Usability and Acceptance of UF-IBA, an Image-Based Authentication System

Paper No. ICCST-2007-xxxx

Piyush Harsh

University of Florida
CISE Department PO Box 116120
Gainesville, FL 32611-6120
USA
pharsh@cise.ufl.edu

Richard E. Newman
Member, IEEE
University of Florida
CISE Department PO Box 116120
Gainesville, FL 32611-6120
USA
nemo@cise.ufl.edu

Abstract – Text-based username-password systems have been traditionally used in authenticating users before allowing them access to online services. Psychological studies [1] have shown users' inability to recall random sequences of alpha-numeric strings, which theoretically make the best passwords. Image-based authentication (IBA) systems [3, 4, 5, 6, 7, and 8] show great promise in circumventing users' inherent weakness. This paper provides design details of a fully functional experimental IBA system deployed at the University of Florida's CISE department, exposes issues faced by the researchers regarding user acceptance, and suggests how to make such IBA systems more usable. This paper further provides key usability insights gained from analyzing the log files of users using the system over a period of more than two semesters.

Index Terms — user authentication, image-based authentication (IBA), usability, security, key logging, symbol entropy.

I. INTRODUCTION

Authentication of humans is based on some combination of what you are (biometric), what you have (token) and what you know (password). This paper considers knowledge-based authentication, which is by far the most common form. Text-based authentication systems are by far the most widely deployed knowledge-based user authentication mechanism, but they suffer from many weaknesses. Alpha numeric passwords are often subjected to dictionary attacks [9] and are vulnerable to social engineering attacks. They are easily compromised by even the most basic key loggers. Some of the above mentioned problems can be averted if strong passwords selection rules are enforced by system administrators. But this often results in frequent password recycling, and often passwords are written down by the user making them vulnerable to loss. Use of RSA secure random number generator dongles have also been used to strengthen user passwords in corporate settings, but economics of deployment becomes a key factor.

In this paper an easy to deploy image based authentication mechanism is described that has been developed over the past two years [2]. The system is being actively used by students of some CISE department courses for accessing course tools and grades securely. This paper also describes practical challenges that were faced and how these were overcome. Key usability parameters that adequately describe any authentication system are discussed. The log analysis results in terms of these key parameters for the UF-IBA system are also discussed. Towards the end, the direction of future research is presented. Comments and suggestions made by various users on how to further increase the system usability of UF-IBA system are also discussed.

II. SYSTEM DESCRIPTION

UF-IBA system is web-based system developed in PHP. The system allows user to register an account, select image password set, and undergo training rounds. Once these steps are successfully completed, users can gain access to the requisite resource after successful authentication attempt. Figure 1 below shows the underlying server – client communication protocol during authentication.

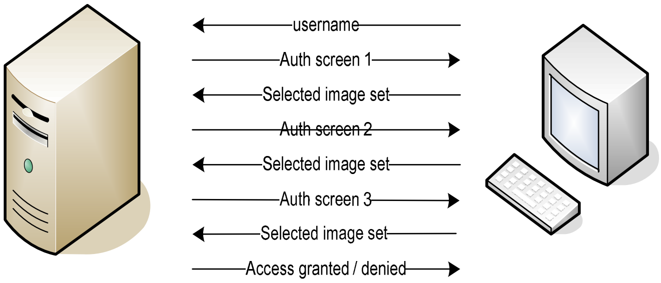


Fig. 1. UF-IBA Authentication Protocol

All messages in the protocol shown above are sent over 128 bit SSL encrypted links. The UF-IBA authentication server is stateful. It keeps state of each user in a separate one-byte state file. This allows the authentication server to track whether or not the user in question has selected the password set, undergone training rounds after selecting the password, and so on. This increases the usability of UF-IBA as the user is not forced to finish all the steps in one sitting. The user can choose to finish remaining steps at a later time, but the steps must occur in proper order. The state variable helps the server in enforcing correct order. Shown below in Figure 2 is the state transition diagram between various states in UF-IBA, after which the meaning of each state is described.

