

Design, Privacy and Authentication of Challenge Questions in Online Examinations

Abrar Ullah

School of Computer Science
University of Hertfordshire
Hatfield, United Kingdom
a.ullah3@herts.ac.uk

Hannan Xiao, Mariana Lilley, Trevor Barker

School of Computer Science
University of Hertfordshire
Hatfield, UK
{h.xiao, m.lilley, t.l.barker} @herts.ac.uk

Abstract— Online examination is an essential part of the online learning and secure authentication is considered vital for the success of online learning. This study is part of an ongoing research on student authentication approaches and the use of challenge questions in online examination authentication. This paper presents the results of an empirical study based on “Profile Based Authentication Framework” (PBAF), which uses challenge questions for student authentication in online examination. The PBAF uses challenge questions related to personal, academic and professional information. These questions inform the usability, security and privacy of PBAF authentication approach. The results presented here summarizes the impact of questions design on the usability based on data collected from challenge questions authentication and a post-experiment survey on the data privacy.

Keywords- challenge questions; online learning; online examination; usability; data privacy

I. INTRODUCTION

Security of online examination is vital to the success of online learning. Student authentication in online examination has been an active research area and a number of authentication procedures have been evolved over time to ensure secure authentication. Online examinations are reported to be more vulnerable to academic dishonesty and authentication attacks due to lack of physical interaction [1]. Recent studies indicate that unethical conduct have intensified in online learning due to un-controlled environment for cheating in the online examinations as a result of use of technology and the Internet [2]. Collusion is one of the potential security threats to online examinations wherein students can collude with a third party to take their examinations for extra grades [3]. In response to authentication attacks, the authors developed a Profile Based Authentication (PBAF) for student authentication in online examinations [4]. The PBAF implements challenge questions to authenticate students before their access to online examinations. Challenge questions authentication has been widely used as a fallback feature for credential recovery and reported with usability issues [5]. The use of challenge questions in online examinations can be an effective authentication solution, if the questions are designed citing usability, security and privacy aspects. The balance of these quality attributes can help in designing reliable authentication mechanism.

Evidence suggests that a question design has proportionate effect on the usability, privacy and security. The questions can be designed using a number of characteristics. We have performed an empirical study based on a solution presented in [4] using PBAF. The study involved participants from the UK and overseas Universities. The results were analyzed based on participant’s answers during online learning and examinations stages. An online survey was conducted on completion of the empirical study to collect participants’ feedback on their privacy concerns while answering the personal and academic questions during the study.

II. BACKGROUND

Student authentication in online examinations are integrated and extremely important component of online learning [6]. Evidence suggests that student authentication in the online examinations can face many security threats. It has been widely researched area and a large number of authentication techniques have been developed, which can be implemented to enhance the security of online examinations. This study is follow up analysis of the previous work [4],[7],[8]. The authors developed PBAF authentication method and analyzed various usability attributes and questions design based on the data received from an empirical trail. In this study, the data was analyzed for the error rate and effectiveness usability attributes. The data was collected from challenge questions for student authentication in the online examinations phase. The findings from participants’ feedback on privacy of questions are presented based on the analysis of post-experiment questionnaire.

The use of *challenge questions* is an important authentication and password recovery technique [9]. It has been widely used for password recovery by corporate email service providers as Google, AOL, Yahoo and Microsoft to name a few [10]. The challenge questions authentication approach has been used in the online and telephone banking sector for verification of customers’ identity[11].

The PBAF approach presented in this study implements challenge questions in combination with user-ID and password to authenticate students. The students are required to submit their answers to questions in the learning phase. It is anticipated that learning is performed in multiple visits by the students in order to access the learning resources. Answers to

profile questions received in the learning phase are used for authentication at a later stage, when access to online examination is requested. The PBAF generates random questions from the stored answers for authentication purpose.

