Technical Report

UCAM-CL-TR-500
ISSN 1476-2986

Number 500

UNIVERSITY OF CAMBRIDGE

Computer Laboratory

The memorability and security of passwords – some empirical results

Jianxin Yan, Alan Blackwell, Ross Anderson, Alasdair Grant

September 2000

15 JJ Thomson Avenue
Cambridge CB3 0FD
United Kingdom
phone +44 1223 763500

http://www.cl.cam.ac.uk/

Technical reports published by the University of Cambridge Computer Laboratory are freely available via the Internet:

http://www.cl.cam.ac.uk/TechReports/

Series editor: Markus Kuhn

ISSN 1476-2986

The Memorability and Security of Passwords – Some Empirical Results

Jianxin Yan, Alan Blackwell, Ross Anderson, Alasdair Grant

Cambridge University Computer Laboratory

Abstract. There are many things that are 'well known' about passwords, such as that uers can't remember strong passwords and that the passwords they can remember are easy to guess. However, there seems to be a distinct lack of research on the subject that would pass muster by the standards of applied psychology.

Here we report a controlled trial in which, of four sample groups of about 100 first-year students, three were recruited to a formal experiment and of these two were given specific advice about password selection. The incidence of weak passwords was determined by cracking the password file, and the number of password resets was measured from system logs. We observed a number of phenomena which run counter to the established wisdom. For example, passwords based on mnemonic phrases are just as hard to crack as random passwords yet just as easy to remember as naive user selections.

1 Introduction

Many of the deficiencies of password authentication systems arise from the limitations of human memory. If humans were not required to remember the password, a maximally secure password would be one with maximum entropy: it would consist of a string as long as the system allows, consisting of characters selected from all those allowed by the system, and in a manner that provides no redundancy – i.e., totally random selection.

Each of these requirements is contrary to a well-known property of human memory. Firstly, human memory for sequences of items is temporally limited [1], with a short-term capacity of around seven plus or minus two items [2]. Second, when humans remember a sequence of items, those items cannot be drawn from an arbitrary and unfamiliar range, but must be familiar 'chunks' such as words or familiar symbols [2]. Third, human memory thrives on redundancy – we are far better at remembering information that can be encoded in multiple ways [3].

Password authentication therefore appears to involve a tradeoff. Some passwords are very easy to remember (e.g. single words in the user's native language), but also very easy to guess with dictionary searches. In contrast, some passwords are very secure against guessing but difficult to remember. In the latter case the

security of a superior password may be compromised due to human limitations, because the user may keep an insecure written record of it or resort to insecure backup authentication procedures after forgetting it[1].

This paper presents an empirical investigation of these tradeoffs in the context of an actual population of password users. Research in cognitive psychology has defined many limits of human performance in laboratory settings where experimental subjects are required to memorise random and pseudo-random sequences of symbols. It is very difficult to generalise from such research to password users, who can select the string themselves, are able to rehearse it while memorising, and need to recall it at regular intervals over a long period of time.

We show that this user context allows the exploitation of mnemonic strategies for password memorisation. There are many successful mnemonic techniques that can be used to achieve impressive performance when memorising apparently random sequences. Password alternatives such as "Pass Faces" exploit superior human memory for faces, for example [4]. However rather than changing the password authentication procedure, we propose changing the advice that is given to the user when selecting a password.

2 Existing Advice on Password Selection

Many large organisations give specific advice to new users about how to select a "good password". A good password, in terms of the above discussion, should aim to be reasonably long, use a reasonably large character set, but still be easy to remember. There are some subtleties about whether the attacker is going to try many passwords over a network or whether she's obtained a copy of the password file and is cracking it offline, but we propose to ignore these for the purposes of the present study.

We made an informal survey of advice given to new users at large sites, by searching on the Web for the terms "choose", "good" and "password". Many sites did not recognise the importance of memorability, merely emphasising resistance to brute-force search. Some typical pieces of advice were:

> "[A good] password should consist of mixed characters or special characters, and should not consist of words found in the dictionary. It should not be written down in an easily accessible place and especially not next to login. It may either be all in capital or small type letters."

[1] This doesn't mean we accept the common doctrine that writing passwords down is always wrong. For machines not in publicly accessible areas, it may be good sense to have a long random boot password written down in an envelope taped to the machine, as one can then have a strict policy that passwords are never under any circumstances to be disclosed over the phone. However, the prevention of 'social engineering' attacks is a separate research topic.

