GeoGraphical passwords

Ziyad S. Al-Salloum

ZSS – Research,
P.O. Box 16211,
Ras Al Khaimah, UAE
E-mail: zss@zss.net

Abstract: It is indeed becoming more challenging for users to maintain different strong passwords for their ever increasing accounts. The lack of secure access credentials has recently led to the compromise of millions of users passwords stored in popular websites, due to guessing, dictionary, or brute force attacks. In this paper, we address the conventional password problem and propose a novel, simple, and practical access credential that would provide secure access to different entities and mitigate many vulnerabilities associated with current password based schemes. We name our proposal GeoGraphical passwords, which is an access credential based on geographical information. The credential utilise the remarkable human ability to remember places as a way to provide safe access, where users can select geographical locations (such as favourite places, mountains, trees, rivers or others) as their access credential to different systems. We develop a prototype to show one possible implementation of GeoGraphical passwords and improve the credential's ability to protect itself from common password threats in an attempt to mitigate the frequent risks associated with – the difficult to remember, construct, and maintain – conventional passwords.

Keywords: passwords; graphical passwords; GeoGraphical passwords; access control; access credential; authentication.

Reference to this paper should be made as follows: Al-Salloum, Z.S. (2014) 'GeoGraphical passwords', *Int. J. Security and Networks*, Vol. 9, No. 1, pp.56–62.

Biographical notes: Ziyad S. Al-Salloum received his BSc (2003) in Computer Science from University of South Alabama in USA and his MSc (2006) and PhD (2012) in Information Security from Royal Holloway, University of London in UK. His research and publications covers different areas of cyber security including, defensive computer worms and authentication. He has worked in both academia and industry before founding ZSS, a cyber security organisation providing novel solutions and consultations on a variety of topics in information security.

1 Introduction

In 2011 a famous website (LinkedIn.com) has announced the exposure of millions of its hashed passwords (BBC News Technology, 2012), it was a matter of days for these passwords to become publicly known after cyber criminals deciphered them. Fifty million passwords, in another breach, have been stolen from the famous Evernote service, leading the cooperation to issue a security notice to rush its clients to reset their – soon to be cracked – passwords (PC World, 2013). Twitter also has been under attacks that made the encrypted passwords of around 250,000 of its users exposed to cybercriminals (Lord, 2013).

Even passwords that were constructed by highly skilled cybercriminals were deciphered, such as the one used to control the Flame Botnet, where the password was: 900gage!@# which happens not to be so obvious (Kaspersky Labs, 2012).

Moreover, a study revealed, after analysing 32 million publicly leaked passwords from the gaming website RockYou, showed that "passwords were generally short, conform to existing language patterns and show a great deal of overlap", unfortunately, the passwords were unencrypted (Devillers, 2010). Even some military personnel – whom are supposed to adopt more restrictive password policies – failed to use strong passwords, as revealed by Booz Allen Hamilton breach incident (Imperva Data Security Blog, 2011).

Furthermore, in another study, by Joseph Bonneau which analysed around 70 million anonymised yahoo passwords, finding that for an attacker guessing the passwords online (using popular guesses), passwords would only provide 10 bits of security, while only 20 bits are available if the attacker brute forced the passwords offline (Bonneau, 2012), providing a very weak protection.

Such incidents indicate that we need to revisit the approaches we use to authenticate users or the ways users construct their passwords – Google has considered authentication as one of the biggest threats towards cloud computing and highlighted the need to displace conventional text passwords (Grosse and Upadhyay, 2013).

Proposing an effective replacement of conventional passwords could reduce 76% of data breaches, based on an analysis of more than 47,000 reported security incidents (The Data Breach Investigations Report, 2013).

In this paper we propose the following:

- a novel access credential based on geographical information
- improve the access credential ability to protect itself from dictionary, brute force, and rainbow attacks
- propose one possible implementation of the access credential and demonstrate it.

The reminder of this paper is organised as follows. In Section 2 we come across few knowledge-based authentication techniques, while Section 3 highlights conventional passwords vulnerabilities and introduces GeoGraphical Passwords. Section 4, thereafter, provides one possible prototype (or implementation) of GeoGraphical passwords. After which Section 5 analyse and improve the guessing entropy of the new access credential, followed by Section 6, which discusses the strengths of GeoGraphical passwords under different protective measures. Finally, our conclusions are then described in Section 7.

