found that users tended to maintain an average of 6.5 passwords, so SimPass agents do the same, the system ensuring that they have a maximum of 6 distinct passwords if they do indeed recycle their passwords.

5.16 Some users will, when required to provide a new password for a system, simply vary the previous one: here this is referred to as reuse. Two studies have reported on the prevalence of this practice (Riley 2006; Hoonakker et al. 2009). 672 respondents out of a total or 1164 admitted to this practice (58%), which is used as the SimPass default.

5.17 Users often resort to writing down their passwords, or recording them in some other fashion. The following studies were consulted: (Zviran & Haga 1993; Brown et al. 2004; Martinson 2005; Hoonakker et al. 2009; Bryant & Campbell 2006; Stanton et al. 2005; Tamil et al. 2007; Riley 2006). Out of a total of 3386 respondents, 1309 admitted to writing their passwords down (39%). Only Hoonakker asked whether they also secured this password record and 18% said they did this. These values are used as the SimPass default settings.
Password Strength

5.18 There is some evidence that users, when forced to change their passwords, will choose a weaker password (Martinson 2005). A default of 68% was chosen based on this study.

System Admin Conscienciousness

5.19 We will assume that 77% of system administrators will patch systems and change passwords, with 23% leaving their systems unprotected. This is based on a study published by Microsoft (Forbath et al. 2005) which stated that only 77% of systems were patched, on average.

Threats

5.20 Threats are classified as internal or external depending on whether they are initiated by agents of, respectively, the malicious or hacker types. In SimPass, if a malicious (insider) agent or a hacker gains access to an agent's account it can decide to leave things as they are, or to change the agent's password. In the former case, details are always retained for later use so that if the hacked agent does not discover the hacker’s activity the door is left open for later access. If the hacker decides to change the victim’s password then the latter is prevented from accessing its account and will have the same choices as it has when it forgets its password.

5.21 If an attacker of either of the above types succeeds in breaching one account, it will try to get into the owner's other accounts using the same password, in the knowledge that many users recycle.

Insider (Malicious) Threats

5.22 Malicious insiders can cause a great deal of damage (Probst et al. 2010). Predd, Hunker and Buford (2008) cite a survey by the Computer Security Institute which reported that the organisations that responded to their survey had attributed 40% of their losses to insider activities. It is much harder to predict the incidence of malicious insiders since many are undetected. Price Waterhouse Cooper's 2010 Information Security Survey found that 19% of large organisations and 5% of small organisations had reported staff using their systems for theft or fraud.

5.23 A 2008 Forrester Research report[2] proposed that 30% of security breaches were caused by malicious insider activity. PriceWaterhouse (2010) reports that organisations experienced an average of 45 incidents last year, suggesting that an average of 13 per organisation were caused by insiders. This argument cannot, however, be used to arrive at an estimation of the number of malicious insiders since multiple incidents could be caused by the same person. As Calder argues, it is difficult to arrive at a reliable estimate of the average incidence of malicious employees in organisations (Calder 1987).

5.24 What is interesting is evidence that, for whatever reason, incidents of this type are increasing year on year (PriceWaterhouse 2010). In 1969 Robin (1969) reported on malicious user behaviour in three companies. The number of employees apprehended was 0.48% per annum. However, Choo and Tan (2007) refer to research at the University of California at Berkeley which reported a 115% increase student dishonesty cases between 1995 and 2000. The former increase could well be attributed to increasing use of computer systems but the latter is less easy to explain.

5.25 Here, it was decided that the incidence of malicious employees would be set at 1%, in order to depict a relatively optimistic scenario but, of course, the figure can be set by the simulation user. Malicious agents may decide target specific individuals at random intervals, attempting to breach their accounts in order to do damage to them, or to use the account to carry out nefarious activities.

Outsider (Hacker) Threats

5.26 The number of hackers that will target a particular company’s systems over a period of interest may vary widely. Here, a default of 3 hackers has been chosen, but this number is configurable. Hackers will attack at randomly assigned intervals, and will target a maximum of 10 agent accounts on any system before retreating to try again another day. This technique is deployed by many hackers, who do not wish their activity to be too easily spotted by vigilant systems administrators.