> "Use the output from a random password generator. Select a random string that can be pronounced and is easy to remember. For example, the random string 'adazac' can be pronounced a-da-zac, and you can remember it by thinking of it as 'A-to-Z'. Add uppercase letters to create your own emphasis, e.g., aDAzac.2"

> "Good passwords appear to be random characters. The wider the variety of characters the better. Mixing letters with numbers is better than letters alone. Mixing special characters with number and letters is better still."

One recommendation that seems increasingly popular is the "pass phrase" approach to password generation. A typical description of this is as follows:

> "A good technique for choosing a password is to use the first letters of a phrase. However, don't pick a well known phrase like 'An apple a day keeps the doctor away' (Aaadktda). Instead, pick something like 'My dog's first name is Rex' (MdfniR) or 'My sister Peg is 24 years old' (MsPi24yo)."

Of course this informal survey does not include sites where no advice at all is given on password selection. We believe that many sites simply tell new users the minimum requirement for a valid password (length and character set), and give no further advice regarding security or memorability. Others, in our experience, enforce rules such as

> "Passwords must be at least eight characters long and must contain at least two nonletter characters. They must also be changed at least once a month."

The usual response of users to such rules appears to be to devise a personal password generation system of which a simple example is 'Juliet03' for March, 'Juliet04' for April, and so on. This is clearly weak. Other attempts to compel user behaviour have backfired. For example, Patterson reports that when users were compelled to change their passwords and prevented from using the previous few choices, they changed passwords rapidly to exhaust the history list and get to their favourite password. A response, of forbidding password changes until after 15 days, meant that users couldn't change possible compromised passwords without help from the system administrator [9].

So the design of the advice given to users, and of the system-level enforcement which may complement this, are important problems which involve subtle questions of applied psychology to which the answers are not obvious.

The existing literature on password selection and memorability is surprisingly sparse. Grampp and Morris's classic paper on Unix security reports that after

software became available which forced passwords to be at least six characters long and have at least one nonletter, they made a file of the 20 most common female names, each followed by a single digit. Of these 200 passwords, at least one was in use on each of several dozen machines they examined [5]. Klein records collecting 13,797 password file entries from Unix systems and attacking them by exhaustive search; about a quarter of them were cracked. Password management guidelines from the US Department of Defense [7] recommended the use of machine generated random passwords.

Zviran and Haga [8] conducted an experiment in which they asked 106 students to choose passwords, writing them on a questionnaire. The questionnaires also assigned a random password to each student, and they were asked to remember both. Three months later, they found:

	Self-selected	Random
Successful recall:	35%	23%
Wrote it down:	14%	66%

However, the students were not actually using the password during the intervening three months. So although this provides a quanitiative point of reference for the difficulty of random passwords, it does not model a real operational environment closely.

3 Experimental Study

In order to investigate these trade-off factors in a real context of use, we have conducted an experiment involving 400 first-year students at our university. The experiment compared the effects of giving three alternative forms of advice about password selection, and measured the effect that this advice had on security and memorability of passwords.

The experimental subjects were students who had arrived in to start a degree in our School of Natural Sciences, which includes physics, chemistry, geology and meterials science. All Natural Sciences students are provided with an account on a central computing facility, using a user ID and randomly generated initial password. They also have access to a number of other facilities. At the time they receive these account details, students are generally advised to select their own password. Some students receive this advice informally from a computer officer in their department or hall of residence. Many students attend an introductory lecture to learn about the central facilities, followed by a tutorial session under the supervision of demonstrators.

4 Method

In October 1999, students attending the introductory lecture were told that they would be subjects (with their consent) in an experiment on password selection.

At the tutorial session they were then asked for consent and randomly assigned to one of three experimental groups. Each student was given a sheet of advice depending on the group that they had been assigned to. The three different types of advice were:

- students in the *control group* were given the same advice as in previous years, which was simply that: 'Your password should be at least seven characters long and contain at least one non-letter'
- students in the *random password group* were given a sheet of paper with the letters A–Z and the numbers 1–9 printed repeatedly on it. They were told to select a random password by closing their eyes and picking eight characters at random. They were advised to keep a written record with them until they'd memorised it.
- students in the *passphrase group* were told to choose a password based on a mnemonic phrase.