2 Background

Searching for alternatives of conventional passwords authentication systems has caught the attention of many researchers. While textual passwords remains the dominant technique in authentication (Herley et al., 2009), other knowledge-based authentication approaches do exist.

In mid 90s, Blonder (1996) introduced the concept of *Graphical Passwords*. In his work the user has to tap an image at different regions in a pre-determined sequence for her to get access. Another knowledge-based alternative of conventional passwords, was presented by Jermyn et al. (1999), which represent a rectangular 2D grid where a user can draw a shape using a stylus as her password. The user should reproduce the shape by going through the same sequence of grids to get access; pen up events are also considered in the graphical password scheme.

Recognising faces has also been used as a graphical authentication approach; PassFace Cooperation, for example, introduced a scheme that would allow the user to get access by selecting a set of faces among different panels (Corporation, 2009); the mechanism employes the human ability to remember faces as a way of authentication. Other graphical based approaches do exist, such as Wiedenbeck et al. (2005), Dhamija and Perrig (2000) and Sobrado and Birget (2002).

However, we do not consider our mechanism to fall under the two knowledge-based categories (conventional or graphical passwords) as neither it uses memorable alphanumeric characters nor it requires graphics, instead it uses GeoGraphical information.

3 GeoGraphical passwords

Humans – in general – do not prefer to memorise characters and if they had to, they do it in the least possible effort (Bensinger, 1998). This human behaviour – in the context of conventional passwords – leads to different vulnerabilities, including:

- using passwords that are vulnerable to dictionary attacks
- using passwords that are short enough to be vulnerable to brute-force attacks
- using the same password for different accounts
- Constructing a password using obvious information, such as birthdays or addresses, making the password easy to guess
- avoid changing the password according to a recommended time interval
- in the event of changing a password, the new password selected by the user is usually not very different from the previous one.

These vulnerabilities have been a main reason to many accounts compromises. To address these vulnerabilities we propose the concept of *GeoGraphical passwords*. We define a GeoGraphical password as:

> "A GeoGraphical password is a password that has been constructed based on GeoGraphical information."

We mean by geographical information the "knowledge acquired through processing geographically referenced data"; that is, data identified according to places on the Earth's surface (Li, 2007).

GeoGraphical information (i.e., lands, rivers, volcanos, mountains) are very familiar to humans, whom have a remarkable ability to remember places they have visited, or wish to visit (Teng and Squire, 1999). The geographical password recognises this characteristic in the human and utilise it for access credentials.

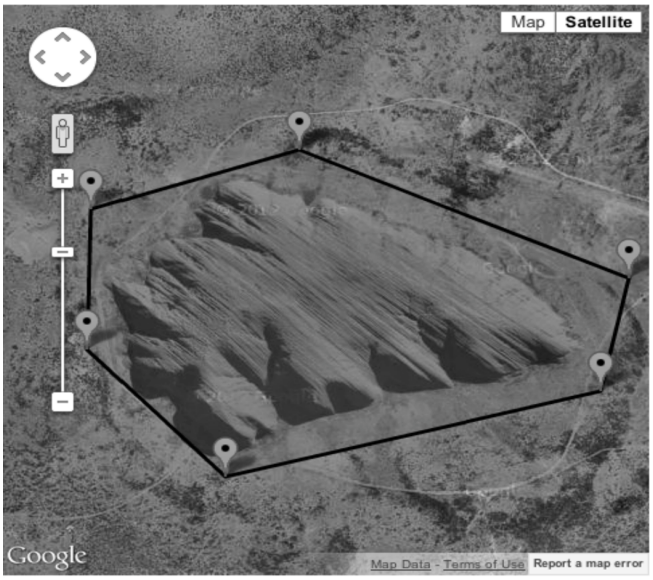
If users were able to select geographical locations as their access credentials then many vulnerabilities of the existing password-based authentication systems can be addressed. That is because geographical locations are:

- easy to remember and hard to forget; especially if there were feelings and memories associated with the selected places
- diverse; there are many geographical locations where the user can select from
- hard to predict; as users choose places based on their preference and experiences.

These elements add strength to the access credential and makes it harder for adversaries to compromise.

Selecting a geographical area can be done using different ways and shapes, a user – for example – can place a circle around his favourite mountain, or a polygon around his favourite set of trees, see Figure 1 for an example. No matter how geographical areas are selected, the geographical information that can be driven from these areas (such as longitude, latitude, altitude, areas, perimeters, sides, angels, radius, or others) form the geographical password.