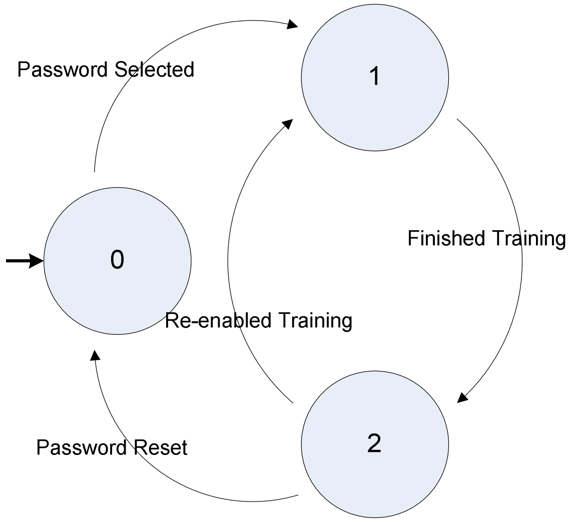
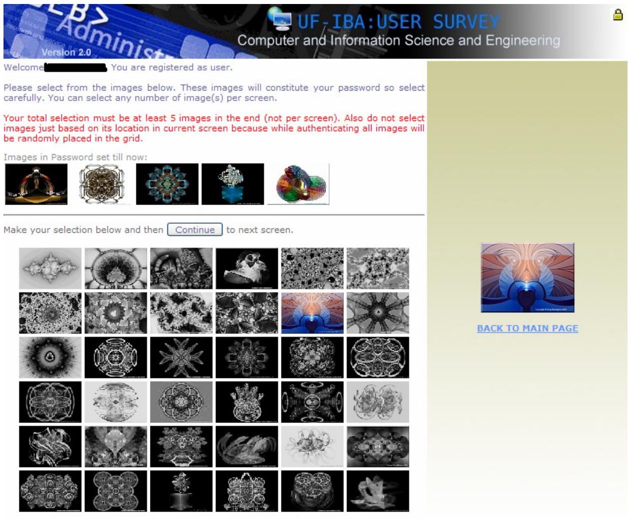


Fig. 2. UF-IBA State Diagram

In the state diagram shown above, state 0 represents that a user has registered successfully with the system but has not selected an image pass set yet. State 1 represents that the image pass set has been selected by the user but the user has not undergone training rounds. State 2 represents that the user has successfully completed the training rounds. Unless the system is in state 2 for a particular user account, that user can not login yet. Once the user successfully logs in, he has options to reset his password image set as well as reactivate the training sessions. These are represented by appropriate state transitions in the above diagram as well.

Presented in Figs. 3-7 are the actual UF-IBA system snapshots for password selection phases as well as authentication sessions for a particular user account.



Welcome [redacted] You are registered as user.

Please select from the images below. These images will constitute your password so select carefully. You can select any number of image(s) per screen.

Your total selection must be at least 5 images in the end (not per screen). Also do not select images just based on its location in current screen because while authenticating all images will be randomly placed in the grid.

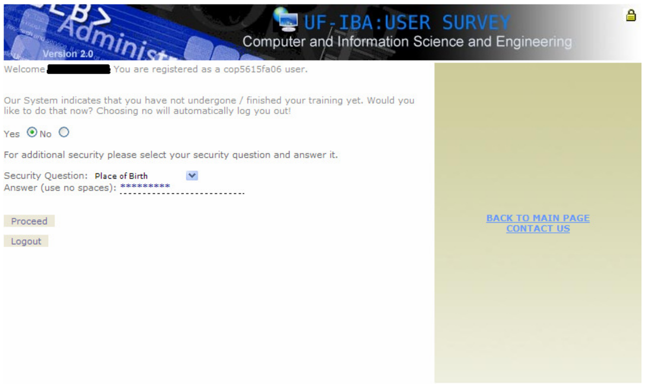
Images in Password set till now:

Make your selection below and then Continue to next screen.

