The Effectiveness of Intersection Attack Countermeasures for Graphical Passwords

Rosanne English
School of Computing Science
University of Glasgow
Glasgow, Scotland
Email:rose@dcs.gla.ac.uk

Ron Poet
School of Computing Science
University of Glasgow
Glasgow, Scotland
Email:ron.poet@glasgow.ac.uk

***Abstract*—Recognition-based graphical passwords are one of several proposed alternatives to alphanumerical passwords for user authentication. However, there has been limited work on the security of such schemes. Often authors state a possible attack combined with a proposed countermeasure, but the efficacy of the counter measure is not always quantitatively examined. One possible attack which has been discussed without examination of the efficacy is an intersection attack. If we can establish which countermeasures for this attack are effective then it will be possible to select the appropriate countermeasure for the level of security required by a given system providing more insight into the security of these schemes.**

Our approach involved creating a simulation of intersection attacks using each of five possible counter measures. The number of attacks which had to be performed before success for each approach was noted and compared to a control where no counter measure was implemented.

Our results show that for three of the five countermeasures there was a significant increase in the number of attacks before success, one showed a significant decrease and the other did not show any statistical significance. We show that it is not decisive that using dummy screens when an incorrect image is selected will increase the number of attacks required. We also show that increasing the number of challenge screens reduces the number of attacks required before success as the number of challenge screens approaches the size of the passimage set. Our results allow a more educated selection of what type of countermeasure one should implement if they wish to reduce intersection attacks.

***Keywords*-recognition-based graphical passwords, authentication, intersect attacks, simulation**

I. Introduction

When considering alternatives for user authentication, recognition-based graphical passwords are often discussed (e.g. [2],[3],[4],[7]). Analysis of the level of security has been discussed on in terms of some of the possible attacks and potential countermeasures e.g. [1]. Whilst this is useful, it would be beneficial to gather numerical data for the proposed counter measures. To combat this, we propose an approach of simulations to analyse how effective countermeasures are against their related attacks. In this work we apply this approach to an intersection attack.

For the purposes of this work, we define an intersection attack as follows. An attacker aims to gain access to a user account. The attacker starts the challenge session which comprises of n challenge screens, each of which has one passimage from the user's passimage set (of size p) and a number of distractors (d). The attacker attacks each screen in turn by noting all the images on the screen (which refer to as the challenge set) and incrementing a count of the number of times each image has been viewed over all the attacks the attacker has launched against this set of passimages. The attacker then attempts to pass the challenge set by selecting the image which has been viewed most frequently.

This is repeated for each challenge screen within the session. If each image selected corresponds to the passimages, then the attacker is successful and the process is complete. If the attacker selects a distractor for any one of the challenge screens, they must start a new attack. This repeats until the attacker is successful. Once the attacker has achieved successful authentication, the number of attacks before success is reported. In the remainder of this paper, we refer to the setup of a recognition-based graphical password scheme as $p-n-d$, where p is the size of the passimage set, n is the number of challenge screens in a challenge session and d is the number of distractors per screen.

A number of countermeasures for intersection attacks have been identified in research, these are discussed in Section II. The aim of this research is to establish whether these countermeasures increase the number of attacks before successfully authenticating for a set of passimages. Additionally, we examine the effect of increasing the number of distractors in countering an intersection attack. Our results show three of the counter measures provide a significant increase in the number of attacks which need to be attempted before success, one countermeasure cannot be established as having a significant effect on the number of attacks and the other significantly reduces the number of attacks required.

II. The Countermeasures

Dhamija and Perrig [7] successfully summarise counter measures for intersection attacks as:

- Use the same distractor images and pass images for each session.
- Repeat a small subset of distractor images for each passimage. This would result in an attacker recording

the same frequencies for these distractors and the passimage and the attacker would be unable to tell which is the passimage. Thus, they would have to randomly select one of the images with the same frequency of occurrence.
- In any given challenge session, if a user selects a distractor on a challenge screen, subsequent screens only display distractor images - "dummy screens".
- Implement a limit on the number of incorrect authentications a user can perform, this stops an impersonator attempting to discover all of the images. (A "three strikes and you're out" approach)