System Characteristics

5.27 Many aspects of systems can be configured (see Figure 6). The default values are grounded in the research literature, as with agent characteristics. When such research is unavailable, the simulation owner may provide settings to explore different scenarios. The following configuration aspects have default values but can be varied to explore the parameter space.

- **System Visibility:** Some systems are visible from outside the organisation, others reside behind the firewall and cannot be accessed. Visible systems are susceptible to outsider attacks. In SimPass, 50% of the systems are visible by default.
- **System-Issued vs Self-Chosen Passwords:** Some systems issue passwords and others allow agents to choose their own. Since there is no published evidence of the distribution, an arbitrary proportion of 10% of systems will do the former
and the rest will allow agents to choose their own passwords.

Some system characteristics are specified by the organisation, to align with their policy directives. Examples of these follow.

- **Password Changes**: This determines whether agents are required to change their passwords and, if so, the number of days between mandatory changes. The default setting is that changes are required and that 30 days will lapse between password changes. It is assumed that agents cannot re-use a previously used password but they can add a numeral to the end, as is commonly the case.
- **Lockout**: This value measures whether authentication attempts are limited. Since many systems apply the practice of three times lockout, this will be applied in SimPass.
- **Password Requirements**: Organisations often impose a minimum password length and complexity on their employees, regardless of what individual systems require. The most common of these is that a password must have a minimum number of characters, that it should contain a numeral, upper case characters and/or a special character.

### Password Corpus

5.28 SimPass agents choose passwords for their systems, and a realistic password corpus is required. Various sites have had their passwords leaked and "post mortem" analyses have subsequently been carried out. For example: Schneier (2006) analysed 34 000 passwords; Calin (2009) reported on an analysis of 9843 Hotmail passwords; Hunt (2011) analysed 77 million Sony passwords; and phished phpBB passwords were analysed by Graham (2009). A summary is given in Table 4.

<table>
<thead>
<tr>
<th></th>
<th>MySpace</th>
<th>Hotmail</th>
<th>Sony</th>
<th>phpBB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Common Passwords</td>
<td>2.7%</td>
<td>2.5%</td>
<td>&gt; 10%</td>
<td></td>
</tr>
<tr>
<td>Numbers Only</td>
<td>1.3%</td>
<td>19%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lowercase Only</td>
<td>9.6%</td>
<td>42%</td>
<td>45%</td>
<td></td>
</tr>
<tr>
<td>Dictionary Word</td>
<td></td>
<td>64%</td>
<td>65%</td>
<td></td>
</tr>
<tr>
<td>Alphanumeric</td>
<td>81%</td>
<td>30%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alphanumeric &amp; Special Char</td>
<td>8.3%</td>
<td>6%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Average Length</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>6</td>
</tr>
</tbody>
</table>

5.29 Disturbingly, 4% of phpBB passwords were variations of the word "password". When a number is used, in 45% of cases it is the number 1, and when a special character is used, it is most often "! ", followed by the ".". What these surveys show is that while the passwords used by a particular population vary significantly, users will often choose simpler and weaker options when given the option to do so. In SimPass the password corpus will include realistic percentages of representative categories of most major identified password types.

5.30 The password collection used by SimPass uses the percentages of passwords in each of the following categories as shown in Table 5. Note that a "people name" category has been included, despite having no evidence from the above analyses to support it. However, Medlin et al. (2005) found that 19.3% of the people in their study chose passwords which reflected family names. Such a significant number was thus worth including.