The text of the instructions given to the three groups is reproduced in the appendix.

The result which we expected was that the random password group would have stronger passwords than the passphrase group, but find them harder to remember and/or easier to forget; while the passphrase group would stand in the same relation to the control group.

So one month after the tutorial sessions, we took a snapshot of all password files, and conducted four types of attack on the passwords:

1. Dictionary attack: Simply use different dictionary files to crack passwords. This attack was attempted against all passwords.
2. Permutation of words and numbers: for each word from a dictionary file, permute with 0, 1, 2 and 3 digit(s) to construct possible password candidates. Also make common number substitutions, such a 1 for I, 5 for S etc. This attack was attempted against all passwords.
3. User information attack: Use user information collected from password files, e.g, userid, user full name, initial substring of name, to crack passwords. This attack was attempted against all passwords.
4. Brute force attack: we made this attack on any passwords that were only 6 characters long

We collected information on the distribution of password lengths, and on the number of cracked passwords, in each group. We monitored the number of times that users requested that their passwords be reset by the system administrators, on the assumption that passwords which were difficult to remember may be forgotten. In such a case the user would either have to ask for their password to be reset, or stop using the central facilities in favour of those provided elsewhere. We also surveyed all experimental subjects by email four months after the tutorial session, asking whether they had had any difficulty remembering their password. This survey asked the following questions:

1. How hard to did you find it to memorise your password, on a scale from 1 (trivial) to 5 (impossible)?
2. For how long did you have to carry around a written copy of the password to refer to? Please estimate the length of time in weeks.

We also tested the validity of our experimental sample by making the same attacks on the accounts of 100 first year students who had not attended the introductory lecture or received any experimental isntructions.

5 Results

Of the 300 students we asked, 288 consented to participate in the experiment. They were randomly allocated to experimental groups as follows:

Control group	95
Random password group	96
Passphrase group	97

The selected passwords were on average between 7 and 8 characters long (7.6, 8.0, 7.9 respectively) with no significant difference between the three groups. All groups chose slightly longer passwords than the further sample of 100 students who had not attended the introductory lecture (mean length 7.3, difference statistically significant at $t = 4.53$, $p < .001$).

The most successful cracking method was the permuted dictionary attack. Cracking based on user information was not successful in any case, probably because of the very limited amount of user information available in these password files (they do not include forenames, for example). All six-character passwords were successfully cracked using a brute-force attack. The summary of the number of cracked passwords is as follows (with brute-force attacks treated separately):

Control group	30 (32%) +3 brute force
Random password group	8 (8%) +3 brute force
Passphrase group	6 (6%) +3 brute force
Comparison sample	33 (33%) +2 brute force

All six character passwords are susceptible to brute force attack. The experimental password selection advice had no effect on this. In each experimental condition a small number of users ignored the advice regarding password length and chose an insecure password. This also occurred among the comparison sample.

Of the passwords that were longer than six characters, far more of these were cracked successfully in the control group than in either the random character or pass phrase group (significant at $\chi^2 = 24.8$, $p < .001$). The proportion of passwords cracked in the control group was lower than in the comparison sample

(for example, 13% in the comparison sample used 6-character passwords versus 5 % in the control group; while 13 passwords in the comparison sample were verbatim doctionary words versus 3 in the control group).

For those passwords that were cracked successfully in the random character and pass phrase groups, all the cracked passwords were dictionary words, or permutations of dictionary words and numbers, that were not compliant with the advice given to the student. These results, together with the number of six-character passwords, provide a reasonable estimate of the level of user non-compliance with password selection advice.

We also observed that nobody used special characters (i.e., neither letters nor numbers) except in the passphrase group, whose instructions had given examples of passwords containing punctuation. So a strong lead in the direction of passwords containing a mix of alpha, numeric and special characters seems to be advisable.

Very few users asked the system administrator to reset their passwords. Within a period of three months after the tutorial session, the number of administrator resets within each group were as follows:

Control group	2
Random password group	1
Passphrase group	3

242 students replied to the email survey, of which 13 responses indicated that the students had not used their accounts, or had dropped out of the course. Of the valid responses, there was a clear difference between the groups:

		Diff	Weeks
Control group	80	1.52	0.7
Random password group	71	3.15	4.8
Passphrase group	78	1.67	0.6

Users assigned to the random password group reported that they found their passwords more difficult to remember (significant at $t = 8.25$, $p < .001$), and that they carried a written copy of their passwords for far longer (significant at $t = 6.41$, $p < .001$). This confirms the results of Zviran and Haga in an operational setting.

