Carleton University, Ottawa, Canada
School of Computer Science
Winter Term, 2020

COMP3008 Project 2
Quantitative Usability Evaluation
Due: April 7th, 2020 at 5pm EDT

This project is to practice and explore the elements of Quantitative Usability Evaluation as presented and discussed in class and the textbook. The application domain is knowledge-based authentication: password systems.

There are two main parts to the project:

1. The first involves using descriptive statics on sample data provided to you. The data is from a user study testing different authentication schemes.
2. In the second part, you will design and develop a new authentication scheme, test it with users, and then use inferential statistics to compare it with an established scheme.

GROUPS:

The project is to be done working in small teams of 4 people. You are responsible for organizing and managing your group. Please use the forum on cuLearn to make connections, and contact the Instructor if assistance is required. You should have your team chosen by **Friday, February 28th, 5PM**. * If not in a group by then, I will randomly assign the students to groups.

ETHICS:

Several steps of the project involve working with study participants to help you in design and in determining usability. The project has been reviewed and approved by the Carleton University Research Ethics Board. At all times, you must adhere to the ethics in user testing presented in the class 5. When working with participants, you must first obtain a participant's informed consent using the consent form and conform to commitments made on the form. This form is available on the COMP3008 cuLearn page. You must not alter or edit the consent form, except where indicated (e.g., putting your name). **You are not allowed to audio record participants, because it is not included in our ethics application.** Completed consent forms must be scanned or photographed and included as an appendix to your report. Physical copies must be dropped off in the assignment box setup in HP 3115.

REPORT FORMAT:

This report must be written according to the requirements in this section.

Cover page: Project title, team name, team member names and student IDs, course ID (COMP3008), and submission date.

Workload Distribution and Summary: Describe the contributions of each team member. This may take the form of an itemized list, but include enough detail to convey the amount of effort and value of the contribution towards the overall milestone goals. In general, teams will earn the same grade. However, significant variability in the perceived effort and value of

contributions will result in adjusted grades for individual team members.

Content formatting:

- 12-point Times New Roman font
- 1.5 line-spacing
- 1-inch margins.
- Page numberings.

Writing quality:

A deduction of up to 5% will be applied for poorly written reports. Reports must adhere to the following writing guidelines:

- Formatted according to project specification.
- No typos.
- No grammatical errors.
- Proper use of heading and sub-headings.
- Report must be organized and have proper structure.
- Images and tables must have captions.
- Images / tables should be part of the text and not included as an appendix.
- No long or run-on sentences.
- Paragraphs should have logical flow from one to another.

Page lengths: Suggested page lengths are provided for each part of the project. A page is approximately 350 words.

PASSWORD AUTHENTICATION:

Authentication is the process of determining the truth of a claim, typically relating to identity or the right to access to services or resources. Authentication is typically determined on the basis of "something you have", such as a physical cards or tokens, "something you are", such as biometrics like fingerprints, or "something you know", such as passwords.

Despite the large number of options for authentication, text passwords remain the most common choice for several reasons:

1. They are easy and inexpensive to implement;
2. They are familiar to essentially all users;
3. They allow users to authenticate themselves while avoiding privacy issues that have been raised about biometrics;

4. They have the advantage of portability without, for example, having to carry physical tokens.

However, text passwords also suffer from both security and usability disadvantages. For example, passwords are typically difficult to remember, and are easy to guess if user-choice is allowed. This is sometimes called the "Password Problem".

In recent years, many novel approaches have been suggested to improve passwords to address the password problem. It is therefore important to carefully assess the security and usability of the new approaches. For more information, please see the following paper, available on the CULearn: R. Biddle, S. Chiasson, P.C. van Oorschot (2012). Graphical Passwords: Learning from the First Twelve Years. ACM Computing Surveys 44(4).

In this project, we focus on schemes that assure a certain level of security, but where we need to assess the usability. In particular, we focus on assigned random passwords with consistent password spaces. Our primary measures of usability will be the memorability of passwords, and the speed of password entry

DATASET:

In this part of the project, you will be required to analyze data from a user study that has been completed. Sample data from the study is provided, and your work is to process the data and interpret the results. The anonymous data comes from a research study testing five different password schemes using the MVP framework, as discussed in class. The MVP framework allows different password systems to be implemented on the same websites and thus compared under identical conditions. Details about the framework can be found in this paper: https://ifca.ai/fc12/pre-proceedings/paper_72.pdf.