Figure 1 A user selecting a geographical location (by drawing a polygon around a sandstone monolith in Australia) as her geographical password



4 GeoGraphical based access credential

We have developed a novel geographical based access credential to demonstrate one possible implementation of GeoGraphical passwords. In our prototype we divide the planet earth into small rectangular geographical areas – see Figure 2(1) – where each rectangle represent a GeoGraphical password – see Figure 2(2). For better user experience and ease of use, we divide earth into different layers where each layer represent a zoom level which has a different rectangular geographical area size.

Let ϕ_{sw} be the longitude coordinate at the south-west angle of the rectangular geographical area and ϕ_{se} be the longitude coordinate at the south-east angle. Let the difference between the two previous coordinates be:

$$\Delta\phi_z = |\phi_{sw} - \phi_{se}|, \text{where } z \text{ is the zoom level.} \tag{1}$$

Let λ_{sw} be the latitude coordinate at the south-west angle of the rectangular geographical area and λ_{nw} be the latitude coordinate at the north-west angle. Let the difference between the two previous coordinates be:

$$\Delta\lambda_z = |\lambda_{sw} - \lambda_{nw}|, \text{where } z \text{ is the zoom level.} \tag{2}$$

So if we assume the point at the south-west angle of the spherical rectangle is $(\phi_{sw}, \lambda_{sw})$ then the point at the north-east angle will be $(\phi_{sw} + \Delta\phi_z, \lambda_{sw} + \Delta\lambda_z)$. Therefore the larger $\Delta\phi_z$ and $\Delta\lambda_z$ are the larger the area the user can select as her geographical password (represented as a spherical rectangle in this mechanism).

We only need to know the south-west and the north-east points to identify the spherical rectangle P; for the sake of our application we will choose those two points as the geographical information that form our geographical password, therefore:

$$P_x = \{(\phi_{sw}, \lambda_{sw}), (\phi_{sw} + \Delta\phi_z, \lambda_{sw} + \Delta\lambda_z)\}. \tag{3}$$

Let P_x denote the rectangular geographical area selected in x order. So P_2, for example, is the second rectangular geographical area selected by the user as part of her geographical password. In our mechanism, the order in which the user selects her geographical locations is considered; therefore, let $GeoGP_q$ denote a geographical password, where q is the sequence number in which the $GeoGP$ has been selected; if $GeoGP_1 = \{P_1, P_2, P_3\}$ and $GeoGP_2 = \{P_2, P_1, P_3\}$, then $GeoGP_1 \neq GeoGP_2$. And since the user can not select the geographical location twice, the mechanism does not allow repetition.

Let r_z be the number of geographical locations selected at zoom level z and let j be the number of zoom levels available in the mechanism. Let R be the total number of selected geographical locations that forms the geographical password, therefore:

$$R = r_0 + r_1 + \cdots + r_j; \quad r_j \geq 0. \tag{4}$$

Let n_z be the number of geographical locations the user can select from at zoom level z; therefore, the total size of the geographical password space is

$$N = n_0 + n_1 + \cdots + n_j; \quad n_j \geq 0. \tag{5}$$

Using

$${}_nP_r = \frac{n!}{(n-r)!}; \quad \text{for } r \leq n, n \geq 0, \text{ and } r \geq= 0 \tag{6}$$

then based on equations (4)–(6), the number of possible ways (permutations) a user can select a geographical password, can be described by:

$$Q = \frac{n_0!}{(n_0 - r_0)!} + \cdots + \frac{n_j!}{(n_j - r_j)!} = \frac{N!}{(N-R)!}. \tag{7}$$

Therefore, as the geographical locations available (N) and the selected geographical locations (R) – as part of the GeoGP – increase, Q would increase as well. Which makes it more difficult for adversaries to guess the GeoGP.

