User Experience Evaluation in the Creation and Use of Graphical Passwords for Authentication in Mobile Devices

Claudia de Andrade Tambascia, Ewerton Martins Menezes, Alexandre Melo Braga, Flávia de Melo Negrão
CPqD Foundation
Campinas – SP, Brazil
{claudiat, emenezes, ambraga, fnegrao}@cpqd.com.br

***Abstract* — This article aims at present the results of a user experience evaluation for the creation and use of graphical passwords on mobile devices as a way to improve usability and security aspects in authentication. The authentication method proposed was defined in a project called Multimodal Biometric and Graphical Authentication for Mobile Devices. These assessments were carried out with thirty users during a period of fifteen days using a prototype that offered a repertory of eighty icons divided into four categories. All users were able to remember their eight-icon password and claimed having a good use experience with this authentication method.**

Keywords - User experience; graphical passwords; mobile authentication.

I. INTRODUCTION

One of the main goals of the use of graphical passwords in authentication processes is to increase the usability of the interaction, facilitating the memorization of passwords for users, ensuring a greater retention of information. This feature relies on the results of various studies and experiments [1][2][3][4] that show the human brain finds it easier to recognize and remember visual information comparing to textual information.

As a recognition-based technique, the graphical authentication with icons demands less cognitive load than recall techniques and tends to increase the usability, the security and the user performance, besides being especially appealing in the mobile context, where pointing to a region of the screen tends to be much easier than typing text.

While some recognition-based systems use faces [5], assuming that the brain has got a special ability to recognize them, other systems use abstract images [6], which are stronger from a security point of view, due to their difficulty of describing. Nevertheless, the use of icons brings a better compromise between usability and security, once it facilitates mnemonic strategies and, consequently, memorization.

The security level offered by such systems depends on many factors, such as the size of the repository available to the user, the password length, the input method, and the icons themselves, which must, ideally, present similar probabilities of choice avoiding possible attacks.

This paper will describe the results obtained in the evaluation of the quality of user experience in the creation and use of graphical passwords for authentication on mobile devices. For this evaluation, a prototype was developed to allow the experience in a real context of use, considering aspects of usability, intelligibility and memorization strategies. This prototype is part of Multimodal Biometric and Graphical Authentication for Mobile Devices (BIOMODAL) project whose main goal is to develop functional prototypes of biometric multimodal authentication and graphical authentication for mobile communication devices.

Section two will present a description of the prototype that was developed for this evaluation, followed by section three that will present the methodology used for the user experience evaluation. Section four will present the main results observed during the exploration, the creation and the strategies of memorization applied followed by section five with the analysis of use data for the different groups of users and interviews. The section six will present the main conclusions and findings related to graphical authentication.

II. RELATED WORKS

As a knowledge-based authentication technique, the graphical authentication requires the user to enter a shared secret as an evidence of their identity. This authentication scheme has been proposed as an alternative to text-based password for over one decade [7]. Surveys on the field [8][9] review some graphical password systems from the usability and security perspective, but Robert Biddle's survey [7] provides us with a comprehensive review of the first twelve years of published research on graphical password systems. According to this study, graphical passwords scheme can be classified in three main categories: based on recall, recognition and cued-recall.

Graphical passwords with icons are a type of recognition-based systems where users are asked to memorize a repertory of images during password creation, and then recognize their images from among decoys to authenticate. Proposed recognition-based systems use various types of images, most notably: faces, random art, everyday objects, and icons. Renaud [10] discusses some specific security and usability considerations, and offers usability design guidelines for recognition-based systems.

In the literature it's possible to find studies such as Antonella [11] that compares the a PIN password against three different graphical passwords schemes in two user studies with 60 participants and verified that graphical

 ISBN: 978-1-61208-229-5

password can be a solution to some problems related to knowledge-based authentication, but poor design can eliminate pictures superiority effect in memory.

Darren study [12] evaluated some security and usability aspects involved in graphical passwords with icons using faces and showed the importance of some rules in the password selection in graphical schemes due to the highly correlation between race and gender of the user.

And due to the specially appealing of graphical password in the mobile context, where typewritten input is less common than pointing at the screen [13] we found studies like the Dunphy's one [14] that evaluate two versions of a recognition-based system on mobile devices where the images used in the application were provided by the users himself.