You will be analyzing data from testing of two different password schemes: Image21 (a graphical password scheme) and Text21 (an alphanumeric password scheme). In the Text21 scheme, passwords consist of 4 random lower-case letters (a-z) and numbers (0-9). Passwords in the Image21 scheme consist of 5 random tiles, chosen from an image made up of 48 tiles. The testing data for each scheme is provided in a separate CSV file: image21.tsv and text21.tsv. Each data set involves approximately 10–20 users. Each user data involves use of 3 different passwords of the same type (e.g., text or graphical). In order to understand the format of the dataset, we first need to understand how the data was collected.

USER STUDY:

The goal of the user study was to compare the usability and security of a text-based and graphical password scheme. Usability was measured using both *quantitative* and *qualitative* measures. Quantitative measures included time-on-task (password creation and log-in time) and memorability. This project only focuses on the quantitative data, so qualitative data such as user opinions and perceptions are considered out of scope. The quantitative data was collected using log-files from the MVP framework.

Procedure: The study used a between-subjects design, where participants were randomly assigned to one of the two password conditions. Participants created and used passwords on three websites over the course of one week. The websites are part of the MVP framework and

are hosted and maintained by the researchers. They have dramatically different appearances and consist of a banking, email, and shopping website. Participants created a password on each website, and then logged into the website at different points in the study: immediately after creating the password, middle of the week, and at the end the week. The log-file collected the following information about participants' interaction with the password schemes, and these fields are included in the CSV files:

1. **time:** timestamp in the format YYYY-MM-DD HH:MM:SS
2. **user:** userid
3. **site:** website identifier
4. **scheme:** password scheme/sub-scheme identifier
5. **mode:** usage mode identifier
6. **event:** event identifier
7. **event details**
8. **data:** data associated with the password event

The password schemes are accessible via this URL: https://mvp.soft.carleton.ca/svp3008. Visit the URL to explore each scheme and familiarize yourself with how the data was logged by the MVP framework. As you interact with the password scheme, the logs created by the interaction are shown at the bottom of the page. The complete details about the study can be found in this paper: https://hotsoft.carleton.ca/~estobert/papers/soups2013_estobert.pdf. I encourage you to read the paper to fully understand the two password schemes, how data was collected and analyzed. Looking at the paper will also give you a better understanding of how the data was logged by the MVP framework.

PROJECT STEPS AND REQUIREMENTS:

Part 1: Descriptive Statistics – 40%

1. Explore each of the two password schemes at: https://mvp.soft.carleton.ca/svp3008

   Report: *Suggest the advantages and disadvantages of each scheme, and provide screen images of you using each scheme from the training page — approx. 1–2 pages.*

2. Download the password data from cuLearn (image21.csv and text21.csv). The files contain log-data in CSV format (comma separated values), and can be inspected in a text-editor or imported into a spreadsheet (e.g., Microsoft Excel or LibreOffice). Each row in the file has the fields below, and is ordered by time:
   - **time:** timestamp in the format YYYY-MM-DD HH:MM:SS
   - **user:** userid
   - **site:** website identifier
   - **scheme:** password scheme/sub-scheme identifier
   - **mode:** usage mode identifier
   - **event:** event identifier
   - **event details**
   - **data:** data associated with the password event

3. Design and implement a software program to process the log data from the two schemes, producing a **single** file in CSV format containing results ready for statistical analysis. For

this processing, you may use any programming language or environment. Think carefully about how to process the log files. In particular, you may wish to re-order the data so that you can process the information by user and by site. The new file must contain at least the following information for each user of each scheme:

- The userid.
- The password scheme (Text21 or Image21).
- The number of logins, successful logins, and failed logins.
- The time taken to enter a password, recorded separately for successful logins, and for failed logins.

Report: *Documentation for your log data processing software, including high-level explanation and pseudocode for your approach, and the documented source code. Also provide the resulting data in CSV format.*

4. Compare the usability of the two schemes, by calculating descriptive statistics and producing graphs. The following descriptive statistics and graphs should be provided:
   - Mean, standard deviation, and median of number of logins per user (total, successful, and unsuccessful).
   - Mean, standard deviation, and median of the login time per user. These should be provided for both successful and unsuccessful login attempts.
   - Histograms for the number of logins (per user, total, successful, and unsuccessful).
   - Histograms and boxplots for the login time per user, successful, and unsuccessful.