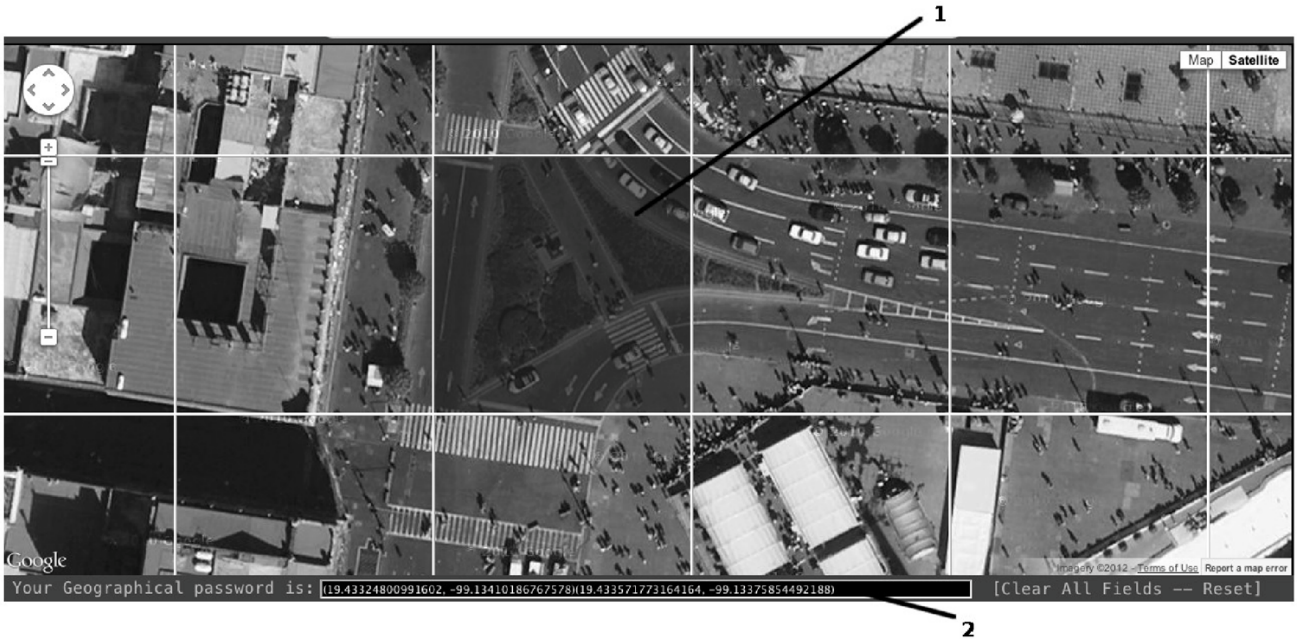
5 Entropy

The harder the adversaries are able to guess the GeoGP the stronger is the mechanism. Since the mechanism is novel, it is difficult to determine the frequency distribution of GeoGPs. However, it is not easy to guess the GeoGP, because due to the nature of the access credential it is more associated with the user's experience and feelings associated with the geographical location which is unique for each user. That is not the case when dealing with text based passwords, which usually adhere to computational linguistics techniques. There are different forms of the term *entropy*, however, in the context of the geographical based access credentials, we define entropy as "an estimate of the average amount of work required to guess" the GeoGP (NIST, 2006).

Since the mechanism (see Section 4) does not allow selecting the same geographical location twice and the order in which the location is selected is considered, then the entropy in bits can be described by the following formula:

$$E = \log_2({}_NP_R). \tag{8}$$

Figure 2 User selecting a geographical location, by selecting a rectangle containing a junction within Mexico City in Mexico (1), to form her geographical password (2)



We choose in our implementation to hash the GeoGP selected by the user to hide the actual rectangular geographical location, see Figure 3, for an example.

$$\text{HASH}(GeoGP_q) = H^q. \tag{9}$$

However, this would not increase the entropy; to increase the password space we can use a keyed-hash message authentication code (HMAC) using a memorable string of characters (i.e., word or a phrase) as a key for each user to hide the selected rectangular geographical location (The keyed-hash message authentication code (HMAC), 2008).

$$\begin{aligned}\text{HMAC}(K_u, geogp_q) = \text{HASH}((K_u \oplus opad)\| \\ \text{HASH}((K_u \oplus ipad)\|geogp_q)) = H_u^q\end{aligned} \tag{10}$$

where K_u is the key for the user u and H_u^q is the keyed hash value of user's u $GeoGP_q$. So the user can type a word or a phrase as her secret key before forming her GeoGP; see Figure 4 for an example. This will help avoid precompiled hashes attacks, such as Rainbow tables (Hellman, 1980). However, because users usually tend to choose short and easy to remember words as their keys and avoid complicated alphanumeric case sensitive keys, the entropy is reduced; we assume 2.5 bits as entropy for each character of the key (IEEE 802.11i-2004, 2004). Therefore after adding the key to the mechanism, the entropy becomes

$$E = \log_2({}_N P_R) + (l \times 2.5) \tag{11}$$

where l is the length of the key. However, allowing the user to pick her own key will make the key vulnerable to redundancy, which might lead to more than one user using the same password hash.