The first approach where the distractors are set for a given passimage will completely negate an intersection attack. However, as Dhamija and Perrig note, reuse of distractor images may result in users recognising the distractors and selecting the wrong image for authentication [7] . Due to this, we feel it is still worthwhile to examine the efficacy of other countermeasures. In addition, as noted by Smith [10, Page 163], there are "different secrets for different uses". That is to say, there are different levels of security required for different environments. Thus it is feasible to consider that one might not wish to go to the effort of eradicating intersection attacks, but merely reduce the risk to an acceptable level. This research aims to establish how effective the different mitigation countermeasures are, specifically the following were examined:

- Repetition of a subset of distractors for a given passimage
- Use of "dummy" screens if an attacker selects the wrong image at any point within a session
- Using a passimage set which is larger than the number of challenge screens in a session (proposed in [5] and extended in [8])

Additionally, there has been no claim that increasing the number of distractors shown per challenge screen or increasing the number of challenge screens would mitigate an intersection attack and so these were also examined for significance. The hypotheses to be examined are reported in the next section.

III. Intersection Attack Algorithm

A collection of 144 images were used in the simulation, the content of the images was unimportant as they were selected randomly for both passimage sets and distractors, minimising any potential effect on the results. The control set up which involved no counter measures was as follows.

The first step was to generate a specified number of passimages from the collection of all images. To create a challenge session, the number of challenge screens to be generated matched the number of passimages. A specified number of distractors were then randomly selected from the remaining images (the complete collection, less the passimages for the current set of passimages) for each of the challenge screens required.

An attack on the set of passimages was then conducted as follows.

A list of images seen by the attacker is created. For each challenge screen presented to the attacker, the images are either added to the list of viewed images, or the number of times they have been seen is incremented. To attack the screen, the attacker takes the most viewed image on the screen (the image on the screen with the highest count in the list of viewed images) and selects that image as the passimage. If the image is the passimage, a counter for the number of screens passed in that session is incremented. If at the end of the session, the number of screens passed is equal to the number of challenge screens in a session,the set of passimages was successfully attacked and the program exits with the number of attacks which were attempted before success (the dependent variable being examined in this research). For each experimental set up ($p - n - d$ and any applicable countermeasure variables), this process was run one hundred times.

IV. Hypotheses

The dependent variable being examined in this research is the number of attacks before success. The independent variables included the number of passimages in the user's passimage set (p), the number of challenge screens per session (n), the number of distractors per challenge screen (d), the number of distractors kept constant per passimage and the use of dummy screens. The hypotheses to test the relationships between the independent variables and dependent variable were established as follows:

- H1 It takes significantly more attacks before a successful intersection attack when there are a subset of distractors kept constant between challenge sessions.
- H2 It takes significantly more attacks before a successful intersection attack when the number of distractors kept constant is increased.
- H3 It takes significantly more attacks before a successful intersection attack when dummy screens are presented if one screen in a challenge set is failed.
- H4 It takes significantly more attacks before a successful intersection attack when a passimage set larger than the number of challenge screens in a session is used.
- H5 It takes significantly more attacks before a successful intersection attack when the number of challenge screens in a session is increased.
- H6 It takes significantly more attacks before a successful intersection attack when the number of distractors per challenge screen is increased.

For hypothesis testing, where an independent variable is being altered, the remaining independent variables are kept constant so that the effect of only one independent variable

is being measured at any given time. We call the independent variable being examined the experimental variable. The corresponding null hypotheses (referred to by the hypothesis number, with a subscript of 0 after e.g. the null hypothesis for H1 is $H1_0$) detail that there is no significant difference in the number of attacks before success. The control set up had no countermeasures implemented and was used to compare to the other configurations where each countermeasure was implemented. A number of different variations were used to test each hypothesis. For each hypothesis, the set ups used were as follows:

A. H1 Experimental Set Ups

The number of distractors used were selected as eight, nine and fifteen. This was to reflect common choices in recognition-based schemes. Eight distractors are used in passfaces (http://www.realuser.com/), nine distractors are used in VIP [4] and fifteen distractors are used in the doodles scheme [9]. The last size provided a comparison of a much larger distractor set.

For this hypothesis, the number of images in the passimage set was kept consistent at four as was the number of challenge screens. The number of distractors changed, but hypothesis testing always compared a control configuration with a corresponding configuration with only the experimental variable changed (in this case the number of distractors kept constant per passimage) . For example the control of no constant distractors with a configuration of 4-4-8 was compared to one constant distractor per passimage with a configuration of 4-4-8.