The differences in response rates were not significant, so we do not believe our results were significantly skewed by students in the random password group finding our advice so difficult that they gave up using the computer facilities.

It is worth noting that many of the random character group were still carrying the written copy of the password at the time of the survey, so they had effectively been unable to memorise the password.

6 Discussion

This study confirms a number of widely held folk beliefs about passwords, and debunks some others.

1. The first folk belief is that users have difficulty remembering random passwords. This belief is **confirmed**.
2. The second folk belief is that passwords based on mnemonic prases are harder for an attacker to guess than naively selected passwords. This belief is **confirmed**.
3. The third folk belief is that random passwords are better than those based on mnemonic phrases. However, each appeared to be just as strong as the other. So this belief is **debunked**.
4. The fourth folk belief is that passwords based on mnemonic phrases are harder to remember than naively selected passwords. However, each appeared to be just as easy to remember as the other. So this belief is **debunked**.
5. The fifth folk belief is that by educating users to use random passwords or mnemonic passwords, we can gain a significant improvement in security. However, both random passwords and mnemonic passwords suffered from a non-compliance rate of about 10% (including both too-short passwords and passwords not chosen according to the instructions). While this is better than the 35% or so of users who choose bad passwords with only cursory instruction, it is not really a huge improvement. The attacker may have to work three times harder, but in the absence of password policy enforcement mechanisms there seems no way to make the attacker work a thousand times harder. In fact, our experimental group may be about the most compliant a systems administrator can expect to get. So this belief appears to be **debunked**.

The work reported in this paper is merely a first step towards a better understanding of the applied psychology aspects of computer security. Many questions remain to be answered, and we plan to continue our experiments with future cohorts of students.

In the meantime, our tentative recommendations for system administrators are as follows.

- Users should be instructed to choose mnemonic based passwords as these are just as memorable as naively selected passwords while being just as hard to guess as randomly chosen ones. So they give the best of both other options.
- Size matters. With systems like Unix which limit effective password lengths to eight characters, users should be told to choose passwords of exactly eight characters. With systems such as Netware which allows 14 characters but are not case-sensitive, one might encourage users to choose passwords of ten or more characters length; perhaps this will further encourage the use of

mnemonics. (This is a topic for next year's experiment, as is enforcement generally.)
- Entropy per character also matters. Users should be told to choose passwords that contain numbers and special characters as well as letters. If such a lead isn't given, then most of them will choose passwords from a very small subset of the total password space.
- Compliance is the most critical issue. In systems where users can only put themselves at risk, it may be prudent to leave them to their own devices. In that case, it must be expected that about 10% will choose weak passwords despite the instruction given. In systems where a user's negligence can impact other users too (e.g., in systems where an intruder who gets a single user account can rapidly become root using well known and widely available techniques), consideration should be given to enforcing password quality by system mechanisms.
- If there is a benefit to be had from the use of centrally assigned random passwords, it appears to come from the fact of central assignment (which enforces compliance) rather than randomness (which can be achieved just with mnemonic phrases).

An interesting and important challenge is to find compliance enforcement mechanisms which work well with mnemonic password choice. We expect that password checkers, which verify that a password isn't part of a known weak subset of the password space, may be an effective tool. An experimental test of this expectation is one of our projects for the next academic year.

