Securing Passwords Beyond Human Capabilities with a Wearable Neuro-Device

Miguel Angel Lopez-Gordo[1], Jesus Minguillon[2(✉)], Juan Francisco Valenzuela-Valdes[1], Pablo Padilla[1], Jose Luis Padilla[3], and Francisco Pelayo[1]

[1] Department of Signal Theory, Communications and Networking - CITIC, University of Granada, Granada, Spain
{malg,juanvalenzuela,pablopadilla,fpelayo}@ugr.es
[2] Department of Computer Architecture and Technology - CITIC, University of Granada, Granada, Spain
minguillon@ugr.es
[3] Department of Electronics and Computers Technology, University of Granada, Granada, Spain
jluispt@ugr.es

Abstract. The election of strong passwords is a challenging task for humans that could undermine the secure online subscription to services in mobile applications. Composition rules and dictionaries help to choose stronger passwords, although at the cost of the easiness to memorize them. When high-performance computers are not available, such as in mobile scenarios, the problem is even worse because mobile devices typically lack good enough entropy sources. Then, the goal is to obtain strong passwords with the best efficiency in terms of level of entropy per character unit. In this study, we propose the use neuro-activity as source of entropy for the efficient generation of strong passwords. In our experiment we used the NIST test suite to compare binary random sequences extracted from neuro-activity by means of a mobile brain-computer interface with (i) strong passwords manually generated with restrictions based on dictionary and composition rules and (ii) passwords generated automatically by a mathematical software running on a work station. The results showed that random sequences based on neuro-activity were much more suitable for the generation of strong passwords than those generated by humans and were as strong as those generated by a computer. Also, the rate at which random bits were generated by neuro-activity (4 Kbps) was much faster than the passwords manually generated. Thus, just a very small fraction of the time and cognitive workload caused to manually generate a password has enough entropy for the generation of stronger, shorter and easier to remember passwords. We conclude that in either mobile scenarios or when good enough entropy sources are not available the use of neuro-activity is an efficient option for the generation of strong passwords.

Keywords: Wearable brain-computer interfaces · Neuro-activity · Secure passwords

J.M. Ferrández Vicente et al. (Eds.): IWINAC 2017, Part II, LNCS 10338, pp. 87–95, 2017.
DOI: 10.1007/978-3-319-59773-7_10

1 Introduction

Currently, humans enjoy on-line services and applications such as social networks, web browsing, email, e-commerce, e-Government services, etc. Wireless technologies contribute to the mobility by enabling access to internet working abroad, such as in the tube, in the bus stop or in the canteen while having lunch. Although people typically use passwords as a simply and straight-forward way to gain access to on-line services, it becomes challenging for humans not only to generate a relatively high number of strong passwords, but at the same time to memorize them.

Passwords strength is related to the probability of guessing it by a repetitive at-tack. It depends on both the entropy of the password and the way the server limits the number of unsuccessful attempts. For a giving number of unsuccessful attempts, the entropy of the password determines the assurance level during on-line authentication.

The entropy of a password determines how difficult it is to discover. It is typically measured in terms of bits (H) and in the case of a password of length n, generated by a set with m equally likely characters; it can be easily measured as formulated in (1).

$$H(bit) = log_2(m^n) \tag{1}$$

Computers can also be used to generate passwords based on pseudorandom binary sequences originated from a seed. One inconvenient of this approach is that the generation of the password is determinist. That is, for a given cryptographic algorithm, any password can be reproduced with the only knowledge of the seed, thus being a vulnerability in the password generation.

In mobile context, generation of strong passwords presents an even a more complicated scenario. On the one hand, mobile devices are typically low-cost electronics with very limited hardware resources unable to support high computational loads and storage capabilities. Then, they have not good enough entropy sources such as those from a PC [1] and this justifies the analysis of other physical sources of entropy such as bio-signals. The latter could be a problem when the user is required for an on-line subscription that requires the strongest possible password. On the other hand, the election and memorization of strong passwords is challenging for humans. A strong password typically requires a large number of characters that, in turn, are difficult to remember. Again, the use of a good entropy source guarantees a high efficiency in the quantity of entropy per character.

In this preliminary study, we propose the use of a wearable brain-computer interface (wBCI) as a source of entropy for the generation of strong passwords. In our experiment we used the NIST test suite to compare strong passwords (i) manually generated by a participant with restrictions based on dictionary and composition rules with (ii) the participants neuro-activity while he was manually generating the passwords and with (iii) passwords generated automatically by a mathematical software running on a work station. The promising results suggest that in mobile scenarios the use of neuro-activity is an efficient option for the generation of strong passwords.