Fig. 3. Password Selection Phase

The password selection phase is shown in Figure 3 above. Here, the user selects images from each screen that will be used as the image set for later authentication.

In order to prevent fraudulent changes to a user's account, any state change process is preceded by an alternate verification mechanism which is primarily done using the user's secret question and answer given during the registration phase. This is shown in Figure 4.



Welcome [redacted] You are registered as a cop5615fa06 user.

Our System indicates that you have not undergone / finished your training yet. Would you like to do that now? Choosing no will automatically log you out!

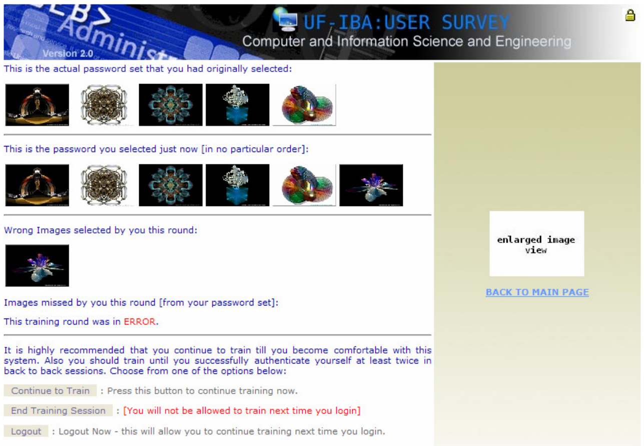
Yes ● No ○

For additional security please select your security question and answer it.

Security Question: Place of Birth
Answer (use no spaces): *********

Fig. 4. Alternate verification before training phase

Each training round provides user comprehensive information such as the original password image set along with the image set selected during the current image round, as well as wrong image (s) selected (if any), and even any missing image from his selected pass set that was not selected. Figure 5 shows a sample screen.



This is the actual password set that you had originally selected:

This is the password you selected just now [in no particular order]:

Wrong Images selected by you this round:

Images missed by you this round [from your password set]:

This training round was in ERROR.

It is highly recommended that you continue to train till you become comfortable with this system. Also you should train until you successfully authenticate yourself at least twice in back to back sessions. Choose from one of the options below:

Continue to Train : Press this button to continue training now.

End Training Session : [You will not be allowed to train next time you login]

Logout : Logout Now - this will allow you to continue training next time you login.

Fig. 5. Feedback screen after a training round

The system in the background logs critical usability data such as time of login, time spent over each screen by the user, success or failure attempt, number of images selected across various image grids, etc. Analysis of these logs is presented later when usability parameters of the system are discussed.

III. DESIGN CHOICES AND REASONS

While designing the UF-IBA system we decided to use fractal images as password elements [2] instead of various other image forms including faces or scenes. The reason for this choice was to ensure the password elements (images) are not easily transferable by word of mouth from person to another. One might argue that this feature might not necessarily be advantageous in some scenarios, but it decreases the chance of accidental password leak. Other image choices including faces and real world object photos and scenes do not provide this capability as a person can easily describe these to another person.

UF-IBA uses three image grids while authenticating a user. This choice of three screens comes from the results of the first pilot study that was conducted almost two years ago. In that study users were divided into four groups and were asked to choose their passwords across a number of screens varying from three to eight. The results of that survey clearly showed that users who had to select their password across only three screens felt the system was more usable compared to other respondents.