The experimental variable of "number of distractors kept constant per passimage" was varied, using one distractor, two and also three. This was varied to establish if the experimental variable had a significant effect on the dependent variable (number of attacks before success). The values for the experimental variable were selected as one, two and three as all values had to be less than the number of distractors per screen. If the number of constant distractors was equal to the number of distractors per screen, no intersection attack would be successful. In total for this hypothesis, nine set ups were used, each of which was run 100 times.

B. H2 Experimental Set Ups

Three configurations were used to test the second hypothesis. The counter measure of distractors being kept constant was compared to an increased number of distractors kept constant. As for H1, the number of distractors were eight, nine and fifteen. For each of these, four challenge screens and four passimages were used and three distractors kept constant were compared to one distractor kept constant and two distractors kept constant. Thus in total there were six set ups each of which was run 100 times.

C. H3 Experimental Set Ups

To test the significance of using dummy screens, three set ups were used, one for each of the values identified for distractors (eight, nine and fifteen). The control set up with four passimages, four challenge screens and each of the distractor values was compared to the results for the dummy screens countermeasure results with the matching number of distractor values, number of passimages and challenge screens giving a total of three configurations for hypothesis testing.

D. H4 Experimental Set Ups

To test the significance of using a passimage set larger than the number of challenge screens per session (thus in any given session, a subset of the passimages is used) six configurations were examined. 1.5 times the number of passimages (six vs. four images with eight and nine distractors) in the set was examined, as was double the number of passimages (four vs. eight with eight and nine distractors), and triple the number of images in passimage set (twelve passimages compared to four with four challenge screens and eight and nine distractors). The number of challenge screens was kept constant and eight and nine distractors were used.

E. H5 Experimental Set Ups

To test the significance of increasing the number of challenge screens the number of passimages was kept constant at ten. This value had to be large enough that the number of screens was always less than the number of passimages in the set, but the number of screens could be increased. The distractors were kept constant at eight per screen and the number of challenge screens was varied using values of five, six, seven and eight.

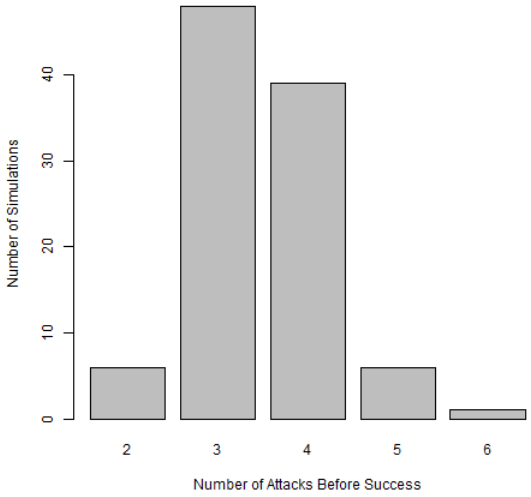
F. H6 Experimental Set Ups

The final hypothesis required configurations which examined the effect of varying the number of distractors when the number of challenge screens and number of passimages was kept constant. Configurations with four passimages, four challenge screens and eight distractors was compared to the equivalent number of passimages and challenge screens but with nine distractors and fifteen. Similarly, four passimages and challenge screens with nine distractors was compared to four passimages and challenge screens with fifteen distractors.

V. Results

The histogram showing the distribution of one hundred simulations for a control setting with four passimages, four challenge screens and eight distractors is shown in Figure 1. It can be seen from this figure that the distribution is skewed to the right, indicating that the use of standard deviation and mean may not be appropriate [6, Page 80]. The frequency

Figure 1. Control Histogram



distributions for the countermeasures configurations also indicated asymmetric distributions. Examples for each of these are provided in Figures 2, 3 and 4, some of which appear more skewed than others (e.g. Fig. 2). The non-normal distribution was also confirmed using a normal probability plot (for the control configuration of 4-4-8), as shown in Figure 5. In a normal probability plot, if the points lie very close to the straight line which represents the normal distribution, the data is normally distributed ([6, Page 115]).

Due to the asymmetric nature of the results, it was decided that a statistical approach which was robust to outliers in data and skewed distributions should be taken. We used the Yuen statistic with 20% trimmed and an alpha value of 0.05, as it is highlighted in [12, Page 157] that this approach yields robust results. The Yuen test examines the hypothesis that two independent groups have equal trimmed means and tests the null hypothesis that two samples come from the same distribution, this allowed us to examine whether the counter measure configurations were statistically significantly different from the control configurations. The Yuen test is used if the population size is small , has outliers and the probability curve is non-normal [13]. Wilcox [12] has written robust statistical functions for the statistical program "R" http://www.r-project.org/. Thus for this analysis, this program was used. In our hypothesis testing, if the Yuen test statistic value was higher than the critical value (which is automatically calculated by the R program for the data input) then the null hypothesis was rejected (this is shown in [11, Page 252]).