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Common Passwords</td>
<td>10%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Numbers only</td>
<td>5%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One Word</td>
<td>30%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Movie Names</td>
<td>10%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>People names</td>
<td>10%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A word followed by a number</td>
<td>20%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two words</td>
<td>5%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two words separated by a number</td>
<td>5%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alphanumeric and a special char</td>
<td>5%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5.31 The list of common passwords was obtained from Whats My Pass (2008) and the words were mined from an online dictionary. The password corpus holds 100 000 passwords altogether. The minimum length of such passwords is 6 characters and the length extends to as many characters as specified by the simulation user.
Simulation Experiment

6.1 To demonstrate the potential of SimPass here the issue of writing down passwords will be addressed to assess its impact in a scenario using mostly the default settings established above. Most organisations officially forbid the recording of passwords in this way but, despite this, many users do so anyway so that they will not forget their passwords. The perceived risk is that other people will find such records and exploit that knowledge, as depicted in Figure 15.

Figure 15. Writing Passwords Down

6.2 Whereas most current information security policies forbid the recording of passwords, an alternative approach might be to focus on reducing the incidence of forgotten passwords thus removing much of the motivation for passwords to be stolen. The expected downside would be an increase in vulnerability, if insecure recording methods were to make it easier for malicious employees to steal passwords should they still wish to do so.

6.3 To test the effect of writing down passwords on the security of the system, two simulations were executed, spanning 100 days with 100 agents using up to 27 systems simultaneously. One malicious and three hacker agents were introduced to attempt to breach visible systems at random intervals. The two simulations varied as follows.

- 39% of agents recording their passwords, either insecurely (in a spreadsheet, for example), or securely (using a password management application).
- 100% of agents recording their passwords, again securely or insecurely.
- The simulation was executed 100 times, and the resulting values for the following were averaged across all 100 simulations.
- Good and bad logins.
- % shared and % stolen passwords.
- Number of lockout events.
- Number of malicious logins and number of hacker logins during the 100 days.

In order to confirm the choice of running the simulation 100 times we calculated the 95% confidence intervals (t-distribution) for 50 and 100 runs as shown in Table 6.

Table 6: 95% confidence Intervals for Number of Simulations

<table>
<thead>
<tr>
<th>% Recording Passwords</th>
<th>Confidence Intervals</th>
<th>Number of Lockout Events</th>
<th>Number of Bad Logins</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 Simulations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>39%</td>
<td>16.71% ± 0.48%</td>
<td>232.02 ± 6.01</td>
<td></td>
</tr>
<tr>
<td>100%</td>
<td>1.42% ± 0.31%</td>
<td>19.06 ± 4.83</td>
<td></td>
</tr>
<tr>
<td>100 Simulations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>39%</td>
<td>16.78% ± 0.35%</td>
<td>233.03 ± 6.30</td>
<td></td>
</tr>
<tr>
<td>100%</td>
<td>1.63% ± 0.23%</td>
<td>22.59 ± 3.35</td>
<td></td>
</tr>
</tbody>
</table>

6.4 These figures show that the extra 50 samples do not really improve the confidence (already very high) that the difference between the 39% and 100% cases is a genuine effect and not just a sampling artefact. We were thus satisfied that 100 samples was sufficient to demonstrate differences due to configuration settings.

6.5 Figure 16 shows that the number of bad logins shrinks from 14.35% to 1.6% when agents record their passwords. The number of lockout events, as shown in Figure 17, shows a 90% decrease. This is a large effect from a relatively small adjustment to the system.
6.6 Has the security of the system been compromised to the same extent? As Figure 18 shows, the number of passwords either shared or stolen has been reduced, with the percentage of unleased passwords increasing from 75.95% to 97.04%. This reflects the efforts of legitimate agents in the system. Figure 20 depicts the numbers of system breaches by malicious and hacker agents. There is indeed an increase in hacker successes. One possible explanation for this could be that forgotten passwords lead to password changes. SimPass systems do not permit use of previously-used passwords, so when passwords are frequently forgotten they change more often presenting hackers with a faster moving target. Even so, the increase in the number of attacks is relatively minor: from 15.34% to 17.95%. On the other hand, the number of malicious agent breaches decreases, the opposite of what is expected. This is due to the fact that malicious agents can and do make use of passwords that have been shared with them previously. That no longer happens since no one forgets passwords any more (the main causative behind sharing).