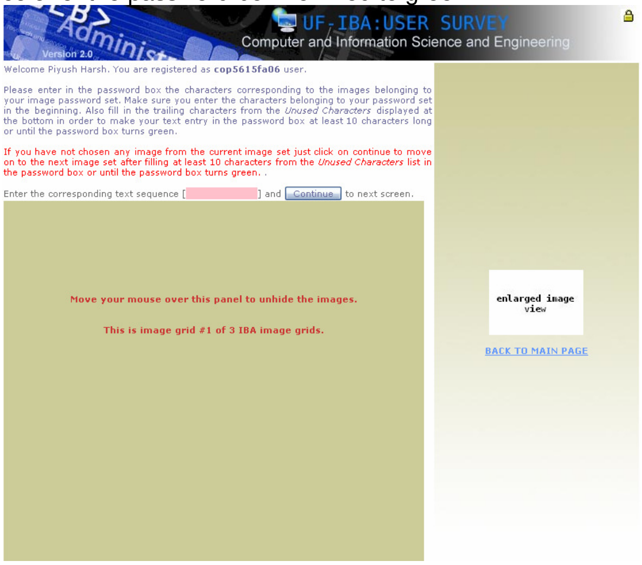
Each grid in UF-IBA contains 36 fractal images, and users are asked to choose at least five images (total) as their password set across all three screens (refer to Figure 3 above). In the worst case scenario, if a person chooses all five images from one set and zero from the remaining image grids, the password strength can be calculated to be $\log_2(C(36,5))$ where C(.) is the choice function. This value comes out to $\log_2\left(\frac{36!}{5!*31!}\right)$ or $\log_2(376992)$ bits, or approximately 18 bits. This makes the password strength comparable to at least an eight character long alphanumeric password based on usual English language letter usage frequency. Theoretically, the maximum equivalent strength of password the UF-IBA system can offer is 3 $\log_2(C(36, 18))$ bits, which comes to approximately 99 bits; this is extremely strong for passwords.

IV. DEPLOYMENT ISSUES / SOLUTIONS

One of the major challenges with any IBA system is their vulnerability to shoulder surfing attacks especially in a public setting. IBA systems that show image screens in the clear to the users and then ask them to select their pass set in the clear are not suitable to be deployed in a public setting. The UF-IBA team had to overcome this problem because most of the students in the CISE department at one time or another are asked to work in teams of up to four people on some project, and they need to access their account for submission in front of their teammates.

This deployment issue was overcome by providing the users the choice of public or private setting before they proceed to the password selection phase. When the public setting is chosen, the system generates a fresh random alphanumeric map for the displayed images in each of the three grids, and displays both the images and the corresponding alphanumeric character assigned to the images together. The images themselves are rendered unclickable.

The user is then asked to enter the characters corresponding to the images belonging to his image set from that screen in the password box, supplementing these with characters from the unused alphanumeric character bucket (which is also displayed alongside) to make his password entry for al the screens at least 10 characters long. The system provides a visual feedback when 10 characters have been entered by changing the background color of the password box from red to green.



Welcome Piyush Harsh. You are registered as **cop5615fa06** user.

Please enter in the password box the characters corresponding to the images belonging to your image password set. Make sure you enter the characters belonging to your password set in the beginning. Also fill in the trailing characters from the *Unused Characters* displayed at the bottom in order to make your text entry in the password box at least 10 characters long or until the password box turns green.

If you have not chosen any image from the current image set just click on continue to move on to the next image set after filling at least 10 characters from the *Unused Characters* list in the password box or until the password box turns green. .

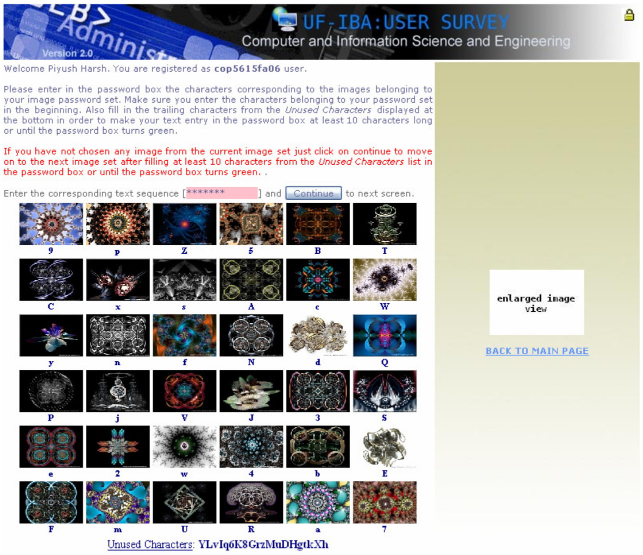
Enter the corresponding text sequence [ ] and Continue to next screen.