2 Related Works

In our days, a large number of the people live connected to networks. They freely commute from home to work, visit leisure spaces and spend their time connected to social networks, or corporative and private services and applications. Every day, we install apparently-free-of-charge applications in our smartphones. Typically we are only required an email account, an username and a password. In occasions we may not need a strong password when we subscribe for a trivial temporal service, but sometimes we would like to provide a strong password, as secure as possible. When a computer is not available, such as in the context of Wireless Body Area Networks (WBANs), bio-signals from wearable devices and bio-inspired solutions can be used to generate cryptographic keys [2,3]. For instance, pulse oximetry, pulse interval, heart rate variability (HRV) or electrocardiography (ECG) [3–6] have been employed as physical sources of entropy in WBANs.

Computers can be used to generate pseudo random seeds for passwords. This is a deterministic process that solely depends on the seed and the generation algorithm. Then, the predictability of the seed is the main vulnerability. Standards of the industry such as the RFC4086 [7] proposes hardware-based random sources to produce unpredictable seeds (e.g. ring oscillators, disk drive timing, system clock, mouse motion, CPU interruptions, etc.). However, under some circumstances, some of these methods (e.g., vectors based in date or time) are very predictable [8]. Sensors integrated in smartphones have been used when a computer was not available. Authors in [9] show that one of the best entropy source could be the microphone. However, it could be hacked or force the random number by means of a simple auditory attack.

Humans can also produce strong passwords. The mayor inconvenient is our inability to memorize them, especially when we need to access or subscribe to many services. When we are asked for a password in an on-line registration, we typically produce a very weak one. Authors in [10] performed a large-scale study about users habits in web password. The study confirmed (i) the poor quality of user passwords; (ii) users reuse them and (iii) they often forget them. Restrictions imposed by services (e.g., composition rules and dictionaries) typically improve the entropy, although still far from being a random password. Another recent study [11] concludes that people habits have changed in different ways from previous studies and surveys. Figure 1 shows an estimation of the bits of entropy in a password when users select it without restrictions, with dictionary and composition rules and completely random [12]. From visual inspection of manually generated passwords (curves with triangles, circles and asterisk), we conclude that (i) a minimum of 8–10 characters is recommended to obtain a substantial benefit using dictionaries and composition rules with respects to no checks; (ii) a completely random password of just 6 characters conveys as much entropy as a manually generated password without checks or either with dictionary of 24 characters length, or with dictionary and composition rules of 18 characters length. In summary, the use of a good source of entropy can help the user to choose an efficient number of characters for a strong password not difficult to remember.

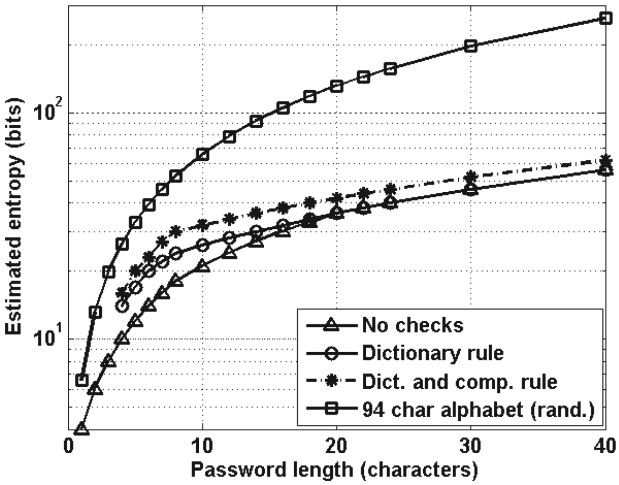

Fig. 1. Estimation of the passwords entropy. The curve with rectangles (upper one) represents the maximum theoretical limit of entropy with 94 equally likely printable characters. The curve with triangles (bottom one) represents a typical password manually generated without any restrictions (i.e., easy to remember). The curves in between represent the entropy when restrictions based on dictionary (curve with circles) and dictionary together with composition rules (curve with asterisks) are imposed. Adapted from [12]