III. QUESTIONS DESIGN

The PBAF is implemented by information from student answers to profile and challenge question, hence, question design is highly important to ensure usability, security and privacy. People tend to share common themes of their personal information and reluctant to provide details of personal information due to privacy and data security concerns. The questions used for PBAF authentication are designed in the various categories.

A. Questions Category

We designed our challenge questions from the following categories and picked 20 questions for the empirical study shown in Table 1:

- *Personal Descriptors*: The information, which describes a common theme of a person. The question can use various attributes including name, age, date of birth, gender, weight, height, eye color, hair color
- *Identification Numbers*: A wide range of numbers which can verify individual's identity including: House number, telephone number, mobile number, student number, PIN number, Social Security number.
- *Ethnicity*: The information regarding individuals ethnicity including race, ethnic origin or nationality, color
- *Health*: The information pertaining individual's health including physical and mental health, e.g. blood type, disability etc.
- *Financial*: The information regarding individual's financial status including income, bank details etc;
- *Employment*: The employment history including job title, role, dates etc;
- *Life*: The information about life includes individual's likes, dislikes, character, general reputation, social status, marital status, religion, political beliefs etc;
- *Education*: The educational information includes current education status and previous education history.

B. Open and Close Questions

As in [5], the two commonly used types of questions are fixed and open questions. The fixed type questions are pre-designed, where the questions are selected from the pool of already created questions by the *system administrator*. In contrast, open questions are created by the users' choice and they enjoy full control over the question building. The PBAF application uses fixed type questions, where the questions are pre-designed. These questions are referred as profile and

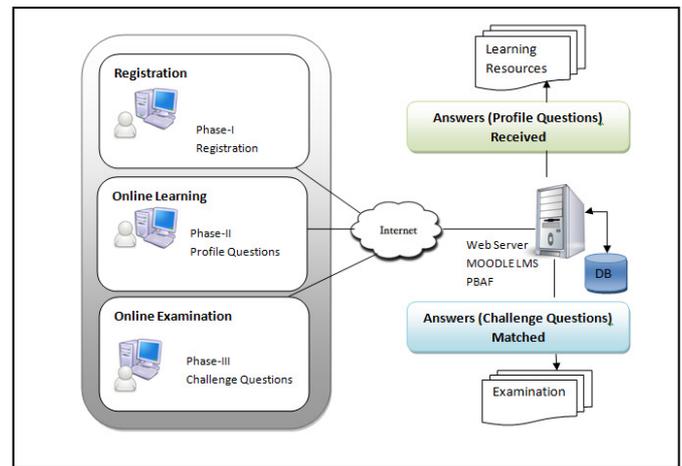


Figure. 1 Study Phases

challenge questions randomly selected both at the learning and authentication phases.

C. Answer type

Like varied question types, the answers are also categorized in different types. The answers can be formatted e.g. date format, formatted numbers, multiple word answers, single word answers etc. The question and answer type most likely defines the usability and security of challenge question approach.

IV. RESEARCH METHODOLOGY

The study was designed to analyze usability, security, and reliability of PBAF in the online examination. As in Fig.1, we performed the experiment in three phases. The questions were designed from the selective categories keeping in view data privacy and security issues. A cohort of 20 fixed type questions were created with 10 questions each from personal and academic themes. The study was organized on a sample size of 34 online students from University of Hertfordshire United Kingdom and Institute of Management Science, Pakistan, where 23 students actively participated in the experiment and all the post experiment surveys. The participants were liaised via email for all correspondence, support and guidance. The participants were communicated the purpose of the study and a step by step guidance to complete the experiment and surveys.

1) Adding Questions to PBAF

The questions presented in Table 1, were designed to collect information pertaining personal descriptions, education and life. In order to minimize data privacy concerns, we avoided to include questions that could lead to possible participants' reluctance to submit genuine information. These questions were added to PBAF at the outset.