Move your mouse over this panel to unhide the images.

This is image grid #1 of 3 IBA image grids.

Fig. 6. UF-IBA Public setting mode – grid hidden.

As shown in Figure 6, the image grid is also hidden except when the user brings the mouse over the grid area. This design choice was made to reduce the efficacy of even the most sophisticated key logger/screen capture program. More discussion on this is deferred until the next section.



Welcome Piyush Harsh. You are registered as **cop5615fa06** user.

Please enter in the password box the characters corresponding to the images belonging to your image password set. Make sure you enter the characters belonging to your password set in the beginning. Also fill in the trailing characters from the *Unused Characters* displayed at the bottom in order to make your text entry in the password box at least 10 characters long or until the password box turns green.

If you have not chosen any image from the current image set just click on continue to move on to the next image set after filling at least 10 characters from the *Unused Characters* list in the password box or until the password box turns green. .

Enter the corresponding text sequence [*******] and Continue to next screen.

Unused Characters: **YLvIq6K8GrzMuDHgtkXh**

Fig. 7. UF-IBA Public setting mode – grid unhidden

As mentioned earlier each image element in the grid is assigned a random alphanumeric character by the authentication server in real time. The image – character mapping changes every time user tries to log in, making the system more robust. Also displayed is an unused characters bucket. Notice the background color of the password entry box in Fig. 7 is red, which will turn green only when at least 10 characters have been entered.

V. INHERENT STRENGTH / WEAKNESS

Most of the proposed IBA systems are impervious to dictionary-based password attacks. A recent study [9] presented a very interesting result where the team used a web crawler to get famous phrases, celebrity quotes, etc. off the web, and then formed strings using the first characters from each word in a phrase. Common symbol substitutions (e.g., ‘3’ for ‘e’) were used to expand these. They concluded that if a mnemonic password-specific dictionary is created, then even mnemonic passwords can be cracked unless users are trained not to use popular phrases posted on the Internet. If this case is not addressed, then in the future mnemonic passwords could become progressively more vulnerable than normal textual passwords. In that study it was found that simple dictionary with basic permutation attack with John the Ripper [10] was able to break 4% of mnemonic passwords and 11% of control passwords. Their study shows even if utmost care is taken in selection of text passwords, they are inherently weak unless truly random characters are used. At least until now, no such dictionary-based attacks are possible for most of the IBA systems.

One area where the UF-IBA system differs from other IBA systems is the choice of fractal images as password symbols. This makes word of mouth transfer of password extremely difficult. Some people may argue that this feature might not always be desirable but it makes this system more secure against any accidental or deliberate password leak. Fractals being abstract images make any social engineering attempts difficult too.

The use of random image placements in the image grid and furthermore random order of display of grids themselves makes the system immune to key logging attacks. And the use of random alphanumeric character association with individual images in the public mode makes over the shoulder snooping in any public environment largely impractical.

Logged data suggests that images used in UF-IBA system do not have significant selection bias in them, thereby increasing the entropy of the password element. This in turn makes even a short password in the UF-IBA system equivalent to a much longer English language text password. This is because of lower entropy in natural language due to biased usage frequency of characters and character combinations.

UF-IBA is still not perfect and has few weaknesses. In order to increase the system usability, it was decided to allow users to recover a forgotten password after undergoing secondary authentication based on a secret question/answer model. This necessitated that users' password be stored in some recoverable form, possibly encrypted at the authentication server. Therefore, if the authentication server itself is compromised, user passwords could be compromised. In this aspect passwords stored at UNIX systems are safer because they are stored as one way hashes together with salt (but are not recoverable, only resettable). Another drawback with any IBA system, including UF-IBA, is increased time to authenticate compared to text based password systems.