Although there are many academicals studies about graphical passwords with icons only a few commercial products are available in the market such as Passfaces and the over take-up is low [14] [15].

III. PROTOTYPE FOR CREATION OF GRAPHICAL PASSWORDS

The evaluation of user experience with graphical passwords was performed by the means of a prototype developed and installed in a Samsung Galaxy S smartphone, with 4 inches (10 centimeters), resolutions of 480×800 pixels and Android operation system version 2.2, as shown in Figure 1.

Figure 1 - Prototype installed in a smartphone

The prototype enabled the evaluation to be conducted in real devices, allowing evaluating not only graphical authentication parameters but also identifying interaction requirements inherent to mobile platform. Users were able to create and use graphical passwords in the device that automatically registered: time of use, number of page scrolls, use of "clear" function and number of authentication attempts, which are parameters needed to measure the quality of user experience.

For the creation of the graphical passwords, a repertory was produced based on the results presented on [16] that adopted mnemonic strategies to favor the use of episodic memory in the moment of password creation. The repertory was composed of eighty icons, displayed in 4 screens of twenty icons each, classified in categories: people, objects, means of transportation and context/place, as shown in Figure 2.

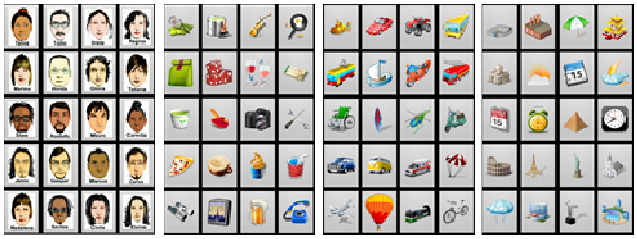

Figure 2 - First repertory of icons

For the usability study it was designed a simple prototype, which not bound by strong password policies. This design decision allowed the study of users unsafe behaviors. A commercial authentication solution based on this technology not only must implement stronger passwords (longer), but must provide security policies that inhibit the passwords considered that would facilitate easy and dictionary attacks. Only for an example, a graphical password cluttered with 9 icons in a grid of 80 icons is equivalent of a password of 8 letters, and a password of nine icons arranged in a grid of 57 icons is equivalent of 8 size alphanumeric password with uppercase letters, lowercase letters, numbers and special characters.

As well as in the use of alphanumeric passwords, it was possible to observe that some of users tried to eliminate apparent or implied variables of the repertories to facilitate their process of memorization. Thus, to enhance the quality of the passwords, besides not allowing choosing repeated icons, additional restrictions were imposed in the moment of creation: not choosing more than three icons in each screen to avoid that all icons are in the same page or belong to the same category.

From the application interface point of view, this evaluation sought to identify and validate the adequacy of the quantity of icons presented per screen, easy and intuitive navigation, and clarity of the actions and feedbacks offered to the users.

IV. PROTOTYPE EVALUATION METHODOLOGY

In order to deeply explore the process of graphical password creation, the user sample was divided into three groups with distinct methods of creation. The first group was able to create the password according to their own personal preferences; the second group was suggested a random theme (adventure, romance, work, celebration, contretemps, tourism, luck and trouble) and the third group received a password that was randomly generated by the application with two icons in each screen.

Training is a factor of great influence in the memorization and performance of any password. In order to understand of the impact of this variable in graphical passwords, half of the participants went through a training session where they used the password three consecutive times right after creating it.

A total of thirty users representing general population participated in the test divided into six groups taking into

 ISBN: 978-1-61208-229-5

account the interaction with or without password training, user created password with or without a suggested theme and the memorization of randomly generated passwords by the application and given to the users right before the first interaction, as shown in Figure 3.

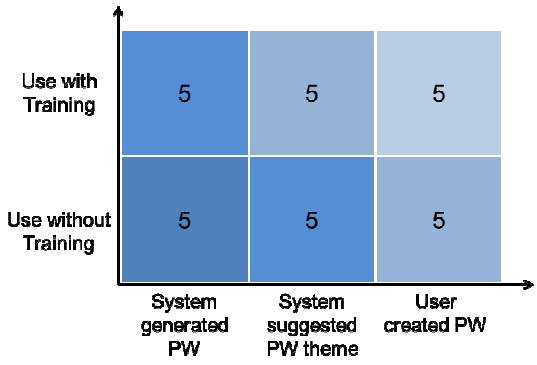

Figure 3 - User distribution according to established groups

The evaluation consisted of six stages, where some stages were common to all groups and others, however, were specific to certain groups. The stages considered were:

- Exploration: in this stage users were requested to visually explore the repertory of icons displayed in the eye tracking device. No interaction with the mouse or keyboard was required since the screens were presented for a preset time.
- Creation: the creation of the password of eight icons was performed in the mobile device instantly after the repertory exploration. A total of twenty users underwent the process of password creation, and ten of them were suggested a theme to help create the password.
- Password generation: the password was generated in an application developed to randomly choose two icons from the repertory in each screen, composing an eight icon password presented to the user right after exploration.
- Training: consisted of entering the password three consecutive times in the mobile device. The user did not need to enter the correct password every time, but after each interaction the system informed if the password was correct or not.
- Password effective use: the graphical password was used four times during the evaluation, in increasing time intervals from the creation until the last use.
- Interview: it was conducted at the end of the evaluation, where some questions were asked to the participants to verify their perception regarding the interaction with graphical passwords. The following aspects were treated: experience, difficulty of use, positive and negative issues, tendency to replace alphanumeric passwords for graphical, passwords created and memorization strategies, category harder to recall, category most liked, experience with smartphones and the mobile device model owned by the participant user.

To evaluate the performance between different groups of users, quantitative and qualitative parameters were collected. The quantitative parameters considered the time of creation and uses of the passwords, the success and failure rates, the number of attempts in each authentication, the use of “clear” function, navigation between screens, icons chosen and icons looked at. The qualitative parameters, in turn, considered the memorization strategies, the evolution of use, the adherence of graphical passwords to the suggested theme and the level of user satisfaction.

Users were divided into gender: male and female; and age: below thirty years old, between thirty and forty-five years old and above forty-five years old, as shown in Figure 4. Such division was necessary once the memorization factor was to be measured, which implied the need of considering different age groups.

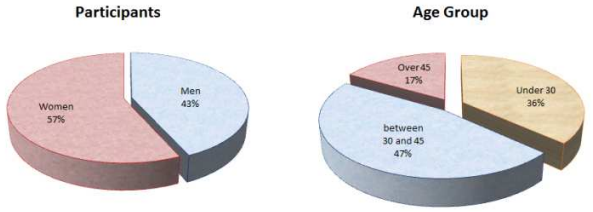

Figure 4 - Test users’ selection

Since one of the main criteria to be evaluated would be the capacity of memorization of passwords after a period of time, it was defined that the test would happen during an interval of fifteen days according to the schedule presented in Figure 5.

Figure 5 – Test schedule

V. RESULTS TABULATION

A. *Users behavior during icon exploration*

Initially, before the creation of passwords and beginning of the evaluation, the users were submitted to a process of exploring the icons from the first repertory in order to become acquainted with them, before the creation of the password itself. This process was performed for one minute, being twenty-five seconds for each category: people, objects, means of transportation and context/place.

With the support of the eye tracking tool it was possible to assess the icons that drew more attention during the exploratory process and the ones that were practically not observed, as shown in Figure 6 for the category “people”.

The Heat Map, a feature of eye tracking tool, enabled the visualization, throughout the variation of shades from green to red, of which icons were more observed, where the user fixed the attention and the icons that were practically not seen. The graph immediately after the Heat Map presents how many users choose each icon of the category. The red

"X" sign in the icons represents the ones that were not selected by the users when creating the password.

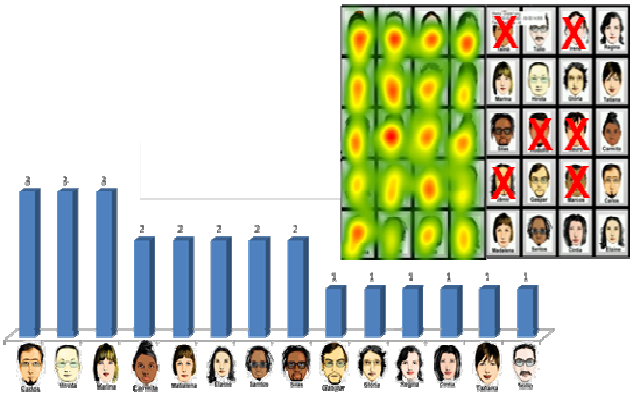

Figure 6 - Users behavior during the exploration of category "people"

B. *Users behavior during the creation of passwords*

The users selected to create passwords freely presented the behavior mapped in Figure 7.