Therefore, sacrificing a bit of flexibility for more strength, by using a unique random key for each user to hide the selected rectangular geographical location, would increase the entropy of the mechanism and make each hash distinctive. The entropy after adding a randomly generated key, can be described by

$$E = \log_2({}_N P_R \times b^l) \tag{12}$$

where the key is generated from an alphabet of b characters, see Figure 5 for an example. Table 1 gives an overview of the guessing entropy of the proposed geographical password prototype.

Table 1 Guessing entropy in bits

	$N = 360 \times 10^9$		
Secret key type	$R = 1$	$R = 2$	$R = 3$
No key	38.39	76.77	115.167
Memorable string of characters ($l = 8$)	58.38	96.77	135.167
Memorable string of characters ($l = 16$)	78.38	116.778	155.167
Randomly generated key (128 bits, $l = 32$, $b = 16$)	166.38	204.77	243.167
Randomly generated key (256 bits, $l = 64$, $b = 16$)	294.389	332.77	371.1

6 Discussion

Stopping the ever increasing password breaches has pushed the research community to look for better password solutions to further protect the user; and the proposal of GeoGraphical passwords comes within this prospective. GeoGraphical passwords should not be confused with graphical passwords, as it does not require graphics, for example another prototype implementation of GeoGPs can be a high-tech glove that can

Figure 3 User selecting a geographical location, by selecting a rectangle containing a small pyramid in Egypt (6) as her geographical password and transforming it into a hash value (7)

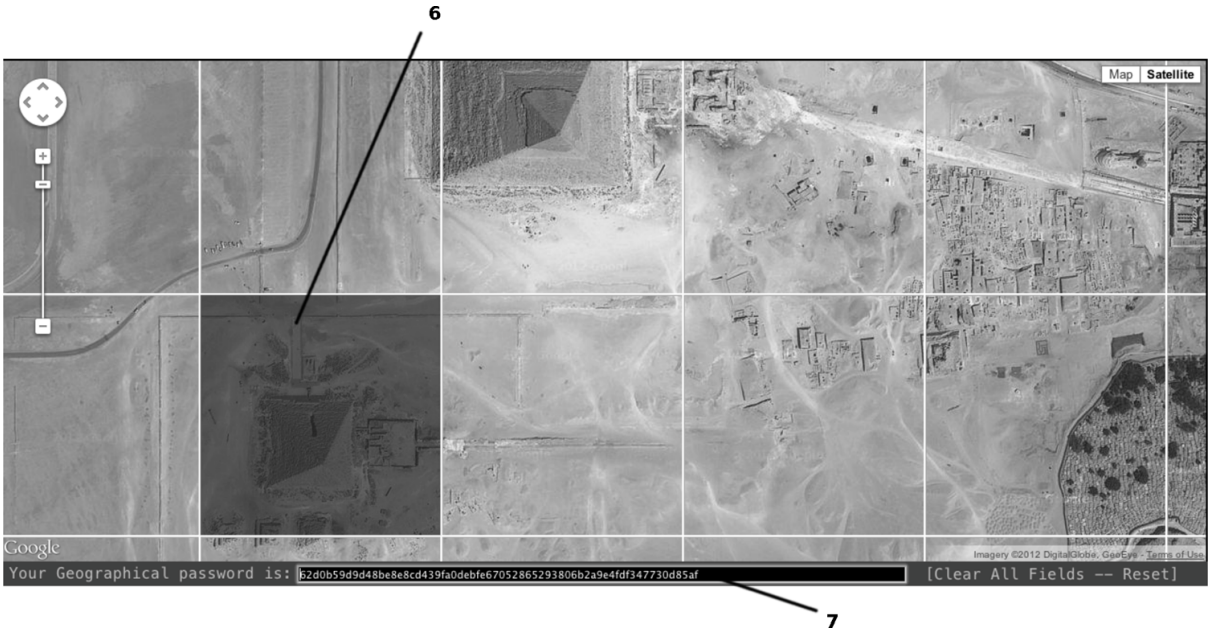
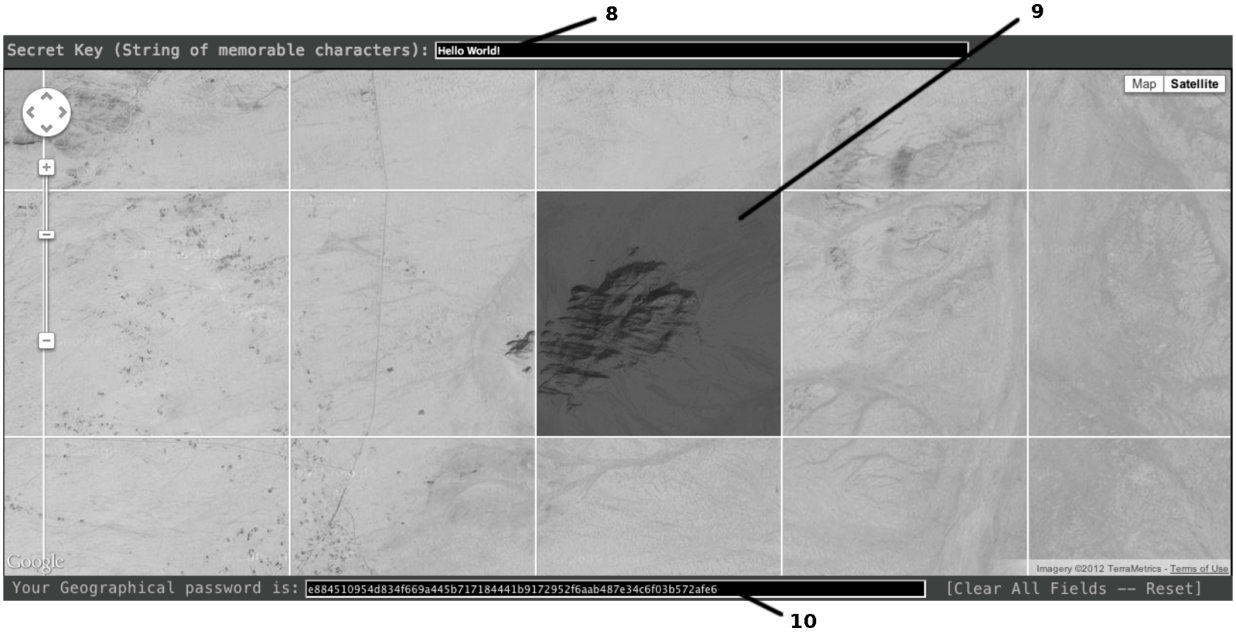