VI. USABIITY METRICS

While comparing systems or reporting any usability data, it is important to define the metrics for evaluation. Metrics such as average time taken while authenticating becomes a very important usability metric. Other metrics such as ease of password recollection, success rates over reasonable period of usage under normal circumstances, and long term memory retention of passwords could serve as good metrics for usability study of any authentication system, including IBA systems. Most of the above metrics can be calculated. But some, for example, ease of recollection, can not be logged or calculated directly unless users are surveyed.

VII. USABILITY DATA FOR UF-IBA

The first pilot study results based on which the current deployed UF-IBA system parameters were chosen will not be provided in this paper, but if desired those results can be made available upon request. In this section some analysis of log files of users using UF-IBA system in fall 2006 semester is provided.

There were a total of 58 students excluding UF-IBA system administrators that used the system on a regular basis to access course tool applications that allowed users to submit home works and projects as well as access their grades and course statistics.

Table I
Demographic Data of Users

Demographic Data			
Age	Freq. (m)	Freq. (f)	Tot Freq.
20	1	0	1
21	1	2	3
22	6	4	10
23	13	4	17
24	8	1	9
25	5	0	5
26	1	1	2
27	2	0	2
28	1	0	1
29	0	0	0
30	2	0	2
31	1	0	1
32	2	0	2
33	0	0	0
34	0	0	0
35	0	0	0
36	1	0	1
37	1	0	1
38	0	0	0
39	1	0	1
Total	46	12	58

Table I shows the demographic distribution of users for which the usability data was analyzed. Figure 8 shows the gender and age demographics in histogram form.

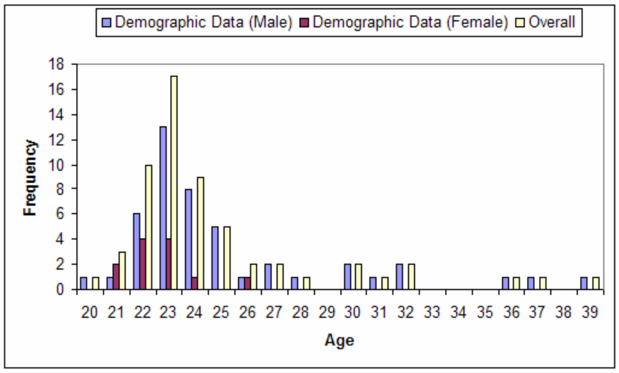


Fig. 8. Demographics Plot

The next graph shows the selection frequency of images from the three image grids. In the password selection phase, users were presented with three image grids of 36 images each in the order bucket 1, bucket 2 and then bucket 3. This greatly explains higher selection frequency of images from bucket 1 and bucket 2 compared to bucket 3. The X-axis represents image names.

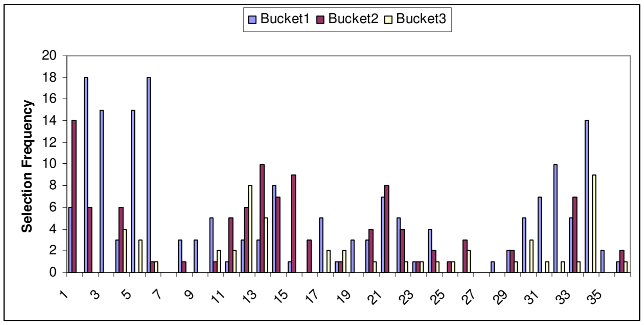


Fig. 9. Image Selection Frequency

Because of slight bias in the image selection frequency, even though it is much better compared to English language letters, there was a slight reduction in symbol entropy per image bucket.

Table II below shows the entropy value computed individually for three image buckets as well as overall symbol entropy of images across all the three buckets. Note that perfect entropy at the bucket level would be 5, while overall perfect entropy would be 6.75.