Figure 7 - Users behavior in the creation of free passwords

According to the Heat Map mapping, it was possible to assert that the most observed icons where indeed chosen for the users passwords. Some icons had a high degree of choice in relation to others due to the easy connection to everyday situation and to cultural factors.

The users that had to choose the icons of the password according to a suggested theme also chose them freely and were not obliged to follow the theme. Since the test had the premise to create passwords that were safer and easier to memorize, the themes were intended to work as support in the creation of a plot for the composition of the password and ease of memorization. Figure 8 presents the passwords created with an associated theme and the adherence of each one of them to what was suggested.

From the ten users that participated in this category only four sought to use password adherent to the suggested themes.

Figure 8 - Users behavior in the creation of free passwords with suggested theme

C. *Memorization strategies*

Using of a repertory of icons separated into four categories, yet allowing for the formation of an episode with subject, actions and context, provided a great variety of memorization strategies and few participants declared the need to write down the password to help retain the icon set and at the end of the fourth use all participants were able to remember their graphical password in less than three attempts. These strategies were obtained by interviewing the users of the test after executing all interactions planned.

For all passwords used in this test, only eight users created a plot for the memorization process, six used the suggested theme and four affirmed not using any strategy.

The users that received the system generated passwords sought to look only at the icons given and not to the other icons in the screen to avoid possible confusions. For these users, the creation of a plot demanded a quick creativity which was not always possible.

Fig. 9 presents some passwords created by the users and the description of the strategies used, according to what was reported during the interviews.

Regarding password (I), the strategy used by the user was selecting icons of sports and food that he liked best. He chose the icon "balloon" because he wished to travel in one, the icon "airplane" because it was related to holidays and places he would like to visit.

Regarding password (II), the user admitted not using any strategy, only the visual memorization and chose the icon "calendar" because the date displayed was her husband's birthday. However, this user did not have a satisfactory performance during the interactions and made mistakes in different attempts.

 ISBN: 978-1-61208-229-5

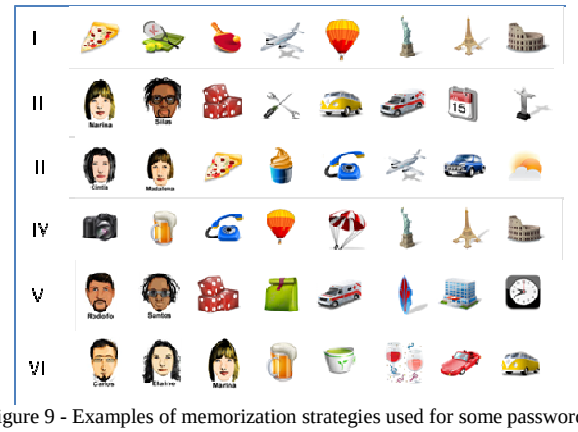

Figure 9 - Examples of memorization strategies used for some passwords created by the user

Regarding password (III), the user claimed to search for icons related to the suggested theme (romance), choosing outstanding pictures and remarkable or familiar names.

Regarding password (IV), the user chose objects he identified himself with and places he would like to visit.

Regarding password (V), the user also did not use any strategy, and since the password was provided by the system, he considered himself lucky by the presence of the icons "ambulance" and "hospital", in addition to two faces he could memorize by the names.

Regarding password (VI), the user claimed to have selected people with familiar names and icons related to two categories only (means of transportation and objects), thus avoiding confusion.

VI. RESULTS ANALYSIS

The results obtained in the data analysis, considering the performance during system interactions for all participants, were compared according to the group they belonged to, as follows:

- Participants with "User-Created passwords with No Training (UCNT)";
- Participants with "System-Generated passwords with No Training (SGNT)";
- Participants with "System-Suggested-Theme passwords with No Training (SSTNT)";
- Participants with "User-Created passwords With Training (UCWT)";
- Participants with "System-Generated passwords With Training (SGWT)";
- Participants with "System-Suggested-Theme passwords With Training (SSTWT)".

Figure 10 shows the average time of interaction after the first, fifth and ninth days from the creation date of the graphical passwords.