Figure 4 User selecting a geographical location, by selecting a rectangle containing a Giant Plateau in the Arabian Peninsula (9) as her geographical password and transforming it into a keyed hash value (10), where the secret key is a memorable string of characters (8)



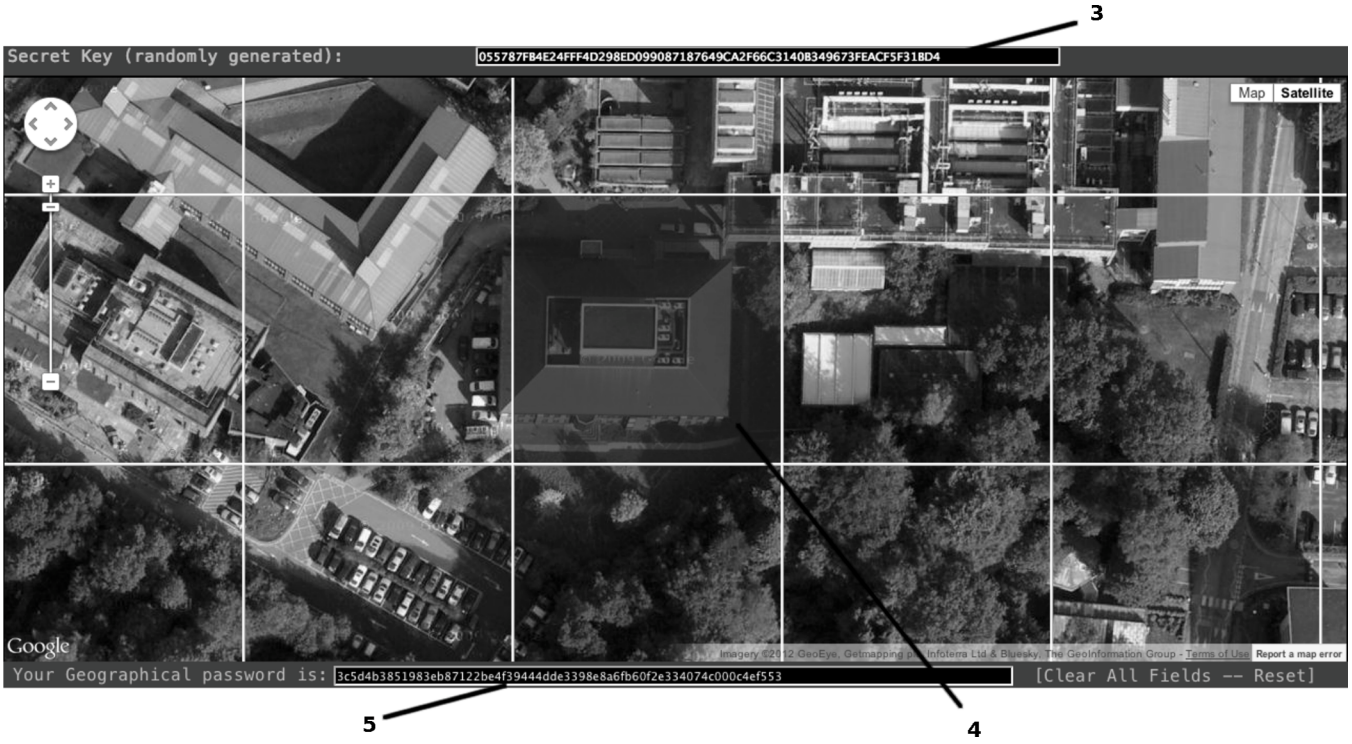
extract geographical information from objects it touches and provide access based on the extracted information.