A. H1 - Constant Distractor Subset Results

A summary of the results for the use of constant distractor subsets is demonstrated by the boxplot in Figure 6 where one sees the effect of the number of distractors kept constant per passimage on the number of attacks required before success

Figure 2. Subset of Constant Distractors Histogram

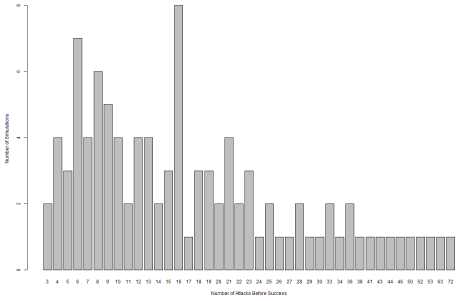


Figure 3. Dummy Screens Histogram

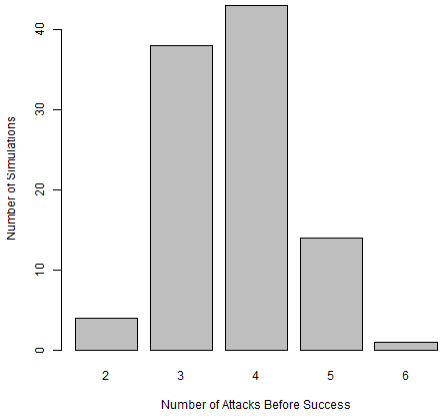


Figure 4. Larger Image Set Histogram

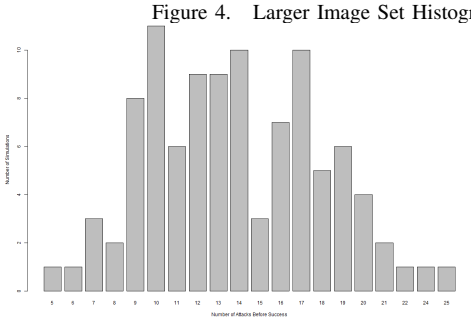


Figure 5. Normal Probability Plot

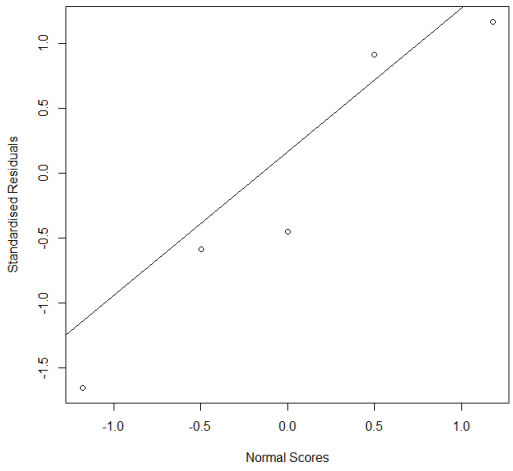


Figure 6. H1 Boxplot- Constant Distractors

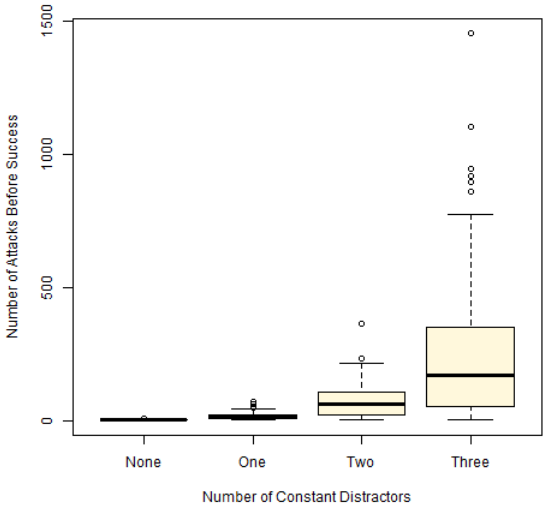


Table I
CONSTANT DISTRACTORS SUMMARY STATS TABLE

No. of Constant Distractors	Min	1st Qu.	Median	Mean	3rd Qu.	Max.
None	2.00	3.00	3.00	3.48	4.00	6.00
One	3.00	8.00	15.50	18.44	23.00	72.00
Two	3.00	23.50	60.50	73.33	103.80	365.00
Three	3.00	54.75	171.50	252.20	347.20	1454.00

for values 0 (control) ,1,2 and 3. It can be seen from this plot that the use of a number of constant distractors reduces the number of attacks required before success when compared to zero constant distractors.