Recent advances in technology have given rise to the development of wearable and mobile BCI that can be part of a wearable WBAN ecosystem [13] (see [14] for a review). Currently, they are wearable and low-cost devices capable to ubiquitously monitor our cognitive and electro-physiological activity. Other WBAN devices such as ECG measurers have been reported as good entropy sources for the generation of cryptographic keys [15]. In comparison with ECG, electroencephalography (EEG) theoretically presents better characteristics for the generation of random binary sequences. For instance, EEG is the result of both endogenous and exogenous cognitive and physiological processes sustained by billions of interconnected neurons. The additive sum of field potentials generated by each of them gives rise to the EEG signals. They depend on the location of the electrode over the scalp and they also need physical contact for their acquisition. In summary, it is very unlikely to hack, reproduce or guess spontaneous EEG activity. EEG is an estimable physical source of entropy that requires a minimum preprocessing and hardware resources usage. Another advantage is the bit rate. ECG can generate circa 16 random bps [3], whereas EEG which is typically acquired at 1000 samples/s with 24 bits of resolution per sample, could offer a maximum of 24 Kbps per EEG channel. Even after a severe process of bit discarding, EEG could be used to generate large data sets of random numbers. For instance, for some NIST tests, a sequence of circa 2Mb is recommended. In this case, ECG and EEG would take more than one day and just a few minutes respectively.

Finally, in [16] the authors proposed the use of a wBCI for the generation of a data set with secure keys to encrypt communication of WBAN devices. This was presented as a contribution to the people-centric Internet of Things. From the best of our knowledge this was the first study in which EEG was evaluated as source of entropy to generate passwords.

3 Experimental Design

The objective of this study was to compare the randomness of binary sequences generated by neuro-activity with the derived from manually generated passwords with severe restrictions and with the automatically generated by a computer. We defined three data sets:

Data Set I: Passwords generated by the neuro-activity exerted by one participant when he was manually generating passwords. An EEG electrode was placed on Cz position of the 10–20 International System [17]. This position (top of the head, see Fig. 2) was chosen because it matches reports of successful studies of random number generation [16]. The channel was referenced and grounded to the left ear lobe. The impedance of the electrodes was much lower than the input impedance of the acquisition system. EEG signals were recorded at 500 Hz and 24 bits per sample with a RABIO w8 developed by the University of Granada.

No preprocessing technique or whitening pre-processing was performed on the EEG data. Even the use of a notch filter to remove power-line coupling was discard-ed. Only raw EEG data was used. Only the eight least significant bits of

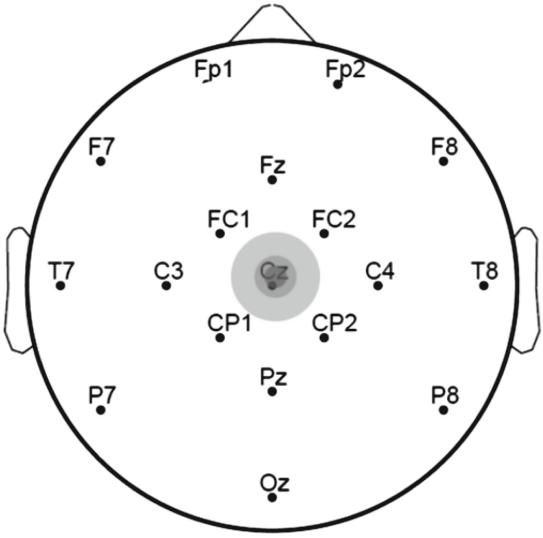

Fig. 2. Top view of the electrical montage. One active EEG channel was located on the top of the head.

each 24-bit data were extracted to generate the random sequence of bits. This generated a random binary sequence at a rate of 4 Kbps. For the comparison with the manually generated passwords (Data set II), only the first 26 Kb were used. For the comparison with the computer (Data set III), the total EEG data were used. The EEG was recorded during the manual generation of passwords. It took approximately 46 min for 40 manual passwords. Then the binary sequence obtained was approximately 11 Mb.

Data Set II: It contains passwords manually generated by a participant with restrictions. The participant (a male) was told to complete 40 independent passwords each one of 12-character length. It took approximately 46 min. He was told to use keys from the 94-character set to produce somehow easy-to-remember passwords. To ensure a high level of randomness, first he had to choose a password classified as strong and without deductions by a password meter available in Internet (http://www.passwordmeter.com/). These passwords were 10-character length. Af-terwards, two more characters were added to complete 12 characters. These two additional characters were chosen to obtain the maximum number of random bits as measured by a tool available in Internet (http://rumkin.com/tools/password/passchk.php). This slow methodology was able to produce only 12-character-length passwords at a rate slower than 1 password per minute. The rest of the passwords of the data set were generated by another participant. He was told to generate passwords, as random as possible. No restrictions about the easiness to memorize were imposed. The whole data set included passwords with a total of 4480 characters. The six least significant bits were extracted to generate the random binary sequence of the same length as Data set I (circa 26 Kb).