Let us assume that we have 360 billion tiles ($N = 360 \times 10^9$) that covers planet Earth in 20 zoom levels (Miller, 2010). In our GeoGP implementation described in this paper, not using a key has resulted in a guessing entropy of 38.39 bits if only one geographical location has been selected. The strength of the password would noticeably increase when the user selects two or more geographical locations as part of her GeoGP; for example 115.167 bits of guessing entropy were the result of a GeoGP that consists of three places.

Entering a secret memorable string of 8 characters along with one, two, or three geolocations has resulted in 58.38, 96.77, and 135.167 bits respectively. Increasing the length of the string to 16, has improved the strength of the GeoGP to be 78.38 bits for one geolocation selected and 155.167 bits for three geolocations.

The strength of the GeoGP highly improves when each user use a unique randomly generated key as part of her GeoGP. A 128 bit hexadecimal key, for example, resulted in 166.38 bits of guessing entropy for one selected geolocation, while 243.167 bits resulted from selecting three geographical

Figure 5 User selecting a geographical location, by selecting a rectangle containing Royal Holloway, Univ. of London library building in UK (4) as her geographical password and transforming it into a keyed hash value (5), where the secret key is 128-bit of length and is randomly generated (3)



locations. Increasing the length of the key to 256 bits resulted in 294.389 bits of guessing entropy for one geolocation selected and 371.1 bits for three. For all results see Table 1.

As it appears, the more geolocations selected – by the user – the stronger the GeoGP becomes; also adding a key either memorable string or randomly generated, further increase the strength of the access credential.

Furthermore, changing the GeoGP to a very different one, can be easily achieved, just by selecting another geographical location. And due to the nature of the password, users can easily choose different GeoGPs for their multiple accounts reducing the suffer from password fatigue, and eliminating many vulnerabilities associated with conventional passwords.

7 Conclusion

Indeed passwords compromises have increased, this led the research community to search for conventional passwords alternatives. In this paper we tackled the password problem by proposing a novel access credential based on geographical information. The credential employees the distinctive human ability to remember places to eliminate many vulnerabilities associated with current password based authentication schemes. The high guessing entropy of the credential makes it very difficult for adversaries to compromise. The geographical human friendly password would change how people deal with their access credentials; just imagine your geographical password to your e-mail or social network is your summer home or the lake you have visited few years ago. Geographical passwords can address the increasing vulnerabilities associated with conventional ones and would further improve online security, paving the way towards a better user protection in an unpredictable cyber world.

References

BBC News Technology (2012) *Linkedin Passwords Leaked by Hackers*, [Online Accessed 12 April, 2013], http://www.bbc.co.uk/news/technology-18338956

Bensinger, D. (1998) *Human Memory and the Graphical Password*, Passlogix, Inc.

Blonder, G.E. (1996) *Graphical Password*, US Patent 5559961 No. 5559961.