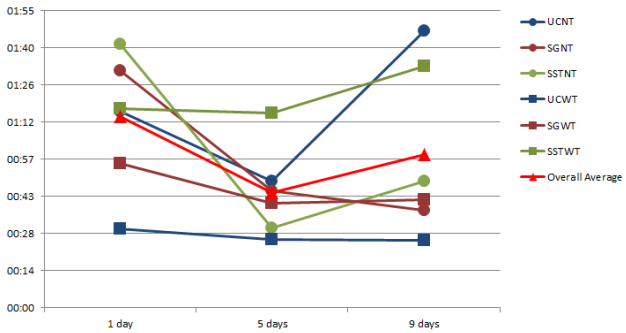

Figure 10 - Average time by user groups

As noted, the best time was achieved by the participants that created their own passwords and were trained (UCWT), while the worst time was achieved by the users that received a theme and were able to practice the password (SSTWT). It is possible to declare by the analysis of the graph in Figure 10 that four of six groups hold their performance close to the average, except for the last interaction for groups UCNT and TSCT, which is significantly deviated from the curve. It is also possible to observe that the groups that were not trained performed worse in the first day of use, but after the fifth day the performance returned to normal.

Regarding the creation method and training, a more detailed analysis done, to verify the behavior of users during the interactions. Figure 11 presents the average time of interaction of the participants, comparing passwords that were created by the users and the passwords generated by the system.

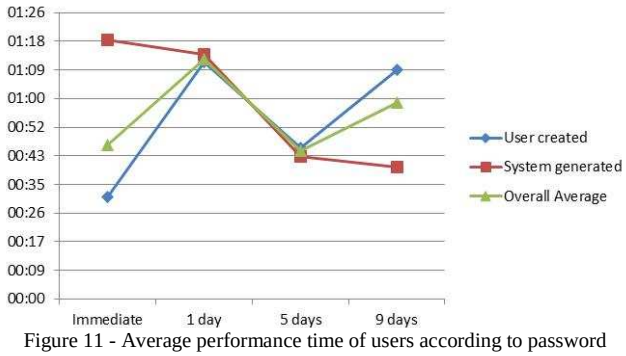

Figure 11 - Average performance time of users according to password creation method

The immediate use of system generated passwords had greater interaction time, as expected because the password was not created by the user and by the need to instantly having to create a mental model for its memorization. The same behavior was observed on the first day of use, after the password creation, where the time periods were considered long, and probably motivated by the user's fear of making mistakes in password selection. The behavior on the following days was considered better, curiously obtaining better performance on the last day for passwords that were not created by the user himself.

Figure 12 presents the behavior of the participants by previous training to the use of the created password. The

 ISBN: 978-1-61208-229-5

variable related to training was considered for this test as means to prove a possible outstanding performance regarding users that were not able to be trained.

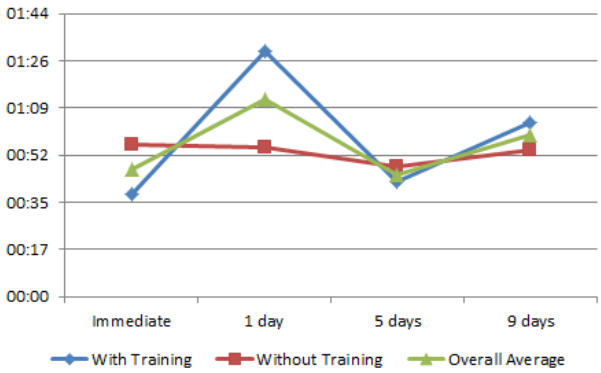

Figure 12 - Average performance of users with and without training

It was possible to observe in this study that only in the first day of interaction there was a significant difference between the participants with training and without training and, in an unexpected way, the users that were not trained had a better performance than the ones that were trained. This situation happened because many users did not use good memorizing strategies and didn't invested much time in creating the password, which means a lack of care during the creation process, making the memorization process difficult. Only after the first use that the memorizing strategy was created.

However, in the following days, the average for the participants with training and without training remained the same, which allows concluding that memorization is not directly related to the fact of password being practiced or not, but rather to the memorization strategy associated.

Regarding the experience in using graphical passwords, most of the users showed preference for this type of authentication, probably motivated by the innovation and ludic aspects of this approach. More than eight percent of the users considered the experience satisfactory and the remaining users since were not able to memorize the password in all interactions, eventually did not consider the experience as satisfactory.