   Interpret the results from the statistical analysis, and discuss which scheme is better in terms of usability.

   Report: *The descriptive statistics, graphs, and your interpretation of them, discussing which password scheme has better usability.*

PART 2: DESIGN, IMPLEMENTATION, STATISTICAL INFERENCE – 60%

This part of the project is for you to design a new authentication system and conduct quantitative usability testing.

1. Design a new knowledge-based authentication scheme. It should assign passwords randomly, rather than let users choose them, and the password space should be approximately 21 bits. That is, there should be approximately 2^{21} possible passwords. The scheme may use characters, words, images, or any combination, or something novel. Try to leverage what you know about human memory. Even simple schemes will earn a passing mark, but extra marks will be reserved for innovative schemes.

   Report: *Design rationale explaining your scheme, why you think it might have good usability, and your calculation of the password space — approx. 1–2 pages.*

2. Implement your password scheme. You may use any programming language and environment you wish, as long as it is possible to test your scheme with users. (This will be easiest if the system can be used on the web.)

Report: *Documentation including screen images of your program in use, both for password creation and password entry, together with notes explaining the images; and the documented source code for your software for us to review (see Assignment Submission, below) — approx. 2–3 pages.*

3. Create a simple framework to assist quantitative testing of your new password scheme. Your system should assign and test 3 passwords. For each password, the system should say what it is for (e.g. "Email", or "Banking', or "Shopping"), show the user the password, and have the user enter the password to confirm they know it. When this has been done for all 3 passwords, the system should request the user enter the passwords again, in a random order. The user may be allowed up to 3 attempts to enter each password before declaring failure. Instrument your software to record details useful for assessing usability, such as login times, success, failures, etc. You should record the event details in a log file, to allow unanticipated analysis later. Make sure you save the log files!

   Report: *Documentation including screen images for your system, together with notes explaining the images; and the documented source code for your software — approx 2–3 pages.*

4. Create a simple questionnaire to investigate user's perception of the new password scheme, in comparison to normal user-chosen text passwords. It should use approximately 10–20 Likert scale questions to assess relevant perception. You should create your questionnaire using the COMP3008 survey system at: https://hotsoft.carleton.ca/comp3008limesurvey Your team will be given a userid on request.

   Report: *your survey questions in PDF (generated by the survey software), and a link to your actual questionnaire on the survey website.*

5. Conduct testing of your password scheme with users, using your testing framework and questionnaire. Your participants should be other students at Carleton, and can include your own team members and other students in COMP3008. You should recruit at least 10 participants.

   You must again obtain informed consent by participants. You should use the consent form for Project 2 provided on cuLearn. Completed consent forms must be scanned or photographed, and provided in an appendix to your report.

6. Compare results of your usability testing with the results obtained by the Text21 scheme. You should produce descriptive statistics and graphs for success and failure, and also for login time, as in Part 1. You should use inferential statistics to determine whether the results are statistically significant. You should use the results of your survey to compare perception of your scheme with a normal password scheme, and produce descriptive and inferential statistics. Interpret the results of your testing and survey, and assess your findings. Do your best to answer the questions: Is your new scheme better than TEXT21? Would users prefer it to a normal password scheme?

   Report: *presentation and interpretation of your study results, and your interpretation and discussion — approx. 2–5 pages.*

ASSIGNMENT SUBMISSION:

Throughout the period of the project, the teaching assistants will be available during their office hours for advice and feedback as you work through the project steps. For example, you may wish to seek feedback on your password design, understanding of the experiment data, or interpretation of results.

For the project report, please write up the steps of your study as described above. Please ensure that the entire report is in a single PDF document. Upload the file following the submission instructions on cuLearn.

All program source code should be helpfully documented internally. In particular, every file must begin with a block comment briefly explaining all the classes or functions within it, and their purpose. All data files should be named to make the contents clear. Your set of files should also have a file "readme.txt" that explains each of the other files and their purpose. The documentation should include sufficient detail for someone else to compile or run the code themselves. Program source code should be submitted as a .zip file.

The project is due at 5PM EDT (as indicated on cuLearn) on April 7th. Projects submitted or updated late will be penalized by a deduction of 10 marks (out of a possible 100) per 24 hour period, or part thereof.