B. PBAF Online Experiment

The PBAF was developed in PHP (PHP:Hypertext Preprocessor) server side scripting language and MySQL relational database. PBAF was integrated with Modular Object Oriented Distributed Learning Environment (MOODLE) Learning Management System (LMS) as a block. The PBAF

offers configuration interface of questions, profile questions, challenge questions and authentication results. We deployed the LMS on an online test server. For the empirical study, we created a mock-up course and an online examination on the online server. Study was performed in phases presented below.

1) *Registration*

In phase-I, the participants were emailed the steps to create their *user-id and password* for login, through a self Registration form. The mock-up course was only available to registered participants. Of the sample of all 34 potential participants, 23 accessed registration phase and created their logins.

2) *Profile Building*

In phase-II, all the participants were requested to access the online mock-up course at different given dates. On each login to the online course, the participants were redirected to submit answers to already configured 3 profile questions in order to access the course. All the answers to *profile questions* were stored as individual participants' *profile*. The profile represents description of individual participant in the form of questions and answers. The information was stored in the database table as alphanumeric text. The participants were emailed on scheduled dates to repeat phase-II until all the 20 *profile*

questions were answered.

3) *Challenge Questions Authentication*

On completion of phase-II, the participants were emailed to access the mock-up online examination. In phase-III, the participants accessed the online examination and were redirected to answer the challenge questions. The PBAF randomly selected 3 challenge questions from the individuals' profile. The participant's answers to challenge questions were compared with the stored answers in the profile using text-to-text string comparison. On successful match, the participants were authenticated and granted access to online examination link. The participants failing to provide correct answers were denied access and reported to the course administrator.

C. *Usability and Questions Analysis*

In the previous work [8], the authors analyzed usability factor of PBAF including user interface design, recall and use of the traffic light system. In another study [7], the authors discussed the impact of questions design to evaluate efficiency usability attribute by comparison of time and anticipated answers length. This study analyses the question type, answer type and authentication results with respect to varying questions design.

Table.1 Summary of Privacy questionnaire and outcome of challenge question authentication

Personal Questions				
Questions	Privacy Mean (median)	Challenge Questions Authentication		
		Impressions	Correct	Incorrect
What is your date of birth?	1.9 (2)	5	80%	20%
What is your country of birth?	1.5 (1)	5	100%	0%
What is your father's surname?	1.8 (1.5)	3	67%	33%
What is best friend's surname?	2.0 (2)	6	100%	0%
Who is the favorite hero of your childhood?	2.0 (2)	3	100%	0%
What was your dream job as a child?	1.9 (2)	2	50%	50%
What is your home address house name or number?	1.6 (1.5)	4	0%	100%
What is your home address town?	3.1 (3)	4	25%	75%
What is your home telephone number including country and area code?	2.9 (3)	2	50%	50%
What is your mobile number including country code?	2.5 (2)	1	0%	100%
Academic Questions				
What is your student number?	1.8 (2)	1	100%	0%
What is the name of your last school attended?	2.2 (2)	5	60%	40%
What was your highest qualification before starting this course?	1.9 (2)	4	75%	25%
What were your grades in your highest qualification before starting this course?	1.9 (2)	2	0%	100%
In which year were you awarded your highest qualification?	2.1 (2)	4	75%	25%
In which month have you started the current course?	1.8 (2)	1	0%	100%
In which year have you started the current course?	1.9 (2)	3	100%	0%
What is the first name of your favorite tutor ever?	1.8 (2)	6	17%	83%
What is your favorite module on this course?	2.2 (2)	3	0%	100%
Where did you find about this course e.g. Internet, Friend etc.?	1.8 (2)	3	67%	33%

V. RESULTS

Of the 23 participants in initial registration, 18 participated in different phases of the empirical study including, profile building and online examination. In the PBAF authentication, participants were required to provide exactly the same answers, submitted in phase-II of the experiment during profile building. The participants were presented random questions from their profiles in the PBAF authentication. A total of 13 participants answered 66 *challenge questions* in phase-III of the experiment. Of the total 66 answers, 38 matched with the individual profiles and 28 were incorrect due to various usability issues. From the data collected in the phase-III, we observed that a large number of incorrect answers were triggered by questions design discussed below.