Table II
Symbol Entropy Values

Bucket	Entropy
1	4.465
2	4.123
3	3.999
Overall	5.714

Examining the entropy values, it can be seen that UF-IBA system offers better security per password element compared to traditional text-based passwords.

An automatic log analyzer program was developed that analyzed each user's log file and reported back critical usage data such as whether an authentication session was a success or failure, the time spent by the user doing authentication, and the time gap between successive authentication sessions. Next the summary view of these reported data is given. The log analyzer program also reported finer data such as time spent choosing pass images per screen, but these data are not reported here.

Table III
Global Success rate vs. Time Gap

Login Gap (Days)	#Success	#Failure
0	1346	194
1	336	26
2	189	18
3	115	8
4	88	12
5	64	15
6	40	8
7	25	5
8	15	4
9	19	3
10	13	4
11	11	4
12	9	1
13	4	3
14	3	0
15	5	2
16	2	0
17	2	1
18	4	1
19	2	1
20	3	0
21	1	0
22	1	0
23	1	0

In Table III above, 'Login Gap' represents the number of days between successive authentication attempts by the users. The data given are global across all users' attempts.

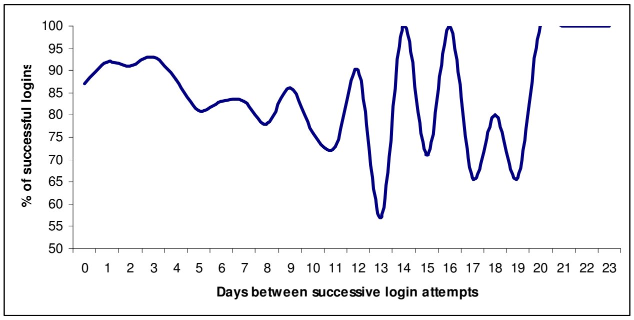


Fig. 10. Global success rate Vs gap between logins

Figure 10 above represents the plot of global login attempts success rates vs. gap between successive login attempts (in number of days). The plot data is encouraging in the sense that users are able to recall their passwords even after long durations of no use. The worst case according to the above plot was success rate of 57% after no use gap of 13 days, which is very promising. Overall, a total of 2608 login attempts were made, out of which 2298 were successes and 310 were failed login attempts. The overall success rate was 88.1%, which is very impressive.

Table IV
Global Median Access Time

Login Gap	Success	Failure
0	28	68
1	35	57
2	36	60
3	41	102
4	38	48
5	38	66
6	39	92
7	44	50
8	38	108
9	37	44
10	37	54
11	76	72
12	40	61
13	40	54
14	43	-
15	27	134
16	108	-
17	150	81
18	41	60
19	51	132
20	65	-
21	25	-
22	33	-
23	79	-

Table IV shows global median values of access time in seconds for both successful login attempts as well as failed login attempts. Comparing these values with the number of login attempts made from Table 3 shows that for majority of access attempts (51.6%), median access time for successful attempts was only 28 seconds. Although this is still large compared to authentication time for text-based passwords, we believe this value can be brought down further by making users do more training rounds in the beginning. More field testing is needed to prove this hypothesis.

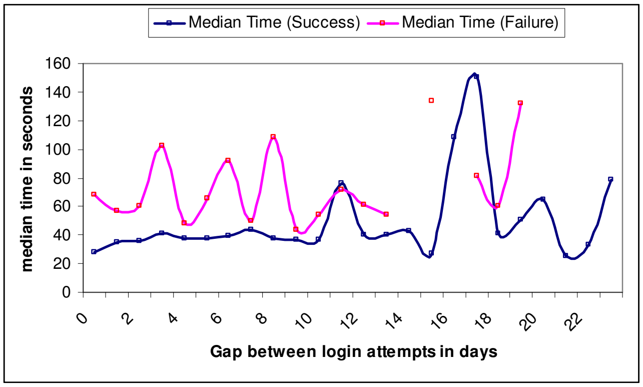


Fig. 11. Median time for successful and failed login attempts vs. gap between attempts in days

As expected, these data (shown as a plot in Figure 11) confirm that median time for an authentication attempt is typically smaller for successful attempts compared to failed attempts.