It was possible to notice that in the last days, the degree of dissatisfaction increased in a significant manner, due to the interval between the previous interaction and the moment of password creation. Many memorization strategies here prove to be inefficient.

The same analysis was performed for the item related to the difficulty of use, where it was possible to observe a greater difficulty of the participants in the last day of interaction, linked to the time elapsed between the creation and penultimate day of use of the password; and also regarding to inefficient memorization strategies.

Since the memorization strategies did not work properly, the difficulty and quality of experience were compromised, not by the application and the solution itself, but often by the participant's own frustration for not accomplishing the task.

Regarding the category that most pleased the users there was a balanced result in categories "objects", "means of transportation" and "context/places" in the item most liked category.

The most significant result was related to category "people" with a larger number of critics mainly by the definition of the faces and confusion among several similar features. There were very few indifferent or totally satisfied users, not that significant to offset further analysis. Figure 13 shows the results obtained and the number of users that expressed their preference for any of the categories.

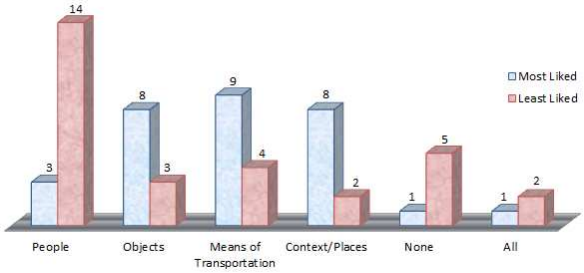

Figure 13 - Experience of use of graphical passwords

Regarding the difficulty of use, it was observed that less than ten percent of the users complained having problems using the prototype, being that this small percentage was composed by users that had no experience with smartphones, showing that the level of familiarity with technology created one of the main barriers in the usage of the proposed solution.

At last, regarding the tendency to replace alphanumeric passwords for graphical passwords, it was possible to observe a favorable result towards graphical passwords, but it was not so outstanding due to the fact of people being culturally adapted and proficient in the use alphanumeric passwords.

Even thou the fact of the experience with graphical passwords being interesting, from the tendency point of view, there is still some resistance, mainly because some considered this authentication method slower than alphanumeric passwords and identified a significant delay in the navigation between the icon screens.

VII. CONCLUSIONS

In the analyzes, it was possible to observe slightly better error rates and elimination of degenerated passwords through a more efficient repertory of icons with more graphically detailed images, greater possibility of creating stories and with fewer errors observed.

Regarding memorization, it was proved a reduction in the error rate and the users were surprised with their own performance, emphasizing that in the case of the category "people", it was easier to memorize the name of the person than the image itself, which many times was easily mistaken by the similarities of the faces presented.

The introduction of the process of training in the evaluation did not produced outstanding differences in the performance of the participants that had their passwords

 ISBN: 978-1-61208-229-5

generated by the system, but it affected in the creation of passwords by the users themselves, that obtained similar average input time and error rates.

Thus, the possibility of the creation of a script for memorization turns out not to be mandatory, since the less chosen icons where the ones with low contrast and difficult identification, while the ones most chosen had cultural appeal and high contrast.

From the usability requirements point of view, the decision to display icons in touchscreens was a good design solution however its performance needs to be improved. On the other hand, the feedbacks related to the quantity of icons selected and the navigation between screens ware not sufficiently clear and generated questions during the evaluation.

The size and quantity of icons available were well accepted by the users, requiring improvements in the process of cleaning the current selected icon which generated confusion and frustration in many users. The addition of an initial help mechanism is of extreme importance in a graphical authentication process, mainly by the paradigm change in authentication.

There was no correlation found between the most viewed icons in the exploration and the most chosen icons in the password composition. This suggests that the factors that made the icons more visually attractive and draw attention of the participants are not necessarily related to the criteria for choosing the icons to create stories.

Fewer users found great difficulties in the use of the prototype and the majority claimed to have a satisfactory experience of use during the tests. The adoption of this technology can be considered high, around sixty percent of the participants claimed they would replace their current alphanumeric passwords for graphical passwords, mainly for convenience of touchscreen and for the ludic experience involved.

The test conducted in the context of this project produced an analysis that is contributing for the definition of the graphical authentication solution to be used in the implementation of a functional prototype of BIOMODAL Project.