In each case represented in the plot, four challenge screens with four passimages and eight distractors were used, the experimental variable (the number of constant distractors per passimage) was varied using zero, one, two and three. The values for minimum, first quartile, median, mean, third quartile and maximum for each set up is given in Table I where it can be seen that the minimum stays approximately equal in each case, but the values for median and the quartiles increase between each value for constant distractors . It should be noted that due to the skew of the distribution, mean is not an accurate measure of spread, it is included for completeness.

Applying the Yuen statistic with 20% trimmed means, the null hypothesis $H1_0$ was rejected for each of the set ups used to test H1 with the test statistic value ranging between 7.39 and 10.99 and the critical value as approximately 2.00 in each instance. This means that the number of attacks before success when a subset of distractors are kept constant is significantly more (as indicated by the Figure 6) than that when no distractors are kept constant.

Table II
DUMMY SCREENS SUMMARY STATS TABLE

Dummy Screens	Min	1st Qu.	Median	Mean	3rd Qu.	Max.
No	2.00	3.00	3.00	3.48	4.00	6.00
Yes	2.00	3.00	4.00	3.70	4.00	6.00

B. H2 - Increasing the Number of Distractors Kept Constant Results

A summary of the results for the use of constant distractor subsets is demonstrated in Figure 6 when examining the second and third and fourth boxes corresponding to one, two and three distractors kept constant respectively. It can be seen from this plot that increasing the number of constant distractors increases the number of attacks required before success. This is also shown in the statistics in Table I where the values for median increase substantially between each value. In each case here, four challenge screens with four passimages and eight distractors were used, the variable of interest (the number of constant distractors) was increased from one to two and then to three.

Applying the Yuen statistic with 20% trimmed means, the null hypothesis $H2_0$ was rejected for each of the set ups used to test H2 with the test statistic value ranging between 4.84 and 8.16 and the critical value as approximately 2.00 in each instance. This result is as expected from examination of the evidence shown in the boxplot in Figure 6. This means that increasing the number of distractors kept constant per passimage significantly increases the number of attacks before success.

C. H3 - Use of Dummy Screen Results

A summary of the results for the use of dummy screens (upon incorrect selection) is demonstrated in Figure 7. It can be seen from this plot that the use of dummy screens when the attacker selects a distractor instead of a passimage appears to have little effect on the overall number of attacks required before success. In each case here, four challenge screens with four passimages and eight distractors were used. The experimental variable (the use of dummy screens) was varied by either being used or not (when it wasn't used, this was the control configuration).

The values for minimum, first quartile, median, mean, third quartile and maximum for each set up is given in Table II. One can see from Table II that the only value which changes between the control set up (using no dummy screens) and the set up using dummy screens is the median, which changes by 0.22. Thus the use of dummy screens appears to show little evidence of an increase in the number of attacks required before success.

Applying the Yuen statistic with 20% trimmed means, confirmed that the null hypothesis $H3_0$ could not be rejected for each of the set ups used to test H3 with the test statistic value ranging between 0.79 and 1.70 and the critical value as approximately 1.98 in each instance.