A. Questions Analysis

Design of profile and challenge questions for authentication can be an important step towards optimizing usability. Analysis of data collected from the empirical study suggests that, of the total 28 unmatched answers received during authentication, 25% failed challenge questions were related to address information. The address information promoted multiple words answers and syntactic ambiguity. The address related question had a generic design to cope with international audience. As an example, 54% participants submitted full home address to profile question “*What is your home address house name of number*”. In the authentication phase, participants failed to produce exact answers in all their 4 attempts. Similarly, answers to “*What is your home address town?*” resulted in 75% incorrect answers in authentication phase due to question clarity. In a learning environment, where learning is a primary focus for the learners, clarity of profile and challenge questions can be important for effective usability. We received 11% of unmatched answers during authentication, where the questions needed better clarity, for example “*what is your favorite module on this course*” can have different answers at different stages of the course. Also, “*this course*” in question in the context of a simulation course was not realistic. Questions related to “*telephone numbers*” and “*dates*” resulted in 14% unmatched authentication. However, 29% unmatched answers were a combination of personal and academic questions with no indication of ambiguity, clarity and capitalization issues. A 21% unmatched answers were caused by capitalization mistakes, which needs further analysis in terms of security and usability to be tweaked in the matching algorithm.

B. Challenge Questions Effectiveness

The data collected from the PBAF authentication phase-III is presented in Table.1 under the columns “Impression”, “correct” and “Incorrect” to show error rate. Since the PBAF generate “*Challenge Questions*” randomly from individual’s profile, the “impression” indicates, the number of times a question was presented to different participants during the online examination authentication. We noted that a highest percentage of matches were found for personal questions related to “*date of birth*”, “*country of birth*”, “*father’s surname*”, “*best friend’s surname*” and “*childhood hero*”. The 20% “incorrect” responses to the “*date of birth*” question was triggered by variation in the syntax of date format. In spite 4

impressions, answers to “*home address name or number*” were 100% incorrect largely because of ambiguity in the questions reported above. We noted that, answers to the same questions at both phases i.e. phase-II profile building and phase-III challenge questions were largely multiple words, realistic but unmatched due to question ambiguity. The results for questions relating “*dream job*”, “*telephone number*” and “*mobile number*” were posed in a small number of impressions and require further analysis on a larger sample size.

The academic questions of subjective nature received a high number of incorrect answers in the authentication phase. The subjective nature question “*your favorite tutor*” received 83% incorrect answers at phase-III during authentication in 6 impressions. The question “*favorite module on this course*”, was unrealistic due to a simulation online course and subjective in nature, resulted in 100% incorrect response. The *favorite module* in question can be variable and the answer can vary at different stages of the learning process.

The results indicate that the questions design is highly important to optimize usability and security of PBAF approach. Since the questions were generated randomly and the sample size is not uniform for all questions, experiment on a large sample size is warranted.

C. Data Privacy Results

Of all the 23 participants, 8 submitted their feedback in a post-experiment questionnaire on data privacy based on their experience of using PBAF authentication. We requested the feedback regarding all the 20 questions used in the empirical study. The online survey questionnaire was based on a five-point *Likert* scale, where 1 indicated strong agreement and 5 measured as strong disagreement.