References

1. GJ Johnson, in *Psychological Review* v 98 no 2 (1991) pp 204–217
2. GA Miller, "The magical number seven, plus or minus two: Limits on our capacity for processing information", in *Psychological Review* v 63 (1956) pp 81–87
3. A Paivio , "The empirical case for dual coding", in *Imagery, Memory and Cognition: Essays in honor of Allan Paivio*, JC Yuille (Ed), Erlbaum, Hillsdale, NJ (1983); pp 307–322
4. H Davies, "Physiognomic access control", in *Information Security Monitor* v 10 no 3 (Feb 95) pp 5–8
5. FT Grampp, RH Morris, "UNIX Operating System Security", *AT&T Bell Laboratories Technical Journal* v 63 no 8 (Oct 84) pp 1649–1672
6. DV Klein, "Foiling the Cracker; A Survey of, and Improvements to Unix Password Security", Proceedings of the USENIX Security Workshop. Portland, Oregon: USENIX Association, Summer 1990; `http://www.deter.com/unix/`; expanded as a technical report from SEI, 1992
7. Department of Defense, *'Password Management Guideline'*, CSC-STD-002-85 (1985)
8. M Zviran, WJ Haga, "A comparison of password techniques for multilevel authentication mechanisms", in *Computer Journal* v 36 no 3 (93) pp 227–237
9. B Patterson, letter to *Communications of the ACM* v 43 no 4 (Apr 2000) pp 11–12

Appendix

Here is the text of the three instruction sheets given to the three groups.

Control group

This sheet offers some advice on how to choose a good computer password. We are giving you this sheet as part of the password security experiment that was described in your introductory lecture. Different people are receiving different advice (but all advice should result in passwords at least as secure as you would choose if not participating in the experiment). Please do not discuss the experiment, this advice, or your choice of password with your friends.

Please log on using the initial password you have been issued, and choose a new password not known to anybody else. The 'Windows NT Tutor' tells you how to do this on pages 1.6-1.7.

Your password should be at least seven characters long and contain at least one non-letter.

If you have already changed your initial password to one of your choice, and your new password meets this standard, then you do not need to change it again. However we strongly recommend that you change your password from time to time – at least once a term. As the experiment will run for the duration of this academic year, please keep this sheet and use this advice again when you choose your new passwords for Lent and Easter.

6.1 Random password group

This sheet offers some advice on how to choose a good computer password. We are giving you this sheet as part of the password security experiment that was described in your introductory lecture. Different people are receiving different advice (but all advice should result in passwords at least as secure as you would choose if not participating in the experiment). Please do not discuss the experiment, this advice, or your choice of password with your friends.

A secure password is one that is very difficult to guess. Words that appear in a dictionary, or the names of people or places, are easy to guess. The most difficult passwords to guess are random sequences of letters. To help you choose a random sequence of letters for your password, we have printed a grid of random letters overleaf. Choose your password by closing your eyes and pointing at a random place on the grid. Choose eight characters this way and write them down on a scrap of paper.

Now log on using the initial password you have been issued, change your password to the new random password which you have chosen. The 'Windows NT Tutor' tells you how to do this on pages 1.6-1.7.

You may find your password difficult to remember at first. Make sure that the scrap of paper on which you have written it is in a secure place, such as the back of your wallet or purse.

You should find that once you have entered it a dozen times or so, you will be able to remember it. Once you are sure you can remember it, destroy the scrap of paper where you wrote it down.

Finally, we strongly recommend that you change your password from time to time – at least once a term. As the experiment will run for the duration of this academic year, please keep this sheet and use this advice again when you choose your new passwords for Lent and Easter.

6.2 Passphrase group

This sheet offers some advice on how to choose a good computer password. We are giving you this sheet as part of the password security experiment that was described in your introductory lecture. Different people are receiving different advice (but all advice should result in passwords at least as secure as you would choose if not participating in the experiment). Please do not discuss the experiment, this advice, or your choice of password with your friends.

To construct a good password, create a simple sentence of 8 words and choose letters from the words to make up a password. You might take the initial or final letters; you should put some letters in upper case to make the password harder to guess; and at least one number and/or special character should be inserted as well. Use this method to generate a password of 7 or 8 characters.

An example of such a composition might be using the phrase is "It's 12 noon I am hungry" to create the password "`I's12&Iah`" which is hard for anyone else to guess but easy for you to remember. By all means use a foreign language if you know one: the password "`AwKdk.Md`" from the phrase "Anata wa Kyuuketsuki desu ka ... Miyu desu" would be an example. You could even mix words from several languages. However, do not just use a word or a name from a foreign language. Try being creative!

Now log on using the initial password you have been issued, change your password to the new password which you have chosen. The 'Windows NT Tutor' tells you how to do this on pages 1.6–1.7. Do not write your new password down.

Finally, we strongly recommend that you change your password from time to time – at least once a term. As the experiment will run for the duration of this academic year, please keep this sheet and use this advice again when you choose your new passwords for Lent and Easter.