Figure 7. H3 Boxplot - Use of Dummy Screens Boxplot

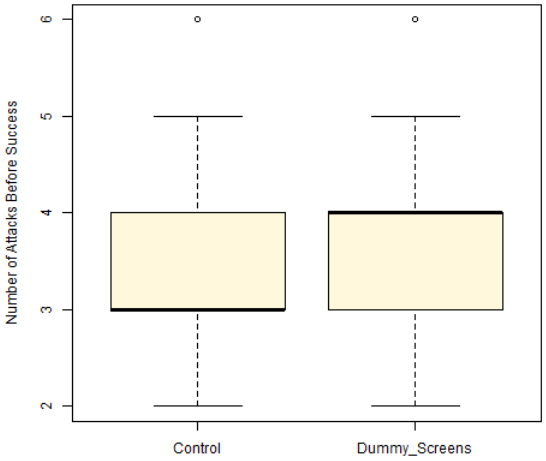


Figure 8. H4 Boxplot - Use of a Larger Image Set Boxplot

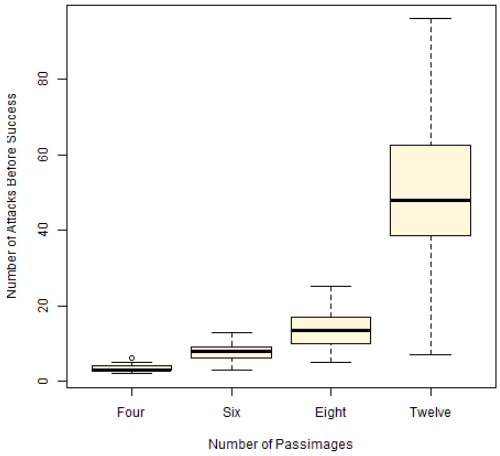


Table III
LARGER PASSIMAGE SET SUMMARY STATS TABLE

No. of Passim-ages	Min	1st Qu.	Median	Mean	3rd Qu.	Max.
Four	2.00	3.00	3.00	3.48	4.00	6.00
Six	3.00	6.00	8.00	7.79	9.00	13.00
Eight	5.00	10.00	13.50	13.85	17.00	25.00
Twelve	7.00	38.75	48.00	49.62	62.25	96.00

D. H4 - Use of a Larger Image Set Results

A summary of the results for the use of larger passimage sets (from which a subset is selected for any given authentication screen) is demonstrated in Figure 8. It can be seen from this plot that the use of a larger passimage set reduces the number of attacks required before success when compared to smaller sets. In each case here, four challenge screens and eight distractors were used, the variable of interest (the number of passimages in a set) was varied using four, eight and twelve.

There is an increase in the number of attacks when moving between four and eight, eight and twelve and four and twelve. This indicates that increasing the number of passimages has a significant effect on the number of attacks required. The values for minimum, first quartile, median, mean, third quartile and maximum for each set up is given in Table III . As expected from the boxplot, it can be seen from Table III that there is a large jump in all the staistics between each of the number of passimages.

When applying the Yuen statistic with 20% trimmed means, the null hypothesis $H4_0$ was rejected for each of the set ups used to test H4 with the test statistic value ranging between 15.89 and 24.17 and the critical value as approximately 2.00 in each instance. This is in line with the results shown in Table III and Figure 8 and means that there is a significant increase in the nuber of attacks required before success when the number of passimages in a user's passimage set is increased.

E. H5 - Increasing the Number of Challenge Screens Results

A summary of the results for the use of larger number of challenge screens is demonstrated in Figure 9. It can be seen from this plot that the use of a larger number of challenge screens appears to reduce the number of attacks required before success when compared to a smaller number of screens. In each case, the number of passimages in the set was kept constant at 10 (since the number of passimages in the set has to be larger than the number of challenge screens in each instance) and eight distractors were used, the variable of interest (the number of challenge screens) was varied using five, six and eight. The values for minimum, first quartile, median, mean, third quartile and maximum for each set up is given in Table III and this also demonstrates the decrease in number of screens as the median value reduces in each case.

We can see in the boxplot (Fig. 9) that the median is consistent when using four and five screens, but as the number of screens approaches the number of passimages in the set this reduces, we conjecture that the reason for this is due to the fact that if the number of passimages is approximately equal to the number of screens then the attacker will see the passimages more frequently, making the attack more successful thus we believe that it is not merely increasing the number of screens producing this effect, but having it close to the number of passimages.

When applying the Yuen statistic with 20% trimmed means, the null hypothesis $H5_0$ could be rejected for four of the five set ups used to test H5. The test statistic value ranging between 0.29 and 5.87 and the critical value as approximately 1.99 in each instance. Where the test statistic was not significant was for the use of five challenge screens, the remaining results established a significant difference in the distributions. This is in line with the results shown in