To summarize, it was found that the UF-IBA system becomes progressively more usable with passage of time even though the system initially appeared very daunting to many users because of its use of abstract fractal images.

VIII. OBSERVATIONS & FUTURE RESEARCH

Fellow peers in CISE have pointed that use of fractal images makes this system somewhat difficult to use in the beginning. Log files also show that this is true. Users' success rates show much improvement over time, with a corresponding decrease in authentication time as experience is gained.

One suggestion made by many students is to allow a few if not all images to be provided by the user. That may reduce the strengths that fractals have to offer and also decrease the symbol entropy significantly but further study has to done to determine if this is the case.

Our next goal is to remove the server password storage vulnerability. We are researching ways to make user password recoverable yet immune to being accessed by attacker in case of server breach. We are also planning to conduct a university wide usability test of UF-IBA.

IX. ACKNOWLEDGEMENTS

The authors are highly indebted to Mr. Pio Saqui, a fellow anthropologist at University of Florida for his valuable comments on UF-IBA system and tips on increasing the system usability. The authors are also indebted to Mr. Panoat Chuchaishri for his constructive criticism of the system throughout the period of its development and deployment. Sincere appreciation for students of CEN5540 – Computer and Network Security, spring 2005 class for taking part in the pilot study that provided UF-IBA team with valuable data that shaped the development of the currently deployed system. We are also thankful to students of COP5615 – Distributed Operating System – fall 2006 class and CEN5501C – Computer Network – spring 2007 class for using the UF-IBA system, providing real usage data.

X. REFERENCES

[1] David Melcher, "The persistence of visual memory for scenes", Nature, 412(6845) p. 401, July 2001

[2] Richard E. Newman, Piyush Harsh and Prashant Jayaraman – "Security Analysis of and Proposal for Image-based authentication", proceedings of IEEE ICCST 2005, p. 141, October 2005

[3] G.E. Blonder, Graphical Passwords, US Patent 5559961, Lucent Technologies Inc., August 1999.

[4] R. Dhamija and A. Perrig, "Déjà vu: A user study Using Images for Authentication", proceedings of 9th USENIX Security Symposium, August 2000.

[5] C. Perra and D. Giusto, "A Framework for Image Based Authentication", IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.

[6] S. Wiedenbeck, J. Waters, J. C. Birget, A. Brodskiy, N. Menon, “PassPoints: Design and longitudinal evaluation of a graphical password system”, International Journal of Human Computer Studies, p. 102-127, 2005.

[7] Realuser, passfaces: web resource link - http://www.realuser.com/

[8] UF-IBA Mirror Site: web resource link – http://www.cise.ufl.edu/~pharsh/iba/

[9] Cynthia Kuo, Sasha Romanosky, and Lorrie Faith Cranor, “Human Selection of Mnemonic Phrase-Based Passwords”, SOUPS, July 2006.

[10] Web Resource Link: John the Ripper - http://www.openwall.com/john/

XI. VITAE

Mr. Piyush Harsh earned his bachelor’s degree in Computer Science and Technology from the Indian Institute of Technology, Roorkee, in May of 2003. He is currently enrolled as a PhD student in the Computer and Information Science and Engineering Department of the University of Florida. He has previously published work with his advisor, Dr. R. Newman, on image-based authentication, including a paper in the 2005 ICCST. He is also working in the area of internet multicast address management.

Dr. Richard Newman earned his bachelor’s degree in Mathematics from New College in Sarasota, FL in 1981, then his Master’s and PhD in Computer Science from the University of Rochester, NY in 1983 and 1986, respectively. Since then, he has been on the faculty of the Computer and Information Science and Engineering Department of the University of Florida. He has published over eighty refereed papers in the areas of computer security, networks, and systems. He has been a member of IEEE since 1986 and has been active in IEEE P1901 and IEEE TCPLC since their inception.