Data Set III: Passwords generated by a computer. A PC (Windows 7, Core i7, Matlab 2013b) was used to generate random binary sequence of the same length as Data set I (circa 11 Mb).

4 Statistical Analysis

We used the NIST Test Suite [13] for the assessment of the randomness of the generated sequences. NIST is a tool largely used for validation of secure keys [9–18]. In this experiment we submitted random binary sequences to compare Data set I with Data set II and III. NIST performed 15 tests. For the sake of simplicity, only the most relevant tests are reported in this article. See [14] for comprehensive information about NIST.

5 Results

The randomness of Data set I (based on EEG) versus Data set III (based on a computer) is presented in Table 1. The corresponding to Data set I versus Data set II (manually generated passwords) are presented in Table 2.

Table 1. Randomness evaluation performed by NIST. The sample size was 112 binary sequences of 0.1 million bits each sequence. * stands for p-values $< \alpha$ or under the minimun pass rate (107). In bold: test values in which Data set I performed equal or better.

Data set I (EEG)		Data set II (computer)		
p-value	**Pass rate**	**p-value**	**Pass rate**	**Test name**
7.2E−01	110/112	6.1E−01	112/112	Frequency
8.4E−01	**112/112**	8.9E−01	109/112	BlockFrequency
1.8E−01	110/112	6.5E−01	112/112	CumulativeSums
4.0E−01	110/112	2.3E−02	112/112	CumulativeSums
4.3E−01	109/112	7.9E−02	110/112	Runs
5.2E−01	**112/112**	2.1E−01	110/112	LongestRun
7.4E−01	**112/112**	1.8E−01	112/112	Rank
6.7E−01	**111/112**	4.5E−01	110/112	FFT
0.0E+00*	**112/112**	0.0E+00*	112/112	Universal
2.5E−01	**112/112**	8.8E−02	110/112	ApproximateEntropy
2.7E−01	**111/112**	7.6E−01	110/112	Serial
5.5E−01	110/112	5.5E−01	111/112	Serial
1.6E−01	**112/112**	6.1E−01	112/112	LinearComplexity

Table 2. Randomness evaluation performed by NIST. The sample size was 12 binary sequences of 2240 bits each sequence. * stands for p-values $< \alpha$ or at the minimum pass rate (10). In bold: test values in which Data set I performed better or the same.

Data set I (EEG)		Data set II (computer)		
p-value	Pass rate	p-value	Pass rate	Test name
8.9E−03	**12/12**	3.0E−06*	10/12*	Frequency
3.5E−01	**12/12**	8.9E−03	10/12	BlockFrequency
1.8E−02	**12/12**	3.0E−06*	10/12*	CumulativeSums
3.5E−01	**12/12**	3.0E−06*	11/12	CumulativeSums
5.3E−01	**12/12**	0.0E+00*	2/12*	Runs
3.5E−02	**12/12**	5.3E−01	11/12	LongestRun
8.9E−05*	11/12	1.0E−06*	12/12	Rank
2.1E−01	**12/12**	4.3E−03	11/12	FFT
0.0E+00*	**12/12**	0.0E+00*	12/12	Universal
0.0E+00*	**0/12***	0.0E+00*	0/12*	ApproximateEntropy
6.7E−02	**12/12**	4.4E−04	8/12*	Serial
3.5E−01	**12/12**	6.7E−02	8/12*	Serial
3.5E−01	**12/12**	2.0E−03	11/12	LinearComplexity

6 Discussion and Conclusion

In this preliminary study we have evaluated the use or neuro-activity as a physical source of entropy that can be useful for the generation of strong password in mobile applications. Passwords generated by means of a wearable brain-computer interface were much more random and complex that those manually generated by the human (see Table 2), even when they were generated with restrictions based on dictionaries and composition rules or even generated as random as possible without restrictions about the easiness to memorize. The required cognitive workload and time was just a small fraction of those required for the manually generated passwords. In this study, only 6 s of neuro-activity contained more randomness than 46 min of manual generation of passwords. Efficiency in terms of entropy per character unit, gives rise to shorter (and easier to remember) random passwords. When we compared the randomness in neuro-activity with the one generated by a computer, the test yielded similar or inconclusive results. Then it is difficult to evaluate which data set performed better. We conclude that in mobile scenarios without good enough entropy sources, the use of neuro-activity is an efficient option for the generation of strong passwords. The current development in wireless brain-computer interfaces makes it a feasible solution.