Figure 9. H5 Boxplot - Increased Number of Challenge Screens Boxplot

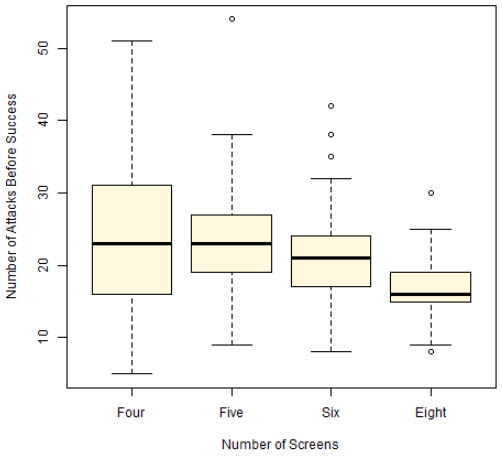


Table IV
MORE CHALLENGE SCREENS SUMMARY STATS TABLE

No. of Challenge Screens	Min	1st Qu.	Median	Mean	3rd Qu.	Max.
Four	5.00	16.00	23.00	24.09	31.00	51.00
Five	9.00	19.00	23.00	23.86	27.00	54.00
Six	8.00	17.00	21.00	20.84	24.00	42.00
Seven	7.00	15.00	18.00	18.27	21.00	29.00
Eight	8.00	15.00	16.00	16.63	19.00	30.00

Table IV and Figure 9, however as seen from the boxplot, it has a detrimental effect on the number of attacks required instead of a positive effect (i.e. it decreases the number of attacks instead of increasing them).

F. H6 - Increasing the Number of Distractors Results

A summary of the results for the use of more distractors in each challenge set is shown in Figure 10. It can be seen from this plot that the use of more distractors in each challenge screen increases the number of attacks required before success when compared to smaller numbers of distractors. In each case here, four challenge screens and four passimages were used, the experimental variable (the number of distractors per screen) was varied using values of eight, nine and fifteen. The values for minimum, first quartile, median, mean, third quartile and maximum for each set up is given in Table V . As expected from the boxplot (Figure 10), it can be seen from Table V that there is an increase in the median values between each of the variations. There is a larger increase of median attacks before success between the use of fifteen distractors and eight and nine, this is as expected as there is a larger difference in the number of distractors.

When applying the Yuen statistic with 20% trimmed means, the null hypothesis $H6_0$ was rejected for each of the set ups used to test H6 with the test statistic value

Figure 10. H6 Boxplot - Increased Number of Distractors Boxplot

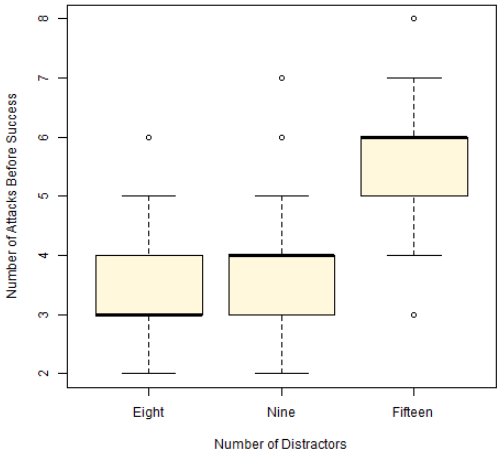


Table V
INCREASED DISTRACTORS SUMMARY STATS TABLE

No. of Distractors	Min	1st Qu.	Median	Mean	3rd Qu.	Max.
Eight	2.00	3.00	3.00	3.48	4.00	6.00
Nine	2.00	3.00	4.00	3.91	4.00	7.00
Fifteen	3.00	5.00	6.00	5.64	6.00	8.00

ranging between 3.84 and 13.67 and the critical value as approximately 1.98 in each instance. This means that increasing the number of distractors per screen significantly increases the number of attacks required before success.

VI. CONCLUSIONS

In this work we aimed to establish the effectiveness of different countermeasures against intersection attacks on recognition-based graphical password authentication mechanisms. This was an important topic to research since alternatives to alphanumerical authentication are arising more (e.g. Windows 8 is purported to be using a variation of graphical authentication for logging onto the system http://blogs.msdn.com/b/b8/archive/2011/12/16/signing-in-with-a-picture-password.aspx) but analysis of security can be limited where recognition-based mechanisms are considered.

One attack which was often discussed (e.g. [3], [4]) but has until now remained a theoretical discussion is that of intersection attacks. Due to the lack of "real world" data on the use of recognition-based graphical passwords, a simulation approach was taken. We simulated the results of attacking a recognition-based system where different countermeasures were implemented and analysed the significance of differences in the number of attacks which would have to be performed before an account was compromised.