We computed the questionnaire results in Table.1, under the column “Privacy Mean (median)” listing the mean and median values. The questions were designed into two main categories personal and academic questions. We asked the participants to submit their feedback, if they were satisfied to answer the questions in term of the data privacy. As in Table 1, to personal questions relating “*date of birth*”, “*country of birth*”, “*father’s surname*”, “*friend’s surname*”, “*favorite hero*”, “*dream job*” and “*house name or number*”, the participants gave positive feedback with the mean value between 1 and 2, regarding their privacy concerns. However, in contrast to “*home country*” the score of 3.1 mean to question “*home town*” indicates a degree of privacy concern. The participants gave 2.9 and 2.5 mean scores to “*home telephone*” and “*mobile number*”. The feedback suggests that some participants were reluctant to provide specific information of their exact locations and contact numbers. The data privacy concern to a question pertaining “*date of birth*” can vary in different regions depending on the use of information in secure transactions e.g. Banking , social security etc;

As in Table 1, under “Academic Questions”, the participants’ feedback indicates their agreement to providing answers to academic questions. To all the academic questions, the participants gave a mean response between 1 and 2.2.

VI. CONCLUSION

The findings from the empirical study reported here suggest, that challenge questions based authentication in online examination can be an effective feature, if the questions are designed keeping a balance of privacy, usability and security. In this study we used a combination of personal and academic questions for student authentication in online examinations. Participants responded positively to a post experiment online survey on the privacy of questions.

Our results revealed that participants posted incorrect answers to questions with clarity and ambiguity issues. Since the challenge questions are selected randomly from the individual's profile, therefore, any shortcomings in the questions design can have detrimental effect on the PBAF authentication approach. The post experiment questionnaire indicates that data privacy needs to be considered while designing challenge questions. Separate analysis is needed involving a broad spectrum of local and international participants on privacy of questions. The views of participants on "date of birth" and "address" information can be different in various countries where this information can be sensitive and used in financial transactions and personal services.

Future work would be directed for questions analysis in term of usability, security and data privacy on a larger sample size. For enhance usability and security, we are aiming to use image questions together with personal and academic questions.

REFERENCES

- [1] Harmon O. R., Lambrinos J., Buffolino J. "Assessment design and cheating risk in online instruction." *Online Journal of Distance Learning Administration*. 2010;13(3).
- [2] Agulla E. G., Rifón L. A., Castro J. L. A., Mateo C. G. "Is My Student at the Other Side? Applying Biometric Web Authentication to E-Learning Environments" in *Eighth IEEE International Conference on Advanced Learning Technologies*; 2008: IEEE.
- [3] Ercole A., Whittlestone K., Melvin D., Rashbass J. "Collusion detection in multiple choice examinations." *Medical education*. 2002;36(2):166-72.
- [4] Ullah A., Xiao H., Lilley M. "Profile Based Student Authentication in Online Examination" in *International Conference on Information Society (i-Society 2012)*; 2012: IEEE.
- [5] Just M. "Designing Secure Yet Usable Credential Recovery Systems with Challenge Questions" in *CHI 2003 Workshop on Human-Computer Interaction and Security Systems*; 2003. Florida, USA: Citeseer.
- [6] Karaman S. "Examining the effects of flexible online exams on students' engagement in e-learning." *Educational Research and Reviews*. 2011;6(3):259-64.
- [7] Ullah A., Xiao H., Lilley M., Barker T. "Using Challenge Questions for Student Authentication in Online Examination." *International Journal for Infonomics (IJ)* 2012;5(3/4):9.
- [8] Ullah A., Xiao H., Lilley M., Barker T. "Usability of Profile Based Student Authentication and Traffic Light System in Online Examination". in *The 7th International Conference for Internet Technology and Secured Transactions (ICITST-2012)* London: IEEE; 2012.
- [9] Just M., Aspinall D. "Challenging challenge questions" in *Socio-Economic Strand*; 2009: Oxford University UK.
- [10] Schechter S., Brush A. J. B., Egelman S. "It's No Secret. Measuring the Security and Reliability of Authentication via" in *30th IEEE Symposium on Security and Privacy*; 2009: IEEE.
- [11] Bignell K. B. "Authentication in an Internet Banking Environment; Towards Developing a Strategy for Fraud Detection" in *Internet Surveillance and Protection, 2006 ICISP'06 International Conference on*; 2006: IEEE.