Bonneau, J. (2012) 'The science of guessing: analyzing an anonymized corpus of 70 million passwords', *2012 IEEE Symposium on Security and Privacy*, http://www.cl.cam.ac.uk/jcb82/doc/B12-IEEESP-analyzing_70M_anonymized_passwords.pdf

Corporation, P. (2009) *The Science Behind Passfaces*, White Paper, http://www.passfaces.com/published/The%20Science%20Behind%20Passfaces.pdf

Devillers, M. (2010) *Analyzing Password Strength*, Tech. Rep., Radboud University Nijmegen.

Dhamija, R. and Perrig, A. (2000) 'Déjà vu: a user study using images for authentication', *Proceedings of the 9th Conference on USENIX Security Symposium – Volume 9, SSYM'00*, USENIX Association, Berkeley, CA, USA, p.4, http://dl.acm.org/citation.cfm?id=1251306.1251310

Grosse, E. and Upadhyay, M. (2013) 'Authentication at scale', *IEEE Security and Privacy*, Vol. 11, pp.15–22, http://www.computer.org/cms/Computer.org/ComputingNow/pdfs/AuthenticationAtScale.pdf

Hellman, M. (1980) 'A cryptanalytic time-memory trade-off', *Information Theory, IEEE Transactions on*, Vol. 26, No. 4, pp.401–406.

Herley, C., Oorschot, P.C. and Patrick, A.S. (2009) 'Passwords: If we're so smart, why are we still using them?', in Dingledine, R. and Golle, P. (Eds.): *Proceedings of the 13th Conference on Financial Cryptography and Data Security*, Springer-Verlag, Berlin, Heidelberg, pp.230–237, http://dx.doi.org/10.1007/978-3-642-03549-4_14

IEEE 802.11i-2004 (2004) *IEEE 802.11i-2004: Amendment 6: Medium Access Control (MAC) Security Enhancements*.

Imperva Data Security Blog (2011) *Military Password Analysis* [Online Accessed 12 April, 2013], http://blog.imperva.com/2011/07/military-password-analysis.html

Jermyn, I., Mayer, A., Monrose, F., Reiter, M.K. and Rubin, A.D. (1999) 'The design and analysis of graphical passwords', *Proceedings of the 8th Conference on USENIX Security Symposium – Volume 8, SSYM'99*, USENIX Association, Berkeley, CA, USA, p.1. http://dl.acm.org/citation.cfm?id=1251421.1251422

Kaspersky Labs (2012) *Full Analysis of Flame's Command & Control Servers* [Online Accessed 12 April, 2012], https://www.securelist.com/en/blog/750/Full_Analysis_of_Flame_s_Command_Control_servers

Li, B. (2007) 'Geographic information services', in Shekhar, S. and Xiong, H. (Eds.): *Encyclopedia of GIS*, Springer, USA.

Lord, B. (2013) *Keeping our Users Secure*, Official Twitter Blog [Online Accessed 12 April, 2013], https://blog.twitter.com/2013/keeping-our-users-secure

Miller, A. (2010) *Under the Hood of Google Maps 5.0 for Android*, Google Official Blog [Online Accessed 12-April-2013], http://googleblog.blogspot.com/2010/12/under-hood-of-google-maps-50-for.html

NIST (2006) *Electronic Authentication Guideline*, Technical Report, http://csrc.nist.gov/publications/nistpubs/800-63/SP800-63V1_0_2.pdf

PC World (2013) *Evernote Hack Shows that Passwords aren't Good Enough* [Online Accessed 12 April, 2013], http://www.pcworld.com/article/2030052/evernote-hack-shows-that-passwords-arent-good-enough.html

Sobrado, L. and Birget, J-C. (2002) *Graphical Passwords*, The Rutgers Scholar, http://rutgersscholar.rutgers.edu/volume04/sobrbirg/sobrbirg.htm

Teng, E. and Squire, L. (1999) 'Memory for places learned long ago is intact after hippocampal damage', *Nature*, Vol. 400, No. 6745, pp.675–677.

The Data Breach Investigations Report (2013) Technical Report, Verizon.

The keyed-hash message authentication code (HMAC) (2008) Technical Report 198-1, NIST.

Wiedenbeck, S., Waters, J., Birget, J-C., Brodskiy, A. and Memon, N. (2005) 'Authentication using graphical passwords: effects of tolerance and image choice', *Proceedings of the 2005 Symposium on Usable Privacy and Security, SOUPS '05*, ACM, New York, NY, USA, pp.1–12, http://doi.acm.org/10.1145/1073001.1073002