ACKNOWLEDGMENT

The authors acknowledge the financial support given to this work, under the project “Biometric Multimodal and Graphical Authentication for Mobile Devices – BIOMODAL”, granted by the Fund for Technological Development of Telecommunications – FUNTTEL – of the Brazilian Ministry of Communications, through Agreement Nr. 01.09.0627.00 with the Financier of Studies and Projects - FINEP / MCTI.

REFERENCES

[1] B. Kirkpatrick. “An experimental study of memory”. Psychological Review, 1894, 1:602-609.

[2] S. Madigan. “Picture memory”. In J. Yuille, editor, “Imagery, Memory, and Cognition: Essays in Honor of Allan Paivio”, cap.3, pp. 65-89. Lawrence Erlbaum Associates, 1983.

[3] A. Paivio, T. Rogers, and P.C. Smythe. “Why are pictures easier to recall than words?” Psychonomic Science, 1968, 11(4):137-138.

[4] R. Shepard. “Recognition memory for words, sentences, and pictures”. Journal of Verbal Learning and Verbal Behavior, 1967, 6: pp. 156-163.

[5] Passfaces Corporation. “The Science Behind Passfaces”. Write paper, http://www.realuser.com/enterprise/resources/ whitepapers. htm, accessed Sep, 2012.

[6] R. Dhamija and A. Perrig. “Déjà Vu: A user study using imagens for authentication”. In 9th USENIX Security Symposium, 2000. Proceeding of the SSYM'00 9th conference on USENIX Security Symposium - Volume 9, Pages 4 – 4. USENIX Association Berkeley, CA, USA.

[7] R. Biddle, S. Chiasson, and P.C. Van Oorschot. 2012. “Graphical passwords: Learning from the first twelve years”. ACM Comput. Surv. 44, 4, Article 19 (September 2012), 41 pages.

[8] F. Monrose and M. Reiter. 2005. “Graphical passwords. In Security and Usability: Designing Secure Systems That People Can Use”. L. Cranor and S. Garfinkel, Eds. O'Reilly Media, Stebastopol, CA, Chapter 9, 157-174.

[9] S. Xiaoyuan, Y. Zhu, and G. Scott. Owen. 2005. “Graphical Passwords: A Survey”. In Proceedings of the 21st Annual Computer Security Applications Conference (ACSAC '05). IEEE Computer Society, Washington, DC, USA, 463-472.

[10] K. V. Renaud. 2009. “Guidelines for designing graphical authentication mechanism interfaces”. Int. J. Inf. Compute Security 3, 1 (June 2009), 60-85.

[11] A. De Angeli, L. Coventry, G. Johnson, and K. Renaud. 2005. “Is a picture really worth a thousand words? Exploring the feasibility of graphical authentication systems”. Int. J. Hum.-Comput. Stud. 63, 1-2 (July 2005), 128-152.

[12] D. Davis, F. Monrose, and M. K. Reiter. 2004. “On user choice in graphical password schemes”. In Proceedings of the 13th conference on USENIX Security Symposium - Volume 13 (SSYM'04), Vol. 13. USENIX Association, Berkeley, CA, USA, 11-11.

[13] C. A. Tambascia, E. M. Menezes, R. E. Duarte. “Usability Evaluation Using Eye Tracking for Iconographic Authentication on Mobile Devices”. Mobility 2011, The First International Conference on Mobile Services, Resources, and Users. Barcelona, Spain (October 2011), 117-122.

[14] P. Dunphy, A. P. Heiner, and N. Asokan. 2010. “A closer look at recognition-based graphical passwords on mobile devices”. In Proceedings of the Sixth Symposium on Usable Privacy and Security (SOUPS '10). ACM, New York, NY, USA, , Article 3 , 12 pages.

[15] Passfaces Corporation. “The science behind Passfaces”. White paper, http://www.passfaces.com/enterprise/resources/ whitepapers.htm, accessed September 2012.

[16] I. Avila, E. Menezes, A. Braga. “Memorization Strategy for Iconic Passwords”. In: IADIS International Conference Interfaces and Human Computer Interaction 2012, Lisboa. Proceedings of the IADIS Intl. Conf. Interfaces and Human Computer Interaction 2012. Lisboa: IADIS, 2012. v. 1. p. 123-132. 1.

 ISBN: 978-1-61208-229-5