We have shown that whilst some of the previously identified countermeasures have a significant impact on the number of attacks before an attacker would be successful,

others cannot claim to have such an effect. In particular, the use of dummy screens (where only distractor images are shown if the user selects the incorrect image on any given challenge screen within a session) did not show significant results. Another approach which showed significant results, but decreased the number of attacks before success instead of increasing them, was that of increasing the number of challenge screens presented to the user to authenticate. It is our recommendation that if a recognition-based scheme were to be implemented (where the programmer is concerned about interference challenge screens are kept constant or is not willing to implement such a measure as it would be excessive given the context of authentication) that dummy screens are not used and the number of challenge screens required for a sessions should not be close the the number of passimages.

Of the methods which achieved significant increases in the number of attacks before success, increasing the number of distractors per screen and the passimage set size and using a subset of constant distractors provide effective results. The most effective countermeasure was established as using a number of constant distractors per passimage. When comparing the median number of attacks before success to the control configuration (with four passimages, four challenge screens and eight distractors) using one constant distractor per passimage resulted in an approximate 5.17 times increase, using two resulted in an approximate 20.17 times increase and using three resulted in an approximate 57.17 times increase The second most successful countermeasure was using a larger passimage set, which when compared to a control of four passimages, six gave approximately 2.67 times increase, eight gave approximately 4.5 times increase and twelve gave an approximate increase of 16 times. The least effective of the significant countermeasures was using more distractors per screen, which when nine distractors was compared to eight resulted in approximately 1.3 times increase and fifteen resulted in a 2 times increase.

Possible future work includes the potential for establishing optimal values for each countermeasure. Another next step in this research is to continue to construct simulations in order to produce a model of the security of recognition-based graphical password schemes. This could then be used to compare any two recognition-based schemes and also to establish the countermeasures required to reflect the level of security needed for a chosen context.

REFERENCES

[1] Robert Biddle, Sonia Chiasson, and P C Van Oorschot. Graphical Passwords : Learning from the First Twelve Years. *Security*, 2011.

[2] Sacha Brostoff and Martina Angela Sasse. Are Passfaces More Usable Than Passwords: A Field Trial Investigation. In *People and Computers XIV-Usability or Else: Proceedings of HCI*, pages 405–424, 2000.

[3] D. Charrau, S.M. Furnell, and P.S Dowland. PassImages: An alternative method of user authentication. In *Proceedings of 4th Annual ISOneWorld Conference and Convention, Las Vegas, USA*, 2005.

[4] Antonella De Angeli, Mike Coutts, Lynne Coventry, Graham Johnson, David Cameron, and Martin H. Fischer. VIP: a visual approach to user authentication. In *Proceedings of the Working Conference on Advanced Visual Interfaces*, pages 316–323. ACM, 2002.

[5] Antonella De Angeli, Lynne Coventry, Graham Johnson, and Karen Renaud. Is a picture really worth a thousand words? Exploring the feasibility of graphical authentication systems. *International Journal of Human-Computer Studies*, 63(1-2):128–152, 2005.

[6] Richard D. De Veaux, Paul F. Velleman, and David E. Bock. *Intro Stats*. Pearson Addison Wesley, second pea edition, 2006.

[7] Rachna Dhamija and Adrian Perrig. Deja vu: A User Study Using Images for Authentication. In *Proceedings of the 9th conference on USENIX Security Symposium-Volume 9*, page 4. USENIX Association, 2000.

[8] Paul Dunphy, Andreas P. Heiner, and N. Asokan. A closer look at recognition-based graphical passwords on mobile devices. In *Proceedings of the Sixth Symposium on Usable Privacy and Security*, pages 1–12. ACM, 2010.

[9] Ron Poet and Karen Renaud. A Mechanism for Filtering Distractors for Graphical Passwords. In *13th Conference of the International Graphonomics Society Melbourne, Australia*, volume 11, page 14, 2007.

[10] Richard E. Smith. *Authentication From Passwords to Public Keys*. Addison-Wesley, 2002.

[11] Rand R Wilcox. *Basic Statistics: Understanding Conventional Methods and Modern Insights*. Oxford University Press, 2009.

[12] Rand R Wilcox. *Fundamentals of Modern Statistical Methods Substantially Improving Power and Accuracy*. Springer New York, 2010.

[13] Karen K. Yuen. The two-sample trimmed t for unequal population variances. *Biometrika*, pages 165–170, 1974.