In future, we plan to extend this study to larger data sets. In addition, a comparison with other available sources of entropy in low-cost wearable devices such as smartphones will be included.

Acknowledgments. This work was supported by Nicolo Association for the R+D in Neurotechnologies for disability, the research project P11-TIC-7983 of Junta of Andalucia (Spain), the Spanish National Grant TIN2015-67020-P, co-financed by the European Regional Development Fund (ERDF) and the Spanish National Grant TIN2016-75097-P (AEI/FEDER, UE).

References

1. Chang, C.-C., Wu, H.-L., Sun, C.-Y.: Notes on secure authentication scheme for IoT and cloud servers. Pervasive Mob. Comput. **24**, 210–223 (2016)
2. Altop, D.K., Levi, A., Tuzcu, V.: Deriving cryptographic keys from physiological signals. Pervasive and Mobile Computing (2016)
3. Zheng, G., Fang, G., Shankaran, R., Orgun, M., Zhou, J., Qiao, L., Saleem, K.: Multiple ECG fiducial points based random binary sequence generation for securing wireless body area networks. IEEE J. Biomed. Health Inf. 1–9 (2016)
4. Venkatasubramanian, K.K., Banerjee, A., Gupta, S.K.S.: PSKA: usable and secure key agreement scheme for body area networks. IEEE Trans. Inf. Technol. Biomed.: a publication of the IEEE Eng. Med. Biol. Soc. **14**, 60–68 (2010)
5. Israel, S.A., Irvine, J.M., Cheng, A., Wiederhold, M.D., Wiederhold, B.K.: ECG to identify individuals. Pattern Recogn. **38**, 133–142 (2005)
6. Poon, C., Zhang, Y.-T., Bao, S.-D.: A novel biometrics method to secure wireless body area sensor networks for telemedicine and m-health. IEEE Commun. Mag. **44**, 73–81 (2006)

7. Eastlake, D., Schiller, J., Crocker, S.: Randomness requirements for security (2005)
8. Rukhin, A., Soto, J., Nechvatal, J., Smid, M., Barker, E.: A statistical test suite for random and pseudorandom number generators for cryptographic applications. Technical report NIST Special Publication 800–22 Revision 1a. National Institute of Standards and Technology (2010)
9. Wallace, K., Moran, K., Novak, E., Zhou, G., Sun, K.: Toward sensor-based random number generation for mobile and IoT devices. IEEE Internet Things J. **3**, 1189–1201 (2016)
10. Florencio, D. Herley, C.: A large-scale study of web password habits, pp. 657. ACM Press (2007)
11. Shen, C., Yu, T., Xu, H., Yang, G., Guan, X.: User practice in password security: an empirical study of real-life passwords in the wild. Comput. Secur. **61**, 130–141 (2016)
12. Burr, W.E., Dodson, D.F., Newton, E.M., Perlner, R.A., Polk, W.T., Gupta, S., Nabbus, E.A.: Electronic authentication guideline. Technical report NIST SP 800–63-1. National Institute of Standards, Technology, Gaithersburg, MD (2011). 10.6028/NIST.SP.800-63-1
13. Lopez-Gordo, M.A., Pelayo Valle, F.: Brain-Computer interface as networking entity in body area networks. In: Aguayo-Torres, M.C., Gómez, G., Poncela, J. (eds.) WWIC 2015. LNCS, vol. 9071, pp. 274–285. Springer, Cham (2015). doi:10.1007/978-3-319-22572-2_20
14. Wu, F.-J., Kao, Y.-F., Tseng, Y.-C.: From wireless sensor networks towards cyber physical systems. Pervasive Mob. Comput. **7**, 397–413 (2011)
15. Zhang, Z., Wang, H., Vasilakos, A.V., Fang, H.: ECG-cryptography and authentication in body area networks. IEEE Trans. Inf. Technol. Biomed. **16**, 1070–1078 (2012)
16. Valenzuela-Valdes, J.F., Lopez, M.A., Padilla, P., Padilla, J.L., Minguillon, J.: Human neuro-activity for securing body area networks: application of brain-computer interfaces to people-centric internet of things. IEEE Commun. Mag. **55**, 62–67 (2017)
17. Jasper, H.: Report of the committee on methods of clinical examination in electroencephalography. Electroencephalogr. Clin. Neurophysiol. **10**, 370–375 (1958)
18. Hong, S.L., Liu, C.: Sensor-based random number generator seeding. IEEE Access **3**, 562–568 (2015)