6.7 Table 5 shows the 95% confidence intervals for the malicious and hacker logins, which demonstrates that the differences are significant. Thus allowing people to write their passwords down does not lead to increased insider attacks - which is counter-intuitive. It is possible that we could mitigate the significant increase in hacker attacks by using other measures such as making passwords longer or more complex (and this will not impact on memorability since forgetting is no longer an issue).

<table>
<thead>
<tr>
<th>Table 5: 95% confidence Intervals for Malicious and Hacker Logins</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Recording Passwords</td>
</tr>
<tr>
<td>-----------------------</td>
</tr>
<tr>
<td>50 Simulations</td>
</tr>
<tr>
<td>39%</td>
</tr>
<tr>
<td>100%</td>
</tr>
<tr>
<td>100 Simulations</td>
</tr>
<tr>
<td>39%</td>
</tr>
<tr>
<td>100%</td>
</tr>
</tbody>
</table>

http://jasss.soc.surrey.ac.uk/16/3/3.html
What the simulations show is that, by removing the need for people to steal and share passwords, i.e., eliminate forgetting, you can actually strengthen the system. If you remove the need of well-intentioned users to engage in these activities one is left with only the efforts of malicious employees and external hackers to compromise the security of the system. These threats are not controlled by security policies, but rather by auditing and other technical and management controls. Moreover, consider the significant reduction in the number of lockouts. Each lockout has an associated expense since the person will not be able to work while waiting for the password to be replaced. If a help desk has to be involved in the replacement the expense will be greater still. These results make it worth returning to Figure 1 and revising it, as shown in Figure 20. The apparently obvious causative link from people writing passwords down, to passwords leaking, and the systems’ security being compromised, is not as clear cut as it appears to be. These findings should give system administrators pause, and make them think again before forbidding the writing down of passwords.

What these simulations show at a more abstract level is that one needs to tackle the cause of the problem rather than the symptoms to increase the overall security of an organisation.
Conclusion

7.1 This paper has described the SimPass simulation model and engine, a first attempt to provide a mechanism for testing the effects of security policy directives. This tool simulates a number of different pressures and impacts on users and allows researchers to experiment with different settings in order to arrive at a particular set of policies which will deliver better security. It is necessary to abandon traditional thinking which mandates and forbids particular user behaviours, especially when such behaviours are effectively natural attempts to cope with the surfeit of passwords that is so characteristic of modern life. Using the tool we can come up with interventions and test such interventions once the simulation has been configured according to the specific organisation. In effect, it supports a systemic meta-approach to the problem of system security. The focus moves away from the user to the organisation and addresses the issue of what policy writers can do to achieve real security improvements.

7.2 In summary, SimPass is a flexible tool with many customisable input parameters. It makes it possible to test the effects on organisational security of varying one or more of these parameters, while holding others constant. It is ideally suited to allow IT managers to project the effect of suggested policy changes, including regulations and recommendations intended to change staff behaviour, on the overall ability of the organisation to resist attack.

Appendix

8.1 The following simulation settings are configurable in SimPass:

- The number of days the simulation should run
- The percentage of agents who write down passwords
- The percentage of agents who secure these records
- The percentage of agents who will share passwords
- Days between password changes
- Number of hackers
- Number of agents
- Initial number of systems to be assigned to agents
- Percentage malicious agents
- Percentage agents who have the potential to be dishonest
- Percentage of trained system administrators
- Percentage systems visible from outside
- Probability that agents choose weaker passwords after a change
- Number of tries before lockout
- Whether lockouts should be implemented or not
- Percentage of systems that enforce password changes
- Percentage of systems that issue passwords (as opposed to allowing agents to choose them)
- Minimum password length
- Whether passwords require numerals
- Whether passwords require uppercase letters
- Whether passwords require special characters
- Whether agents work in open plan offices or not (can password entry be observed?)

Acknowledgements

We thank Joerg Denzinger for his very helpful comments on earlier drafts of this paper.

Notes

1 http://www.schneier.com/blog/archives/2006/12/realworld_passw.html


